SYNTHETIC STUDIES TOWARDS MOLECULAR RECOGNITION OF GUANOSINE QUADRUPLEXES

A Thesis

Submitted to the Graduate Faculty

In Partial Fulfillment of the Requirements

For the Degree of

Master of Science

In the Department of Chemistry

Faculty of Science

University of Prince Edward Island

Monica A. Gill

Charlottetown, P.E.I.

August, 2004

© 2004 M.A. Gill

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

National Library of Canada
Bibliothèque nationale du Canada

Acquisitions and Bibliographic Services
Acquisisitons et services bibliographiques

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 0-612-93859-X
Our file Notre référence
ISBN: 0-612-93859-X

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou aturement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this dissertation.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de ce manuscrit.

While these forms may be included in the document page count, their removal does not represent any loss of content from the dissertation.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The author has agreed that the Library, University of Prince Edward Island, may make this thesis freely available for inspection. Moreover, the author has agreed that permission for extensive copying of this thesis for scholarly purposes may be granted by the professor or professors who supervised the thesis work recorded herein or, in their absence, by the Chair of the Department or the Dean of the Faculty in which the thesis work was done. It is understood that due recognition will be given to the author of this thesis and to the University of Prince Edward Island in any use of the material in this thesis. Copying or publication or any other use of the thesis for financial gain without approval by the University of Prince Edward Island and the author's written permission is prohibited.

Requests for permission to copy or to make any other use of material in this thesis in whole or in part should be addressed to :

Chair of the Department of Chemistry

Faculty of Science

University of Prince Edward Island

Charlottetown, P.E.I.

Canada C1A 4P3

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

SIGNATURE

PAGE(S)

iii - iv

REMOVED

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Abstract

The telomere is a guanine-rich region of DNA located at its termini. Each time a DNA molecule is replicated, the telomere gets shortened. Cell death occurs after a telomere has been shortened to a critical length. The telomere may, however, be re-lengthened by the action of the enzyme telomerase. It has been found that in many cases of cancer, the level of telomerase in the cells is extremely high when compared to normal cells. The telomeric region may be linear, or may adopt higher-order structures such as the G-quadruplex. When the telomere is in G-quadruplex form, telomerase is unable to bind to the region and re-lengthen it. By designing molecules that can recognize the G-quadruplex and bind to them, therefore stabilizing them, it is theorized that telomerase inhibition may occur.

Through the use of computational methods, an oligomer has been designed to recognize the minor groove of a G-quadruplex. The oligomer has a recognition unit of an aminopyridine group and a peptide nucleic acid (PNA) backbone. The synthetic route to the monomers required for solid-phase synthesis of this oligomer is described in detail.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Acknowledgments

I would like to thank Dr. Barry Linkletter for his guidance and support during my two years of graduate work. My learning experience was greatly enhanced by always being encouraged by him to think for myself and by being able to guide the direction of the project.

I would also like to thank the members of my supervisory committee, Dr. Nola Etkin and Dr. Kevin Smith for many helpful conversations and suggestions. The 'lightning speed' at which they read my thesis was also greatly appreciated!

I am extremely grateful to Dawna Lund for her technical assistance, especially for the 2D NMR training. Thank-you also to Sharon Martin for all of her administrative assistance.

Thank-you to Laurel Murphy for assistance in the lab in October, 2003 and thanks also to Wade White for staying with me during evenings and coming in on weekends during the last weeks of my lab work.

I truly appreciate the time that Dr. Robert Hudson from the University of Western Ontario took out at the CSC/IUPAC Conference in August, 2003 to talk to me with respect to my research. His suggestions were very helpful in directing the course of my research.

Finally, I would like to thank my family and friends, in particular, my mother and sister Angela, whose tireless support helped me reach my goals.

Thank-you!

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table of Contents

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

List of Figures

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

List of Tables

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

List of Abbreviations

AM	aminomethyl polystyrene
BHA	benzhydrylamine
Boc	*tert*-butoxy carbonyl
Boc_2O	di-*tert*-butoxy dicarbonate
n-BuLi	n-butyl lithium
Cbz	benzyl carbamate
CIC	N,N'-cyclohexyl isopropyl carbodiimide
Cy	cyclohexyl
DCC	N,N'-dicyclohexyl carbodiimide
DCU	N,N'-dicyclohexyl urea
DIC	N,N'-diisopropyl carbodiimide
DMAP	4-(N,N-dimethyl)aminopyridine
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
EtOAc	ethyl acetate
GMP	guanosine monophosphate
HF	Hartree-Fock
HMBA	4-hydroxymethyl benzoic acid polystyrene
HPLC	high-performance liquid chromatography
M	molar
MeOH	methanol
MHz	megahertz

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

MMFF	Merck molecular force field
MMTr	monomethoxy trityl
mRNA	messenger ribonucleic acid
NBS	N-bromosuccinimide
NMR	nuclear magnetic resonance
NOE	Nuclear Overhauser Effect
NOESY	Nuclear Overhauser Effect spectroscopy
PNA	peptide nucleic acid
RNA	ribonucleic acid
T_m	melting temperature
TFA	trifluoroacetic acid
THF	tetrahydrafuran
TLC	thin-layer chromatography
UV	ultraviolet

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1. Introduction

1.1 Molecular Recognition

Molecular recognition can be broadly defined as a specific inter- or intramolecular interaction involving only non-covalent bonds, such as hydrogen bonds, van der Waals forces and dipole-dipole interactions.[1] One of the most well-known examples of molecular recognition is the Watson-Crick base pairing of guanine to cytosine and of adenine to thymine. This recognition is specific in that guanine and cytosine recognize each other only. The same is true for adenine and thymine, as illustrated in Figure 1.

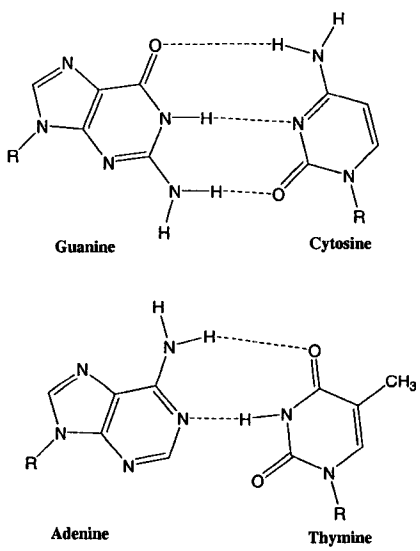


Figure 1: Watson-Crick Base Pairing of Nucleotides

1.1.1 Molecular Recognition in Hydrophobic Cavities

There has been extensive research done on molecular recognition by molecules with hydrophobic cavities, such as cyclodextrins (see Figure 2), which are composed of cyclic oligomers of glucose. This branch of chemistry is known commonly as "host-guest chemistry." There are three main types of cyclodextrins, each with a particular number of glucose units. α-Cyclodextrin has six glucose units, β-cyclodextrin has seven, while γ-

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

cyclodextrin has eight. Both the primary and secondary hydroxyl groups point away from the center of the cavity, therefore creating a cavity that is non-polar relative to the outside of the molecule.[2] In aqueous medium, a relatively non-polar molecule could exclude water molecules from the cavity and interact with the inside of the cavity in a non-covalent fashion.

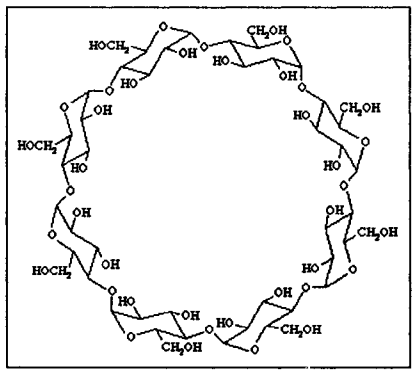


Figure 2: Structure of γ-cyclodextrin

1.1.2 Molecular Recognition by Hydrogen Bonding & π-stacking

Rebek was one of the pioneers in studying molecular recognition. He designed many synthetic receptors that would be used to study the mechanics of molecular recognition. One very elegant model he designed is shown in Figure 3.

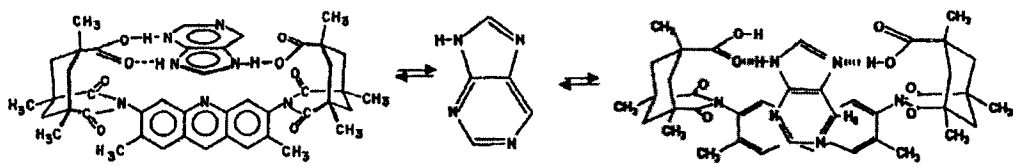


Figure 3: Comparison of Purely Hydrogen Bond Recognition with Combination of Hydrogen Bond and π-stacking Recognition (from reference 3)

He designed 'cleft'-like molecules which had various hydrogen-bond donors and acceptors which could recognize a variety of heteroaromatic molecules.[3] In the case shown, recognition of the purine molecule can be achieved by purely hydrogen bonding,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

as shown with the equilibrium to the left, or by a combination of hydrogen bonding and π-stacking, as shown by the equilibrium to the right. The interaction of the aromatic rings constitutes another form of molecular recognition.

1.2 Antisense & Antigene

Antisense agents are molecules that have been designed to bind to specific mRNA sequences in order to prevent translation to proteins, which exerts a desired therapeutic effect. Antigene agents are quite similar to antisense agents in their action, however their target is DNA sequences in the genome. Therefore, replication of the strand is inhibited. It was in 1978 that Zamecnik and Stephenson first proposed the use of synthetic antisense oligonucleotides for therapeutic purposes.[4] Using an oligonucleotide with 13 repeats (termed a '13-mer') that was complementary to a specific mRNA sequence, they were able to inhibit the growth of the Rous sarcoma virus *in vitro*. The inhibition was based on Watson-Crick base pairing between the nucleotide bases of the antisense oligonucleotide and the viral nucleic acid.

A second method by which antisense oligomers can cause growth inhibition is through activation of an enzyme known as RNase H.[5] As can be seen in Figure 4,

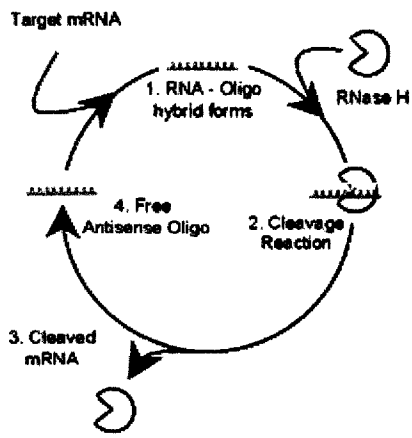


Figure 4: Mechanism of RNaseH Action

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

an antisense oligomer can form a duplex structure with RNA, which activates the enzyme. RNase H actually cleaves the mRNA strand, releasing the antisense oligomer to form duplex complexes with more mRNA. Additionally, once activated, the RNase H can actually stay active, thereby increasing potential degradation.

A great deal of potential was seen in this new technology. However, in its beginning stages, antisense and antigene therapies were not useful *in vivo*. The oligonucleotide used to inhibit the Rous sarcoma virus used a deoxyribose phosphodiester backbone, i.e. a DNA backbone. This type of oligonucleotide would face several challenges if it tried to enter a human cell. First of all, the DNA backbone is highly negatively charged and would be repelled by the phospholipid bilayer of the cell membrane. Second, if the oligonucleotide were able to enter the cell, it would be quickly degraded by enzymes called nucleases designed to seek out and destroy broken DNA strands. In order to further antisense and antigene technology, researchers turned to modified backbones in an attempt to resolve these difficulties.

1.3 Modified Backbones

1.3.1 Studying Melting Temperature Behavior

To study a modified backbone oligomer's affinity for RNA, the technique usually used is the study of melting temperatures. At lower temperatures, DNA and RNA strands are annealed to one another. The planar stacked bases of paired, or duplex, nucleic acids causes a π-π interaction between the aromatic rings. This interaction quenches the absorbance of UV light at 260nm. When the annealed strands are heated, they dissociate and become single stranded, thus destroying the π-π interaction and increasing the absorbance at 260nm. Under ideal conditions, a plot of absorbance *versus* temperature

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

results in an S-shaped curve. The midpoint of this curve is defined as T_m. This value denotes the temperature at which equal fractions of duplex and single-stranded nucleic acids exist at an equilibrium. T_m can be used to determine thermodynamic parameters of the system. Therefore, if it is stated that a modified backbone oligomer has increased affinity to RNA when compared to DNA, that indicates that the T_m value was higher for the modified oligomer/RNA combination than for the DNA/RNA combination standard.[6]

1.3.2 First-Generation Modified Backbones

There has been a great deal of research effort devoted to the synthesis of DNA backbone analogue oligonucleotides. Normal DNA has negatively charged

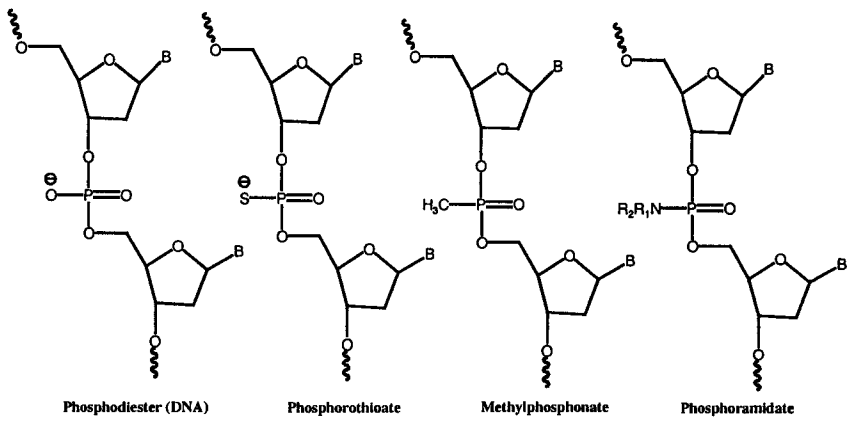


Figure 5: First Generation of Backbone Modified Oligonucleotides

phosphodiester linkages. The first generation modified backbones kept the phosphorus atom, but changed one or both of the attached oxygen atoms. These included changes to sulfur, methyl, and amines (see Figure 5). [7]

Each of the first generation modified backbone oligomers showed increased resistance to nucleases when compared to DNA backbones, however they showed decreased affinity for RNA.[8] It was theorized that this decrease in affinity arose from the fact that substitution on the phosphorus atom with a non-oxygen atom would result in the elimination of a plane of symmetry. This results in a mixture of diastereomers. In a 20-

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

mer, for example, there would be over a *half-million* individual stereoisomers. Therefore, there may have been some stereoisomers that had a higher affinity for mRNA than DNA, however this fact would have been completely overshadowed by the lack of affinity the other several hundred thousand stereoisomers had.[7]

1.3.2.1 Phosphorothioates

Phosphorothioates were one of the first backbone modifications made. In fact, they are still widely used. In these backbone anlogues, one of the non-bridging oxygen atoms is replaced by a sulfur atom. The corresponding RNA duplex structure has a slightly decreased T_m when compared to DNA. Interestingly, there is a dramatic difference between the T_m values for a modified backbone oligomer with all *R* configuration compared to all *S* configuration. The R_p configuration causes a great deal more destabilization (T_m = 16.6°C) than does the S_p configuration (T_m = 33.5°C). This vast difference in T_m's is attributed to the steric and electronic differences between the two enantiomers. The S_p isomer has the negatively charged sulfur directed away from the center of the double helix, thereby by-passing repulsion of both a steric and electronic nature.

Phosphorothioate oligonucleotides have increased cellular uptake in comparison to phosphodiester oligonucleotides, but this depends highly on cell type. In addition, this type of backbone modification, along with phosphodiesters and phosphoro*di*thioates, are the only type of backbone modified oligomers that are capable of activating RNase H.

When both non-bridging oxygen atoms of a phosphodiester linkage are replaced with sulfur, the result is non-chiral phosphoro*di*thioate. The duplex stability of the phosphorodithioate is less when compared to phosphorothioate oligonucleotides and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

DNA. However, phosphorodithioates have been shown to demonstrate less non-specific protein binding and also have shown more resistance to nucleases. Reaction to these results has been somewhat mixed. Most researchers felt that there was little advantage in using the disubstituted backbone, while others felt that phosphoro*di*thioates had potential for development.[8]

There have also been phosphorothioates prepared where the *bridging* oxygen of the phophodiester linkage at the 3' or the 5' end has been replaced by a sulfur atom. Phosphorothioates of this type are quite stable to nucleases, however their main use has been to initiate site-directed strand cleavage. These oligomers are easily cleaved with iodine, silver and mercury ions.

1.3.2.2 Methylphosphonates

Methylphosphonates were the first modified backbone oligomers to be neutral. This lack of charge facilitates greater ease in penetration of the cellular membrane. In fact, it is believed that the oligomer is taken into the cell by passive diffusion. For methylphosphonate oligomers with all thymine bases, there is only a 3°C difference between the T_m's of the S_p and the R_p enantiomers. With other base combinations on a methylphosphonate backbone, the duplex stability appears to be independent of stereochemistry; the T_m values are very similar to DNA/RNA duplexes.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1.3.2.3 Phosphoramidates

Phosphoramidates stand out in the first-generation of backbone modifications due to their high resistance to hydrolysis by nucleases. Generally, the duplexes formed between phosphoramidate oligomers and RNA are weak. It has been shown that an oligonucleotide with alternating phosphodiester and phosphoramidate linkages (as shown in Figure 6) shows similar duplex stability to standard DNA/RNA interactions. These

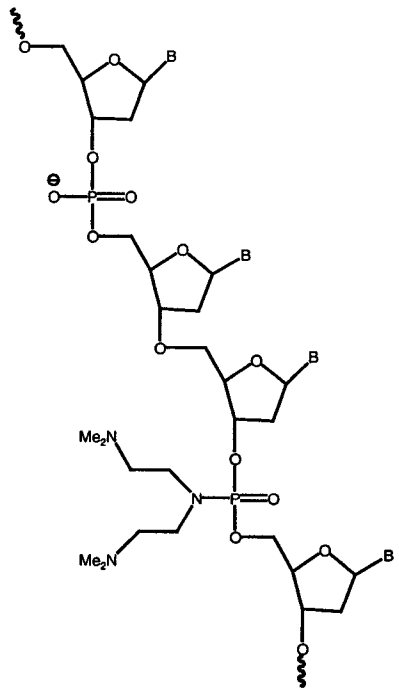


Figure 6: Oligomer with Alternating Phosphodiester & Phosphoramidate Linkages

alternating oligomers are termed 'chimeras.' It is theorized that the non-alternating phosphoramidate oligomer duplex with RNA would be forced to adopt unfavorable conformations, whereas the alternating phosphoramidate would not.

Of interest is the fact that the T_m of the alternating backbone has been found to be independent of ionic concentration. In most instances, the T_m of a duplex system increases as the ionic strength of the surroundings increases. This is due to the negatively charged phophodiester backbone. This negative charge is dampened somewhat by the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

presence of positively charged metal ions, such as sodium and potassium. The partial relief of the charge on the backbone diminishes some of the natural repulsion of like charges.

Phosphoramidates with either the 3' or the 5' oxygen replaced by a nitrogen atom have also been synthesized. When a 10 unit DNA strand with one single phosphoroamidate linkage (with the 5' oxygen replaced) was combined with poly(dA), there was almost no observable decrease in stability when compared to DNA/RNA duplexes. When several phosphoroamidate linkages of this type were incorporated, a large decrease in T_m was observed with ΔT_m= -1.0 to -4.0°C. When a 10 unit oligonucleotide with the 3' oxygen replaced by a nitrogen, there was a great increase in T_m (+3°C). Oligonucleotides with alternating phosphodiester and phosphoramidates linkages with the 3' oxygen replaced have been found to have very little effect on duplex stability.[8]

1.3.3 Second-Generation Modified Backbones

Emerging in the mid-1990's were the second generation modified backbones. These were distinct from the first generation in that they were phosphorus-free. The phosphorus atom had been replaced by a variety of groups, including oxygen, sulfur, methylene, silicon and amines. Excitingly, these oligomers demonstrated greatly increased resistance to nucleases when compared to DNA backbones. Most, however, actually demonstrated a decreased affinity for RNA targets.

Moser's research group[7] undertook a study to examine the effects of completely replacing the phosphodiester linkage in the DNA backbone. The linkage was replaced with amides, ureas, carbamates and alkyls, as shown in Figures 7 and 8.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

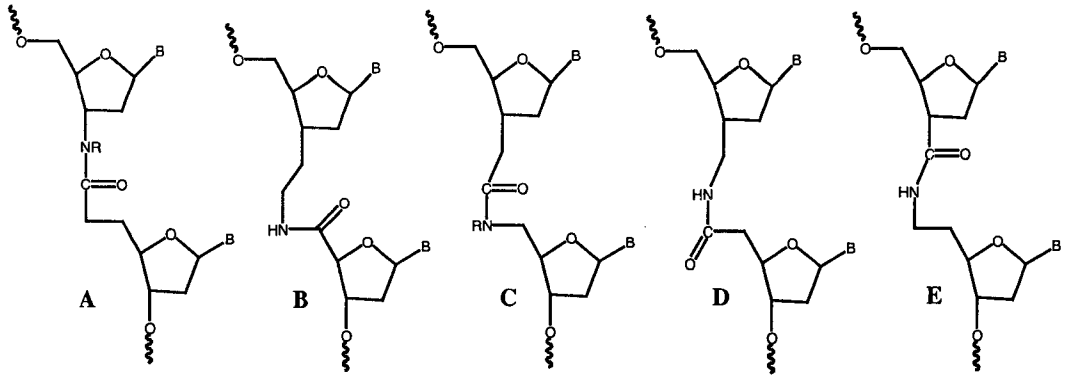


Figure 7: Synthetic oligonucleotides with Amide Linkages as replacements for Phosphodiester Linkages

Amides were chosen as one of the phosphodiester linkage replacements due to the simple and well-known syntheses of amides. Also, amide linkages are achiral and thus there would be no problems with mixtures of diastereomers. In addition, it was thought that due to the neutral nature of an oligonucleotide with an amide linkage that it would be more likely to be able to cross the cell membrane.

The first amide depicted in Figure 7, (A), induced a shift of -3.2°C in T_m, a highly significant drop in duplex stability. Combinations of the first (A) and fifth (E) amides in Figure 7 each with RNA result in similar T_m's. These modified backbones adopted similar geometries when in duplex form. It was theorized by the authors that the restricted rotation of the amide adjacent to the 3' end of the deoxyribose sugar may have a destabilizing effect. If the rotationally restricted amide linkage is adjacent to the 5' end, as in the second amide in the above series, a lesser destabilization is observed (ΔT_m = -2.3°C) The third (C) and fourth (D) amides in the above series likely adopt similar geometries as demonstrated through their similar ΔT_m's with a mean value of +0.15°C. Through this study of amide modified backbones, it was concluded that the crucial factor in determining affinity to an RNA target was geometry of the backbone.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Hybridization properties of synthetic oligonucleotides to RNA targets depend on conformational flexibility and also structural preorganization. Moser's group also investigated linkages that imparted greater rigidity to the backbone, such as ureas and carbamates (see Figure 8). This rigidity is caused by conjugation of an additional heteroatom with the amide bond. When the nitrogen of the urea group adjacent to the 5' end of the sugar moiety was substituted with a methyl, there was a dramatic decrease

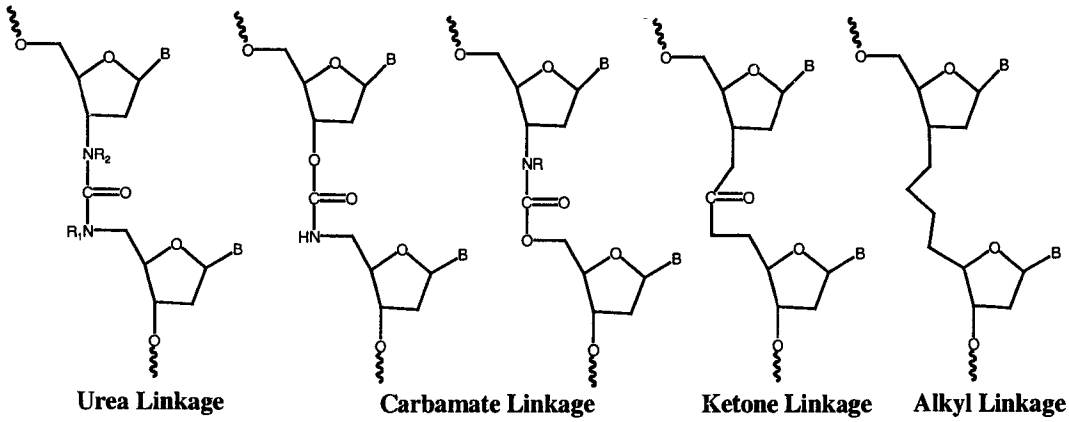


Figure 8: Other Second Generation Backbone Modifications

in the corresponding oligomers' ability to form duplex structures when compared to DNA (ΔT_m = -4.4°C). Carbamates were found to have similar thermal stability. To determine the effect of a lack of rigidity, flexible four carbon linkages were also studied. Not surprisingly, the lack of structural pre-organization of the modified backbone oligomer caused a decrease in thermal stability of the duplex.

Taking into consideration all of these experiments, it was concluded by Moser's group that the "restriction of rotation has to occur preferentially in the middle of the four-atom backbone, with a preferred torsion angle of approximately 180°." [7] With this conclusion in hand, researchers believed that they could fine tune the distance between sugar units and also the rigidity of the backbone to find the optimal degree of preorganization.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1.4 Peptide Nucleic Acid (PNA)

1.4.1 Historical Perspective

In 1991, Nielsen and co-workers introduced a completely new type of modified backbone in which the entire phosphodiester-sugar backbone was replaced. This backbone was the spawn of the Nielsen group's desire to synthesize reagents that bind in a sequence specific manner to double-stranded DNA. Phosphodiester sugar backbones are not suitable for these types of reagents because it is difficult to synthesize these oligonucleotides on a large scale (in the millimole to mole range) and also because it is difficult to introduce modified bases or to conjugate the backbone to other ligands. Simply put, there was little opportunity for diversity in the synthesis of DNA oligonucleotides. [9,10]

It was envisioned that the deoxyribose phosphate backbone could be replaced with a polyamide backbone. The number of backbone bonds would remain the same, as would the distance between the backbone and the base. Starting with a computer model of a T:A-T DNA triplex, the deoxyribose phosphate backbone of the third, triplex forming strand was removed. A polyamide backbone was constructed in its place so as to fit the empty space. The new backbone consisted of repeating units of (2-aminoethyl)glycine as shown in Figure 9.

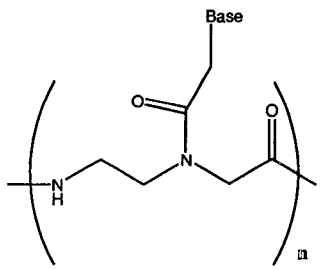


Figure 9: General Structure for a PNA Oligomer

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Nielsen's group has stated that they chose to call this type of backbone PNA because "it was a chimera between a peptide and a nucleic acid".[6] The PNA backbone is quite different from that of DNA. The primary difference is, of course, the fact that the PNA backbone is neutral. Also, PNAs are achiral and not susceptible to enzymatic cleavage.[11]

Two seemingly unbelievable observations were made of the interaction of PNA and DNA. Firstly, PNA oligomers showed an incredibly high binding affinity for single-stranded DNA. The $\Delta T_m = +50°C$ when comparing a T_{10}-PNA interaction with deoxyribose phosphate backbone A_{10}! T_{10}-PNA's interaction with double-stranded DNA was a second intriguing observation. A ten base-pair section of thymine/adenine Watson-Crick pairs within a 248 base pair double stranded DNA fragment could actually be invaded by T_{10}-PNA. Instead of forming the usual triplex, the PNA oligomer would actually displace the T_{10} DNA strand.

These observations caused a great deal of excitement and speculation amongst the antisense community, despite the fact that the strand displacement mechanism only occurred at much lower salt concentrations than are found in the human body. One of the functions of cations such as Na^+ in the body is to stabilize the negatively charged phosphodiester backbone of DNA. Since both strands are negatively charged, there is a certain degree of repulsion that occurs between the strands. The presence of cations relieves some of the negative charge on the strands, allowing them to have a higher affinity for one another. The fact that this PNA could displace a DNA strand was therefore not completely surprising considering the concentration of salt in the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

experiment. There would have been a significant amount of strand repulsion occurring already.

PNA oligomers with all pyrimidine bases have been found to bind to DNA in a 2:1 PNA:DNA ratio, to form triplexes. Within this triplex, the first PNA strand can bind in either a parallel (with N-terminus of PNA at the 5' end) or antiparallel (with N-terminus at the 3' end) manner. The stability of the triplex does not appear to be affected by the orientation of the first PNA strand. To date, the orientation of the second PNA strand in the triplex has not been determined. When PNA sequences with mixtures of nucleobases are used, they bind in a complementary fashion *via* Watson-Crick base pairing to single stranded DNA and RNA in a 1:1 ratio. Again, the duplex stability is much greater than that of the corresponding deoxribose phosphate backbone duplex. [8]

A number of PNA complexes have had their structures solved. The duplex structures of a PNA-RNA and a PNA-DNA have been solved by use of NMR, while a PNA_2-DNA triplex and a PNA-PNA duplex have been solved through X-ray crystallography. Careful examination of these solved structures shows that PNA does have sufficient flexibility to adopt to the A- and B-form helices preferred by DNA and RNA.[12] While PNA *can* adopt the A- and B-form helices, it actually prefers to form a much wider (28Å) helix and with a much larger pitch (18 base pairs). The major conformation of DNA in solution is B-DNA; its helix has a width of 23.7Å with a pitch of 10 base pairs.[13]

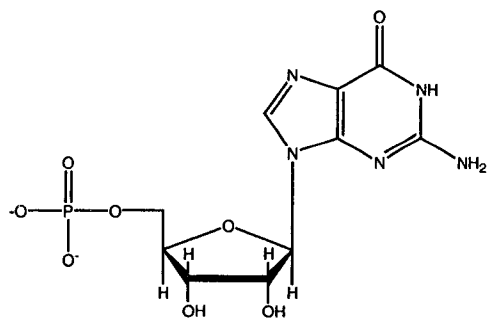
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1.5 Quadruplexes

1.5.1 Historical Perspective

1.5.1.1 X-ray Diffraction Studies

It has been known since early in the twentieth century that concentrated solutions of guanosine monophosphate (GMP) (see Figure 10) formed gels. It had been long proposed that the GMP molecules must form some sort of higher order structure. This structure was finally deduced in the early 1960's by Gellert and co-workers[14] through the use of UV absorption studies, examination of optical density and rotation of the gels and also through X-ray diffraction studies. The X-ray diffraction pattern obtained for 5'-GMP was characteristic of a helix. From the data, it was inferred that there were four units per turn of the helix. It was found that there was 3.36 Å between bases, and the helix repeat was 3.25Å.



5'-GMP

Figure 10: Structure of 5'-guanosine monophosphate

Prior to this work, it had been noted by Donohue[15] that there are four dimers that could conceivably form between two guanines. While guanine dimers would make a contribution to the overall stability of a system, it was not believed that this was sufficient to cause the great stability observed in the guanosine gels. It was noted, however, that if one particular dimer reported by Donohue paired with itself, they would associate

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

through two hydrogen bonds per guanine base, for a total of eight hydrogen bonds (as shown in Figure 11). Due to the high number of H-bonds in this proposed system, one could expect a great deal of stability.

Gellert *et al.* further hypothesized that due to the high potential for van der Waals attraction, this planar arrangement of guanine bases could form higher order structures by stacking on top of each other. They also reasoned that a structure composed of stacked tetramers would be approximately cylindrical and that the cavity in the center of the structure could likely encapsulate one small molecule, such as water, for every tetramer in the stack.

Armed with these hypotheses, the X-ray diffraction data were re-examined. It was realized that the data could be interpreted as stacked tetramers rather than a helix. The difference between the inter-base distance and the helix repeat distance was resolved quite simply by tilting the bases slightly. In fact, it was reasoned that a helix

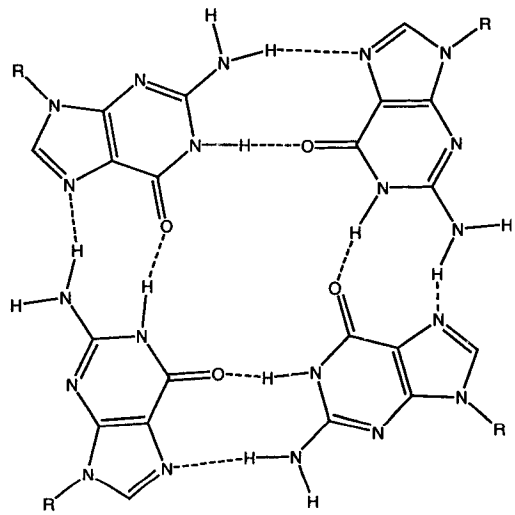


Figure 11: Predicted Higher Order Structure of 5'-GMP (R = sugar phosphodiester)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

with right-handed rotation of 22.5° would be impossible due to steric repulsion between adjacent ribose-phosphate groups.

1.5.1.2 Early NMR Evidence for Tetramer

The first reports of NMR evidence supporting the proposed tetrameric structure appeared in 1975.[16] Prior to this report, all nucleotide complexes, including those formed from complementary association, underwent rapid chemical exchange and only gave time-averaged NMR spectra. Pinnavaia and co-workers presented data demonstrating slow chemical exchange for self-assembled 5'-GMP in neutral solution. This observation was unprecedented at the time. Figure 12 illustrates the numbering convention used when describing atom position in the guanine unit.

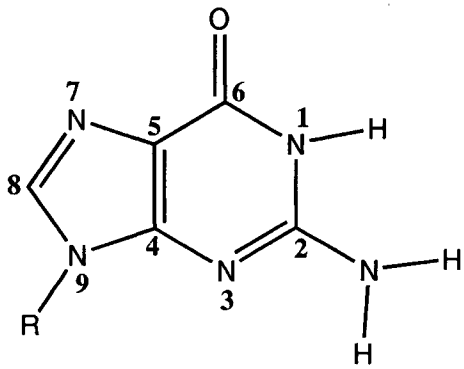


Figure 12: Numbering Conventions in Guanine

Solutions of sodium salts of 5'-GMP in D_2O were prepared at various concentrations and the ^1H NMR spectrum was collected. They found that at low (0.23 M) concentrations of Na_2[5'-GMP], there existed one single sharp peak at δ 8.17 ppm for the proton attached to C(8), which is termed H(8). This was consistent with the spectrum for a disordered dilute aqueous nucleotide system. However, as the concentration of Na_2[5'-GMP] was increased, the peak at δ 8.17 ppm broadened and three new peaks due to H(8) appeared in the region, for a total of four peaks.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The assignment of these peaks to H(8) was verified by replacing only the H(8) protons with deuterium, and not the protons in the ribose sugar. The intensities of the peaks decreases after exchange with D_2O. A similar development of a set of broad peaks from a single sharp peak exists when the NMR experiment is performed by varying temperature instead of salt concentration. At 33.5°C, there is a single peak, whereas at 1°C, there are four broad peaks.

These data provided even more evidence for the existence of a regular and ordered structure that can be formed by guanine monomers. The existence of four H(8) peaks in the ^1H NMR at high salt concentrations and low temperature was indicative of four guanines associating in a repeating structure. This report also began to illustrate the importance of salt ions for the formation of these higher order structures.

1.5.1.3 First Evidence of Metal Ion Specificity

In 1978, there came a publication that reported dramatic results of selective alkali metal ion complexation by the tetrameric structure. Pinnavaia and co-workers found that the ability of 5'-GMP to self-assemble depended on the nature of the metal ion. They claimed that this was the first example of "the ability of alkali metal ions to direct structure formation of a nucleotide through a size-selective coordination mechanism."[17] Li^+, K^+, Na^+, Cs^+ and Rb^+ were included in this study. In dilute (approximately 0.02 M) solution all of the salts exhibit the standard single sharp peak for H(8) at 8.17 ppm (see Figure 13).

When the study was conducted at concentrations nearly at saturation, however, there were interesting results. In the spectra collected with lithium and cesium salts in solution with 5'-GMP, the peak shifted slightly upfield and broadened somewhat.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Importantly, though, there was still only one peak. The sodium complexes exhibit four sharp resonances, as previously described.

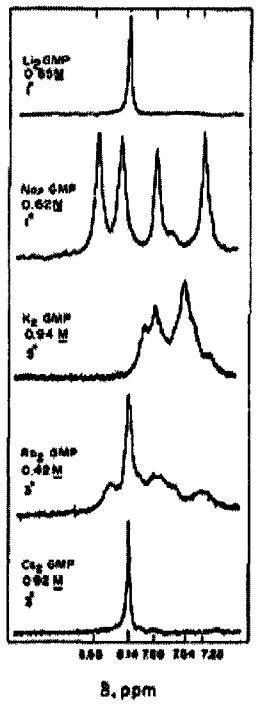


Figure 13: ^1H NMR Spectra (aromatic region) of Alkali Metal Salts of 5'GMP (from reference 17)

The spectrum collected for K_2[5'-GMP] is quite different than that of Na_2[5'-GMP]. Concentration and temperature dependence studies were performed on K_2[5'-GMP] and the presence of at least three different structures is detected. Of special note is the apparent higher stability of the tetrameric structure with K^+ ions as opposed to Na^+ ions. At 53°C for the K^+ solutions, there are still three apparent structures visible in the NMR spectrum, while at only 33.5°C for the Na^+ solutions, there is a single peak visible.

1.5.2 Structural Features of the Quadruplex

Early studies of G-quadruplexes were limited to monomeric GMP. As the study of DNA became more sophisticated, and genes began to be sequenced, it was found that a

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

number of regions of DNA contain regions rich in guanine repeats. It was theorized that DNA strands with repeating guanine nucleotide bases would be able to form intermolecular and intramolecular G-quadruplex structures.

1.5.2.1 Alignment of Strands

There are four common motifs that G-quadruplexes adopt. Four strands may associate in a parallel fashion, with all 3' termini at one end and all 5' termini at the other end. In the case of a parallel quadruplex, the alkali metal ion may be between two quadruplexes or may be coplanar with the quadruplex. There are two ways in

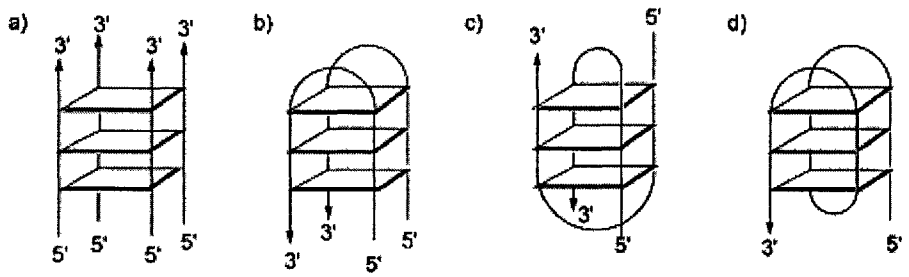


Figure 14: Possible Strand Orientations for G-Quadruplex (from reference 18)

which two strands could associate to form a quadruplex. A strand with stretches of G-repeats can fold back on itself to form what is called a 'hairpin' structure. Two hairpins may dimerize with the two loops adjacent to one another, which is termed 'edgewise loops', or two hairpins may interact such that the loop of one hairpin is located between the termini of the other hairpin. This is termed "diagonal loops." Finally, a single DNA strand with multiple G-repeats may fold on itself to form a unimolecular G-quadruplex. [18]

Between each adjacent strand of a G-quadruplex, a groove is formed. These grooves are analogous to the major and minor groove observed in DNA. If the strands are all parallel, the grooves will be equal in size. However, with the other alignments of strands that are possible, a major and minor groove will occur.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Smith and Feigon [19] reported that the repeating unit in *Oxytricha* $d(G_4T_4G_4)$ could adopt different quadruplex structures based on the alkali metal ion present. NMR analysis showed that in the presence of Na^+, a G-quadruplex with diagonal loops was formed. However, in the presence of K^+, a G-quadruplex with edgewise loops was formed.

1.6 Telomeres

1.6.1 Definition

Telomeres are DNA and protein structures that form complexes at the ends of linear chromosomes. The 3' end of DNA is single-stranded and is therefore susceptible to action of exonucleases and can be recognized incorrectly as damaged DNA. Therefore, in order to protect the ends, they are capped with telomeres. Chromosomes are also anchored to the nuclear envelope through telomeres.[20] In most eukaryotes, the DNA component of telomeres is composed of repeating units of 5-26 base pairs. For example, the human telomere repeating sequence is 5'-TTAGGG-3'. These repeating units serve as binding sites for proteins. A suite of proteins cap directly to the telomeric DNA, thus coating and protecting the region. [21]

Shortening in length of the telomere is a natural consequence of the 'end-replication problem.' The lagging strand in DNA replication always loses a number of base pairs at its termini, because polymerases act only in the 5' to 3' direction. Polymerases must bind to the DNA strand several base pairs down in the 5' direction in order to synthesize the Okazaki fragments on the lagging strand. At the terminus of the strand, there would be no extra base pairs at the 5' end for polymerase to bind to and finish the synthesis of the last few bases. Therefore, there will always be an incomplete end. [13]

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1.6.2 Telomerase

Telomeres can be relengthened by the action of the enzyme telomerase. Human telomerase consists of the catalytic subunit human telomerase reverse transcriptase (hTERT) and the human telomerase RNA template (hTR) which contains a sequence complementary to the telomeric repeat. As well, there are the accessory proteins Hsp90, p23 and TEP-1. [22] Telomerase is found in higher concentration in cancer cells than in normal cells. A normal cell dies after a certain number divisions. It is theorized that telomerase disrupts the normal cell cycle and allows for over-proliferation of cells, leading to a cancerous tumor.

1.6.3 Quadruplex Telomeric Structures

Due to their rich G-content, telomeric regions are known to form quadruplexes. So far, this has been found *in vitro* only, however molecules that are known to stabilize the G-quadruplex also inhibit telomerase. When a telomere adopts the quadruplex formation, the enzyme telomerase is not able to bind to the telomere and re-lengthen it. The telomere must be in the linear form for telomerase binding to occur. There have also been a number of proteins identified that bind to the G-quadruplex structure. In fact, a number of these proteins actually promote the formation of the quadruplex. [22]

1.6.4 G-Quadruplex as a Therapeutic Target

An enzyme could be deactivated in two distinct fashions. The enzyme itself could be deactivated, or the target of the enzyme could be altered somehow so that recognition would not occur. With the knowledge that telomerase is inactive when a telomere is in the quadruplex form, researchers began to theorize that synthetic molecules designed to promote and stabilize the formation of the quadruplex. It was believed that these

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

molecules could be of great therapeutic potential. By inhibiting telomerase, it is believed that cancer cells would be able to undergo natural cell death. This technology could be of great anti-cancer potential. Inhibition of telomerase by inactivation of the enzyme has also been studied extensively, however this discussion will be limited to the G-quadruplex as the target.

1.6.4.1 Anthraquinones

The first report of a molecule capable of recognizing the G-quadruplex came from Neidle *et al.* in 1997. [23,24] This lab began a structure-based search for non-nucleoside compounds that could selectively inhibit telomerase by recognition of the G-quadruplex. This group had previously synthesized a series of 1,4- and 2,6-diamidoanthraquinones that were found to interact selectively with DNA triplexes.[25] Molecular modeling studies showed that this group of molecules might also interact with G-quadruplexes. A "threading" intercalation mode was proposed for this interaction, where both the major and minor groove of the quadruplex are occupied by the two side chains of the disubstituted anthraquinones. [26] The figure shown below gives the general structure of the 2,6-diamidoanthraquinones that were studied. The chain length was varied between n=1 or 2, and the amino group was also varied.

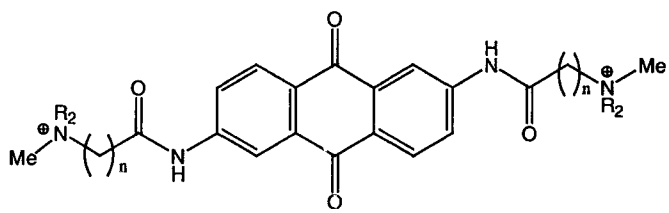


Figure 15: General Structure of 2,6-disubstituted Anthraquinones

Figure 16 illustrates a representative synthesis of a diamidoanthraquinone with substitution, in this case, at the 2- and 6- position. [25]

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

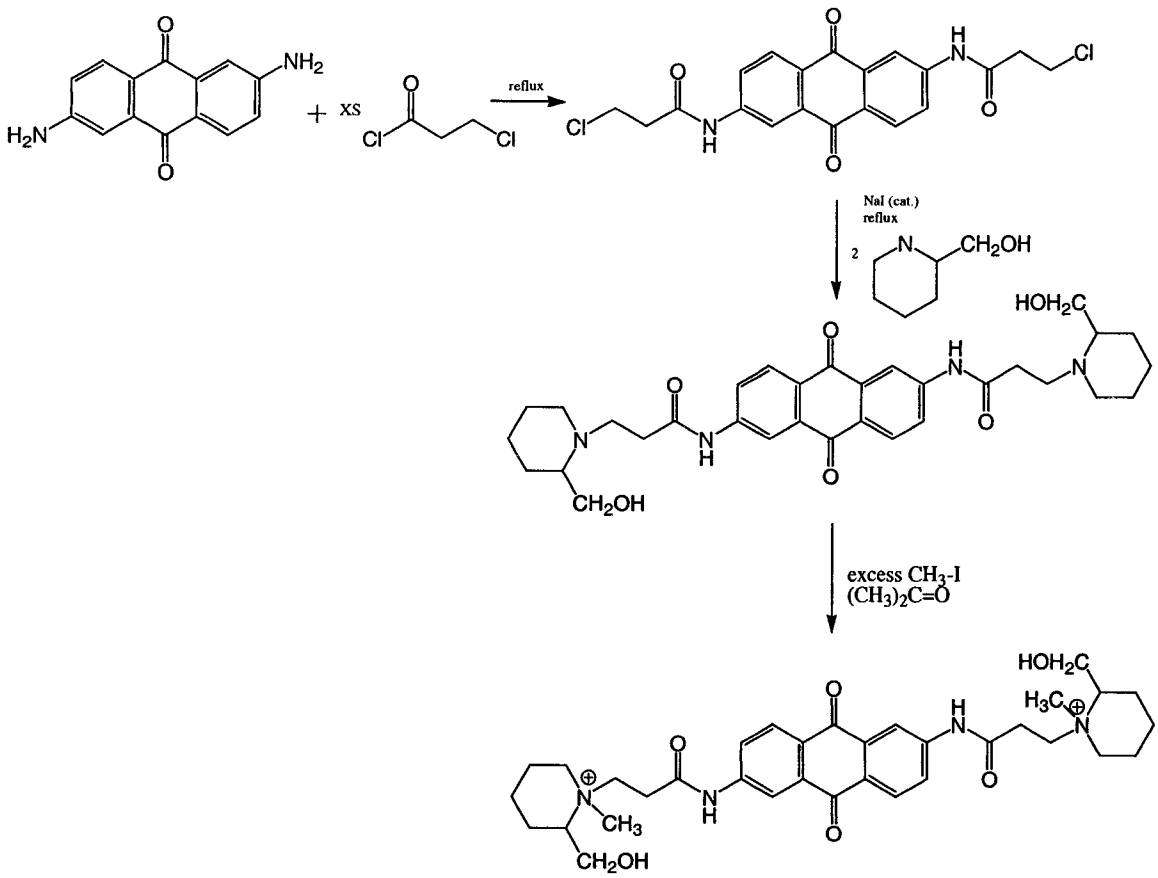


Figure 16: Synthesis of a di-substituted Anthraquinone

One year after their first report, the Neidle group reported synthesis of and telomerase inhibition by additional regioisomeric disubstituted anthraquinones. Substitution at the 1,5-, 1,8- and 1,7- positions were reported with variation of chain lengths and amido group. [27]

By studying the ability of these amido-anthraquinones to inhibit telomerase activity, a set of structure-activity relationships were developed.

(1) Two side chains attached to the amido-anthraquinone are required for activity

(2) Optimally, there is a chain of two carbons between amide carbonyl and amine nitrogen

(3) The amido group is essential

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(4) For optimal activity, the terminal amine group should be piperidine or pyrrolidine

(5) The cationic charge on the terminal amine is essential

(6) The substitution pattern on the anthraquinone ring system is not crucial

(7) As telomerase activity decreases, there is not a significant increase in cytotoxicity.

The optimal terminal amine was found to be either piperidine or pyrrolidine. Interestingly, since the pK_a of these amines is greater than physiological pH, they are protonated *in vivo*. It is not necessary to pre-form the ammonium salt in these cases. The substituent which performed the worst was the morpholino substituent, which would not be protonated at physiological pH. It is inferred that the positive charge on the side chains alleviates somewhat the negative electrostatic potential focused at the center of the quadruplex.

1.6.4.2 Acridines

With the aim of improving telomerase inhibition, Neidle *et al.* [28] designed a ligand analogous to the 2,6-diamido anthraquinone ligand. These new ligands were 'acridine' based and substituted at the 3- and 6- position. Molecular modeling

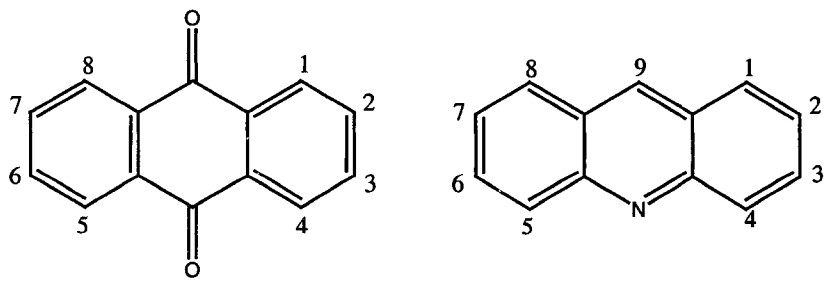


Figure 17: Comparison of Anthraquinone vs. Acridine System

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

studies predicted the affinity of the acridine ligands to the G-quadruplex would be at least as great as that of the anthraquinone ligands. The nitrogen heteroatom also provided another site for protonation, which was predicted to increase water solubility when compared to anthraquinones. Protonation of the heteroatom would also decrease electron density on the chromophore, therefore allowing for better interaction with the G-quadruplex.

The starting material for the synthesis of 3,6-substituted acridines is proflavine, which is simply 3,6-diamino acridine. Through an analogous synthesis to that of bis-substituted anthraquinones, disubstituted acridines were easily synthesized. However, when it was envisioned that 3,6,9-substituted acridines would be even better at recognizing the G-quadruplex, a slightly different strategy was necessary. The challenge was to functionalize the 9-position in the acridine ring.

Diphenylmethane was used as starting material in this synthesis, as outlined in Figure 18.[29] The alkyl linker acts as an *ortho/para* director, thus aromatic electrophilic nitration conditions result in four nitro groups in the 2- and 4- position of each ring. Oxidation creates the carbonyl linker between the two aromatic rings. Reduction of the nitro groups yields an intermediate with four amino groups, which in the presence of acid, quickly reacts intramolecularly to yield the tricyclic product. Addition of 3-chloropropionyl chloride and amination of the alkyl chloride is reminiscent of the synthesis of the disubstituted anthraquinones. Phosphorus oxychloride converts the carbonyl linker to a chloride which allows for the amination of the 9-position in the acridine ring.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

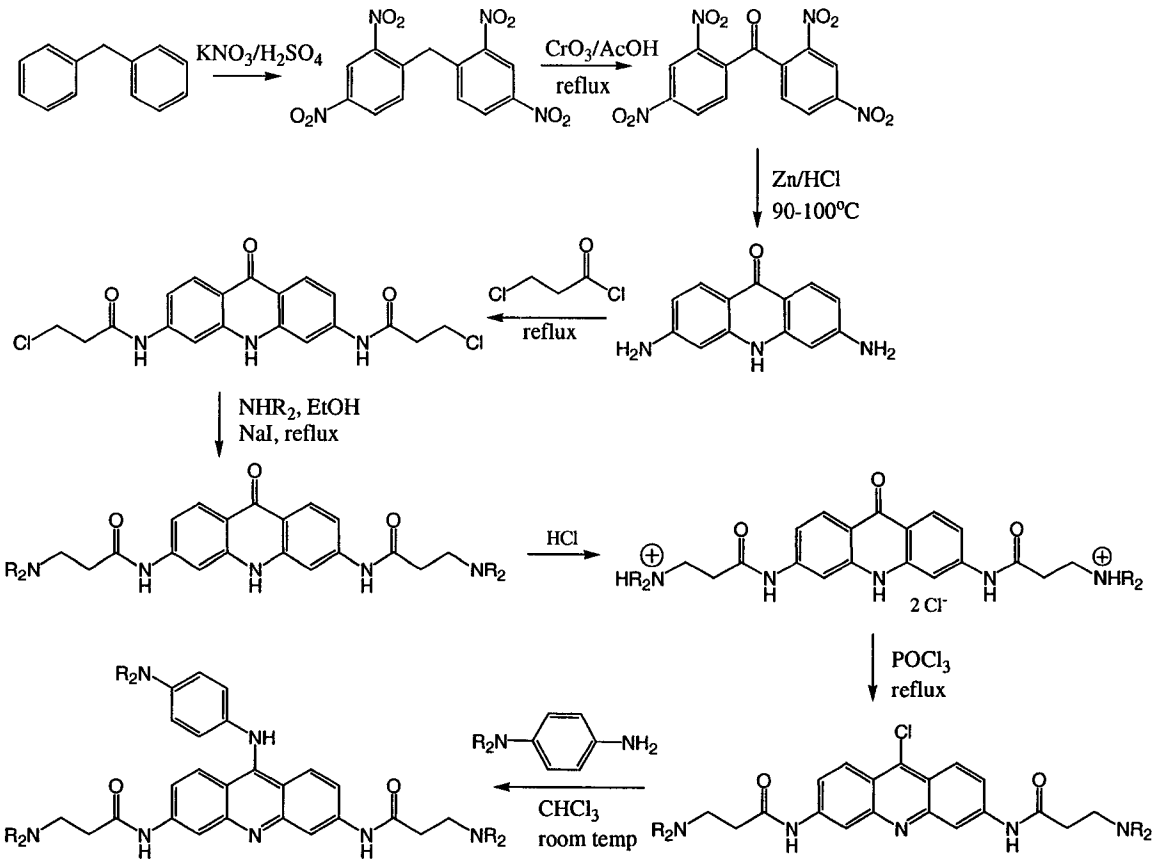


Figure 18: Synthesis of a Trisubstituted Acridine

Neidle *et al.* reported in 2003 the first-ever crystal structure of a G-quadruplex-ligand complex.[30] Up until that point, characterization was based mostly on NMR studies. The ligand that was observed in the crystal structure was a 3,6-disubstituted acridine, with a linker length of two carbons and a pyrrolidine ring as the terminal amine, as shown in Figure 19. The telomeric sequence used in this study was $d(G_4T_4G_4)$, which formed a hairpin dimer.

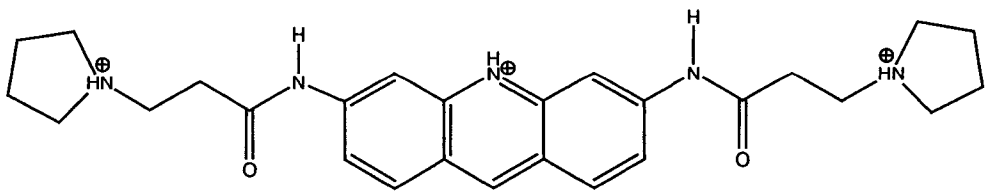


Figure 19: Disubstituted Acridine Ligand Used when Crystal Structure was obtained

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1.6.4.3 Porphyrins

Almost simultaneously in 1998 came two reports of the use of cationic porphyrins to stabilize G-quadruplexes. In fact, identical porphyrins were reported. [31,32] The Sheardy research group felt that the porphyrin shown in the figure was ideal for G-quadruplex binding for a number of reasons. Primarily, the porphyrin has approximately the same dimensions as a G-quartet. Also important was the fact that the porphyrin had groups that

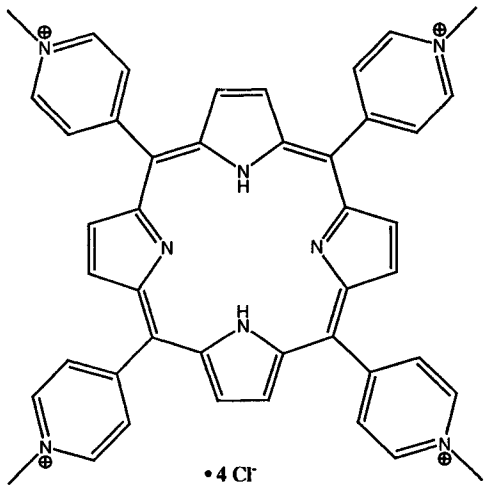


Figure 20: Structure of Porphyrin found to Bind to G-Quadruplex

could interact with a G-tetrad in hydrophobic and electrostatic fashions. Finally, it had been shown previously that this porphyrin could intercalate into double-stranded DNA. This DNA interaction was successfully monitored by visible spectroscopy, and it was felt that this technique could be extended to G-quadruplexes. The Sheardy research group[31] utilized visible spectroscopy, circular dichroism and energy transfer experiments to study the interaction of the porphyrin with the G-quadruplex, while the Hurley research group[32] made use of 1D high-field NMR experiments and UV spectroscopy.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1.6.4.4 Telomestatin

Telomestatin, shown in Figure 21, was isolated from the bacteria *Streptomyces anulatus* 3533-SV4 in 2001.[33] It was the first natural product found to be an effective telomerase inhibitor. Using the repeating sequence $d[T_2AG_3]_4$, which was

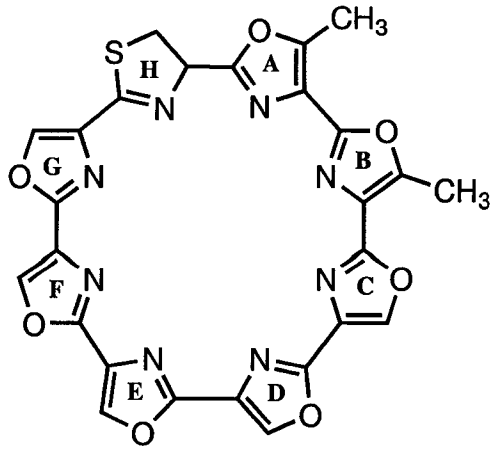


Figure 21: Structure of Telomestatin

known to form intramolecular G-quadruplexes, telomestatin's inhibitory ability was studied. Samples containing this DNA sequence were exposed to varying concentrations of telomestatin. When the concentration of telomestatin was zero, a band of moderate mobility was observed with gel electrophoresis (see Figure 22). As concentration of telomestatin was raised, a band of higher mobility emerged and increased in intensity as concentration of telomestatin increased.[34] The identity of the higher band was assigned as intramolecular G-quadruplex DNA based on previous studies.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

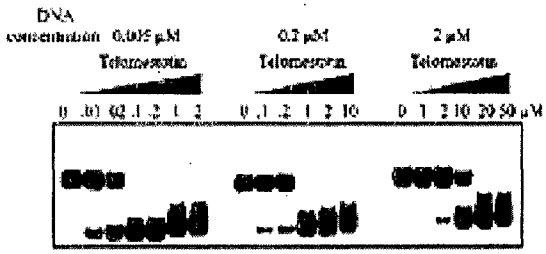


Figure 22: Gel Electrophoresis Demonstrating Telomestatin's Ability to Assist in Formation of Intramolecular G-Quadruplexes (from reference 34)

In an attempt to deduce the mode of binding of telomestatin to the G-quadruplex, DNA 29-mers, each containing six consecutive guanines, were combined with varying concentrations of telomestatin. The choice of DNA sequence was such that an *inter*molecular G-quadruplex would likely form. Interestingly, telomestatin was found to be at least ten times less effective at assisting in the formation of intermolecular G-quadruplexes when compared to those which were intramolecular. This finding pointed towards the loops of the intramolecular G-quadruplex as potential binding sites for telomestatin since these loops are not present in intermolecular G-quadruplexes.

To investigate futher the binding mode of telomestatin to the quadruplex, computational studies were performed using a simulated annealing (SA) docking approach. The sequence used in this study was $d[AG_3(T_2AG_3)_3]$ which forms an intramolecular G-quadruplex as shown in the figure. There are two distinct loop structures in a quadruplex such as this. The lateral loop is formed such that two consecutive GGG triplet repeats on the strand are adjacent to one another in the quadruplex. The diagonal loop is formed by the crossing of two ATT loops such that two consecutive GGG triplet repeats on the strand are diagonal with respect to each other in the quadruplex.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

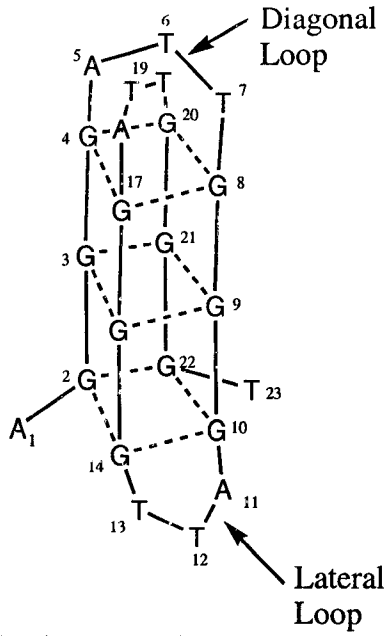


Figure 23: Structure of G-Quadruplex Formed by the Sequence d[$AG_3(T_2AG_3)_3$]

It had been found that each intramolecular G-quadruplex will bind two telomestatin molecules. Computational studies were performed on four models: (1) telomestatin bound to the diagonal loop (1:1), (2) telomestatin bound to the lateral loop (1:1), (3) telomestatin bound to both loops, and (4) telomestatin intercalated between the G-quartets. Of these four binding modes, the 2:1 motif had the most negative calculated binding energy. Intriguingly, though, it was observed that during the formation of a 1:1 diagonal loop complex, the entire complex experienced a dramatic rearrangement. This led to considerable changes in relative position, orientation and potential energy of telomestatin and the G-quadruplex. [34]

1.7 Peptide Coupling

The typical method for PNA synthesis is the use of peptide coupling. Standard peptide coupling involves activation of a carboxylic acid followed by reaction with an amine. There are a wide variety of peptide coupling reagents available for use. Some of the most popular reagents are the carbodiimides. Also crucial in the peptide coupling

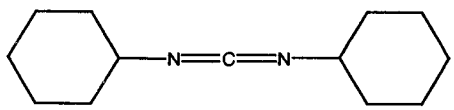
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

reaction is the base used. The base activates the carbodiimide, thus assisting in the deprotonation of the carboxylic acid.

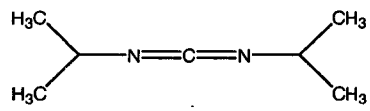
1.7.1 Peptide Coupling Reagents

1.7.1.1 Carbodiimide Reagents

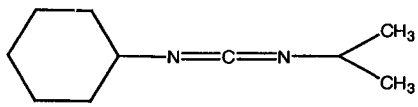
One of the most common carbodiimide reagents is dicyclohexylcarbodiimide (DCC), which was first reported in 1955 by Sheehan.[35] One of the main advantages of using this reagent is that the by-product of the peptide coupling (dicyclohexylurea, DCU) is insoluble in most organic solvents and is therefore easily separable. Other carbodiimide reagents include those depicted in Figure 24.[36] The commonality between all of the above reagents is the carbodiimide functionality. The reagent may be either symmetric (DCC



DCC



DIC



CIC

Figure 24: Carbodiimide Coupling Reagents

and DIC) or asymmetric (CIC). A further advantage to the carbodiimide reagents is that they are relatively inexpensive.

Figure 25 illustrates a typical peptide coupling employing DCC/DMAP. DMAP (4-dimethylaminopyridine) is commonly used as the base with DCC. Historically, carbodiimide reagents were developed to allow for ease of separation of the by-product (as already mentioned) and also to prevent the formation of the N-acylurea by-product.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

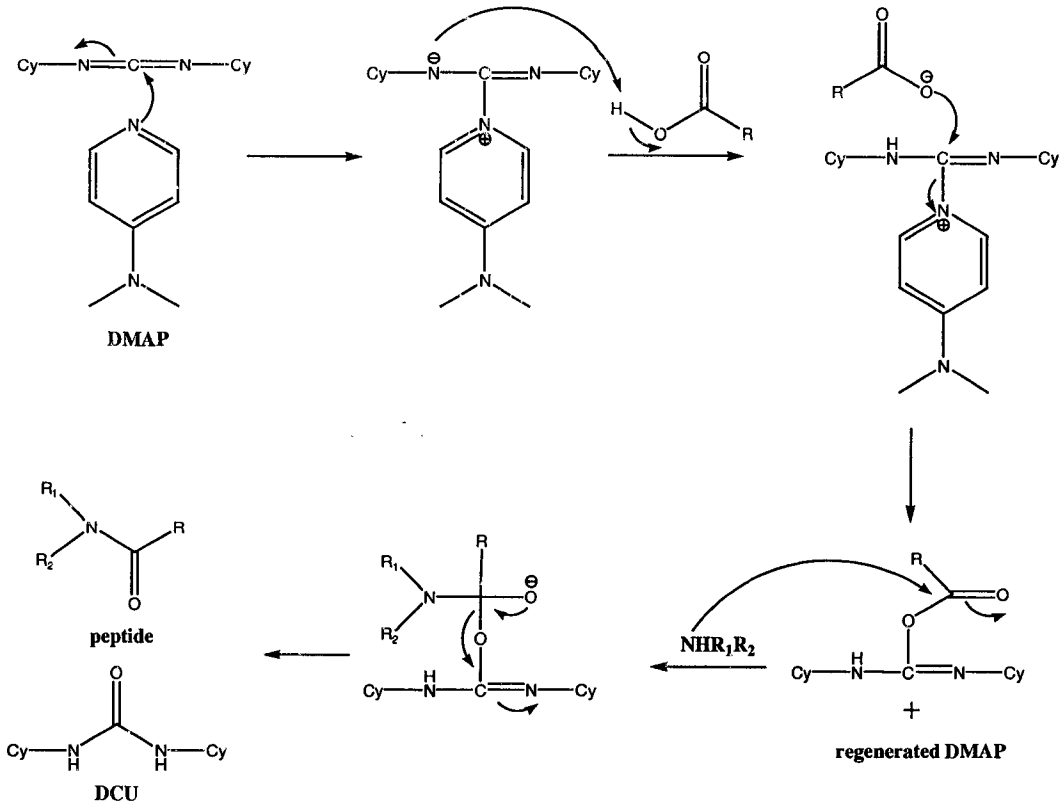


Figure 25: General Mechanism of DCC/DMAP Peptide Coupling (Cy = cyclohexyl)

Many of the bases used in peptide coupling are pyridine-based. DBDMAP has the same structure as DMAP with *tert*-butyl groups in the 2- and 6- position. It may be advantageous to increase bulk around the pyridine-ring nitrogen to decrease its nucleophilicity.

1.8 Solid-Phase Synthesis

1.8.1 Historical Perspective

The term 'solid-phase synthesis' was first introduced by Merrifield in 1963.[37] His introduction of this new and exciting method of synthesis later earned him the 1984 Nobel Prize in Chemistry. The development of this technology stemmed from the need for better methods to synthesize long-chain polypeptides. At the time of publication of his results, synthesis of polypeptides could only be performed in the solution phase. As

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

polypeptide length increased, solubility and the ease of purification decreased. The difficulty with respect to purification is illustrated in the following example.

Suppose one begins with 100 amino acids. If one adds 100 more amino acids to couple to the original amino acids, one could *theoretically* get 100 dipeptides. However, coupling reactions are not usually quantitative, so consider a 95% coupling success. That would result in 95 dipeptides and 10 amino acids after one step. It is crucial in solution phase peptide coupling that non-coupled amino acids are removed before continuing, or varying lengths and combinations of chains would be possible.

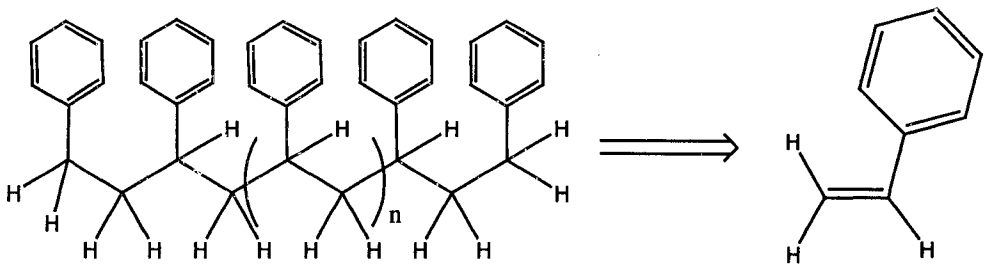
Merrifield desired to develop a method in which solubility of long polypeptides and purification were trivial. He envisioned attachment of the first amino acid to an insoluble polymer support. He reasoned that the peptide chain could be "grown" on this solid support, and that the excess reagents and by-products could be filtered away leaving only the polypeptide attached to the solid support. Simple cleavage from the solid support would yield pure polypeptide.

The support used had to be insoluble in all of the solvents used in coupling chemistry and had the requirement of being of a stable physical form so that it could endure the filtration process. Perhaps most importantly, the support had to have a suitable functional group so that the first amino acid could be attached. Merrifield investigated a number of polymers as possible supports for this new technology, including polyvinyl alcohol and sulfonated polystyrene. For his initial studies, he used chloromethylated polystyrene. This support took the form of 200-400 mesh beads. Their surface was porous which allowed for entry of solvent and reagents. The entry of solvent into these

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

pores was termed "swelling" the resin, since the size of the bead is increased significantly in this process.

Figure 26 shows a small segment of polystyrene and it is derived from the polymerization of the styrene monomer.



Polystyrene **Styrene**

Figure 26: Structure of Polystyrene and its Monomer, Styrene

The polystyrene was functionalized by methylating the benzene rings, followed by the chlorination of the methyl groups. The benzene ring then underwent an electrophilic aromatic nitration. In Figure 27, the sphere represents the rest of the polymer. It is important to note that these reactions are happening simultaneously to each benzene ring in the polymer.

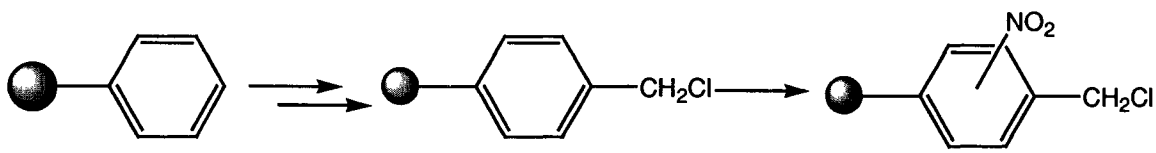


Figure 27: Functionalized of Polystyrene

To attach the first amino-protected amino acid to the chloromethylated and nitrated polystyrene resin, the triethylammonium salt of that amino acid was added, resulting in an ester linkage between the amino acid and the resin, as shown in Figure 28. The protecting group (carbobenzoxy, -Cbz) was removed by treatment with HBr in glacial acetic acid. This method of deprotection was the reason for the nitration of the benzene ring on the polystyrene resin. HBr in glacial acetic acid can also attack the ester

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

linkage between the growing chain and the resin, resulting in premature cleavage. Nitration of the ring deactivates the ester, making it much less susceptible to attack. A second protected amino acid was then introduced in the presence of a DCC to couple

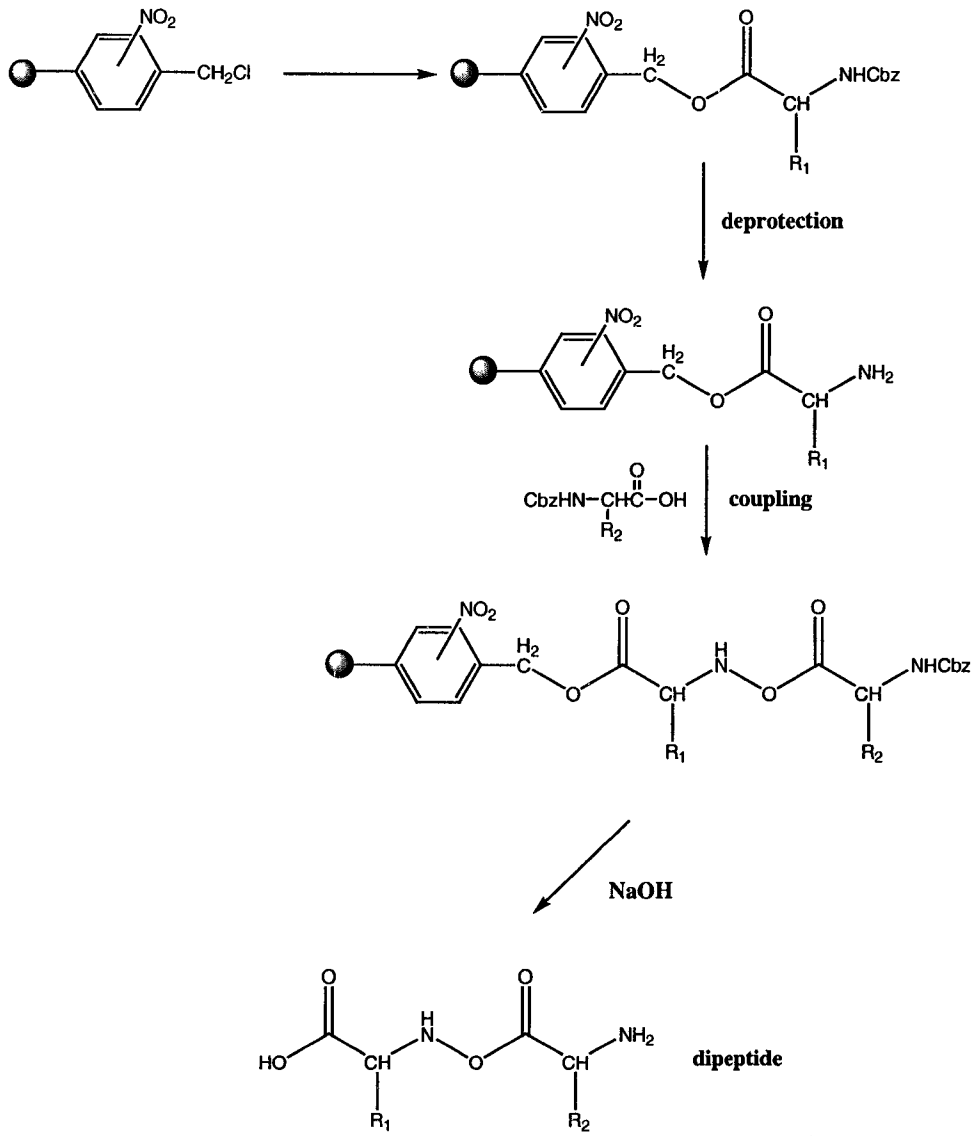


Figure 28: Generalized Scheme of Merrifield's First Solid-Phase Peptide Coupling

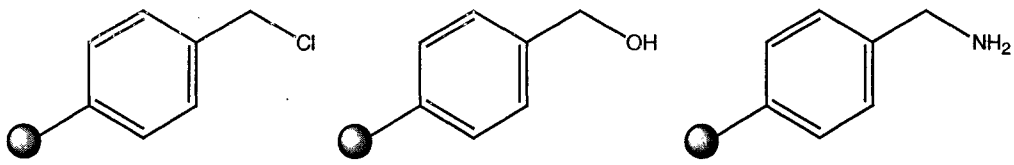
the amino acids. Exposure to NaOH then simultaneously deprotects the terminal amine and cleaves the dipeptide from the resin.

1.8.2 Solid Supports

In modern solid-phase synthesis, the typical solid support has not changed very much since Merrifield's introduction. Supports usually consist of polystyrene with 1-2%

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

cross-linking with divinylbenzene. There are three particular polystyrene supports that are in common use today. They are illustrated in Figure 29.



chloromethylpolystyrene "Merrifield Resin" **hydroxymethylpolystyrene** **aminomethylpolystyrene "AM Resin"**

Figure 29: Modern Solid Supports

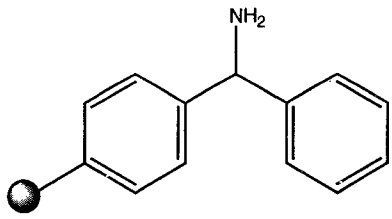
Both the hydroxymethylpolystyrene and the aminomethylpolystyrene can be prepared from the Merrifield resin. AM resins are perhaps the most dynamic of the above resins due to a myriad of possible linkers and spacers which could form amide bonds. The benefit of the amide linkage is, of course, its stability in the presence of strong acid.[38]

1.8.3 Linkers

As of the year 2000, there have been over 200 reported linkers for solid-phase synthesis. In essence, a linker can be considered to be a protecting group attached to a solid-support. Linkers help protect the connection between the growing chain and the solid-support until cleavage conditions are induced. Bradley *et al.*[38] classify linkers into two convenient categories: (1) Integral linkers and (2) Nonintegral or grafted linkers.

Integral linkers are those in which part of the actual solid support forms the linker. Most integral linkers have been synthesized through derivitization of the three common supports mentioned in Figure 30. An example is presented in Figure 30.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

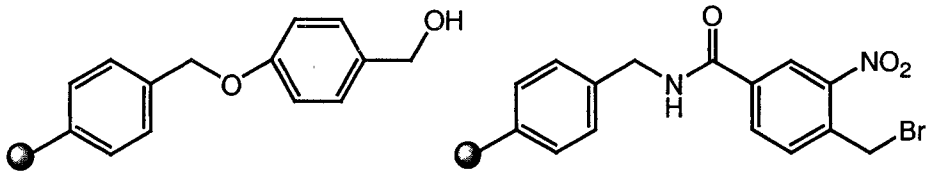


benzhydrylamine (BHA) linker

Figure 30: Example of an Integral Linker

The BHA linker can be derived from chloromethylpolystyrene. Note that with the BHA linker a terminal amine is present, as in AM resin. It is this characteristic that sets integral linkers apart from grafted linkers. The major disadvantage of using an integral linker is that a reaction takes place very close to the resin. Therefore, steric and electronic effects must be considered. It is also difficult to control the extent of the reaction when functionalizing the resin.

Grafted linkers are those in which the terminal functional group in one of the basic resins has been derivatized. These linkers are commonly joined to the support through ether linkages and amide bonds. Two examples are shown in Figure 31.



ether linkage **amide bond**

Figure 31: Examples of Grafted Linkers

There are two important points to note in these examples. Firstly, there still exists the hydroxymethyl or aminomethylstyrene unit, and secondly that there is another functional group at the terminal end of the grafted linker for attachment of the first unit. Without this functional group, grafted linkers would be of no use.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1.9 Project Goals

The goal of this research program is to design molecules that would act as telomerase inhibitors. This research is unique in that the telomerase inhibitors would be designed to recognize the groove of a G-quadruplex. Computational methods would be used to aid in the design and to determine if there is sufficient space for this molecule. A synthetic strategy towards the synthesis of the molecule also needed to be developed.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2. Computational Studies

2.1 Computational Models

The use of computational methods to study chemical and biochemical systems has added an important dimension to research. There are a variety of computational models available for use, each with its own characteristics. Models generally fall into one of two broad categories: quantum chemical models and molecular mechanics models. Quantum mechanical models can be further divided into *ab initio* and semi-empirical models. Each of these will be described in the following sections.

2.1.1 Quantum Mechanical Models

2.1.1.1 *Ab initio* Models

Ab initio literally means 'from the beginning'. These calculations are performed using only quantum mechanical constructs, most notably the Schrödinger equation, and include no reference to experimental data. The Hartree Fock (HF) calculation is the most common type of *ab initio* calculation. Since the Schrödinger equation cannot be solved exactly for any diatomic or larger system, certain approximations must be made. The first Hartree-Fock approximation eliminates the electron repulsion term from the equation. This elimination means that Coulombic factors are not included taken into account. The second Hartree-Fock approximation involves the wave function. Since the exact wave function is known only for a few one electron systems, linear combination of atomic orbitals (LCAO) is used to determine the wave function, or total energy. Density functional theory (DFT) is another popular *ab initio* method. In contrast to HF, total electron density represents the total energy, rather than the wave function.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

These methods provide some of the best qualitative results available computationally, however they are very cost and time consuming. In the HF method, for example, if N is the number of atomic orbitals, the time taken to perform the calculation is N^4. Therefore, if one doubled the size of a molecule, the length of the calculation would increase by 16! Therefore, these methods tend to be applied to small systems only.[39]

2.1.1.2 Semi-Empiricial Models

The most commonly used semi-empirical models are AM1 and PM3. These models are set up the same way as HF calculations, only with more approximations. Semi-empirical models, however, compensate for these approximations by parameterization. Semi-empirical models make use of extensive parameter sets to fit a molecule to its most likely structure. These parameters include bond lengths between specific atom types, bond angles, etc. These data are based on crystal structures of known molecules.

These models, due to the increased number of approximations, are considerably faster than *ab initio* methods. Unfortunately, semi-empirical models are only useful if the molecule under consideration compares well with the molecules that were used to create the parameter sets. Also, semi-empirical models cannot be used practically on very large molecules.

2.1.2 Molecular Mechanics Models

In sharp contrast to *ab initio* and semi-empirical methods, molecular mechanics avoids the use of quantum mechanics completely. One of the most commonly used molecular mechanics model is MMFF. Molecular mechanics expresses the total energy of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

a system as an algebraic equation. The components of this equation are the energy associated with bond stretching and rotation and with intermolecular forces, including van der Waals interactions and hydrogen bonding.

In equation (1), the total steric energy is given by E_{pot}. This is defined as the energy difference between an "ideal molecule" and a "real molecule". E_{bnd} is the energy

$$E_{pot} = \sum E_{bnd} + \sum E_{ang} + \sum E_{tor} + \sum E_{oop} + \sum E_{nb} + \sum E_{el} \quad (1)$$

that results from stretching or compressing a bond length from its equilibrium length. Similarily, E_{ang} is the energy resulting from altering a bond angle from its equilibrium position and E_{tor} is that which results from a change in the equilibrium torsional or dihedral angle. E_{oop} is the component of steric energy resulting from out-of-plane bending, E_{nb} is the energy due to non-bonded interactions, while E_{el} is the energy due to Coulombic forces. The above equation, when expanded, contains a number of force constants and equilibrium values. The values for these terms are taken from experimental data from X-ray, NMR, IR and *ab initio* calculations on specific classes of molecules. This is similar to the parameterization that takes place in semi-empirical calculations. [40]

This model has the advantage that it can model extremely large systems because there is no need to solve a complicated wave function. However, molecular mechanics does not take into consideration any of the interactions between electrons. There can be no information gained from molecular mechanics about a molecule's electron distribution and its reactivity.[41]

2.2 Goal of Computational Studies

The groove of the G-quadruplex was chosen as a target for a synthetic oligomer. Upon examination of the grooves of the G-quadruplex, one can observe that there are

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

many opportunities for hydrogen bond formation. The key was to design an appropriate recognition unit that could recognize the groove of the quadruplex through H-bonds. It was reasoned that a 2-aminopyridine unit would recognize the groove, as shown in Figure 32.

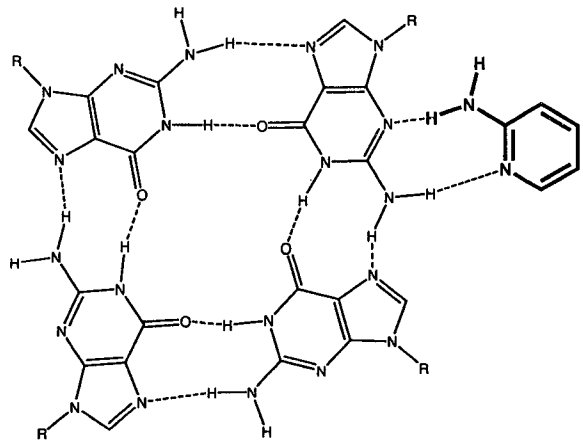


Figure 32: Aminopyridine unit recognizing G-quadruplex

It was envisioned that an oligomer with an aminopyridine recognition unit would be able to lie within one of the grooves of the G-quadruplex .

There are many different backbones available for use, such as DNA and RNA backbones and their modified descendants. However, a peptide nucleic acid (PNA) backbone was chosen to be investigated with computational studies with an eye towards ease of synthesis, flexibility and known cell permeability. Also, the distance between bases on DNA and PNA is virtually identical, making PNA an excellent DNA analogue.

Classically, PNA oligomers have nucleotides attached to their backbone through an acetic acid linkage. It was therefore decided that the repeating unit of the oligomer would be as shown in Figure 33. There are three important features of the repeating unit. (1) the aminopyridine recognition unit, (2) the peptide nucleic acid (PNA) backbone, and (3) the acetic acid linker between the recognition unit and the backbone. Typically, the acetic acid linkage is connected to the secondary amine function of the PNA backbone.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

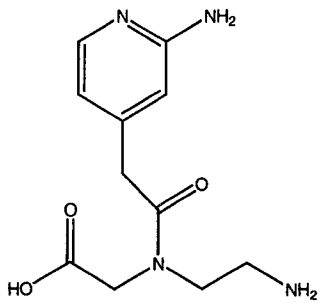


Figure 34: Anticipated Monomer for PNA Oligomer

It was hoped that the computational experiments would address a number of issues. First, would the designed recognition unit be free of steric clashes with the sugar-phosphate backbone? Secondly, would there be room for the recognition units in the groove if they were attached to a PNA backbone. It should be noted that the following experiments were performed in the absence of water molecules, therefore meaningful association energies cannot be calculated. These experiments were for qualitative information only.

2.3 Computational Experiments

A crystal structure of a quadruplex with a coordinated ligand was obtained from the Protein DataBank (PDB)[42]. This quadruplex had four G-quadruplexes stacked on top of each other with three potassium ions, one between each quadruplex. Due to the size of the quadruplex, it was initially thought necessary to use molecular mechanics to calculate the equilibrium state. Using the software package Spartan, the ligand was removed and the hydrogen bonds between adjacent guanine residues in the quadruplex were constrained to 2.0 Å. An MMFF calculation for the quadruplex was performed. This resulted in a minimized structure in which the G-quartets were largely planar and the potassium ions were aligned very well.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

A PNA oligomer with the 2-aminopyridine recognition unit was constructed along the outer edge of the quadruplex structure calculated by the MMFF force field. First, the aminopyridine units were constructed alongside of the G-quadruplex groove. The aromatic ring was then aligned so that the G-quartet and the aminopyridine ring were co-planar. The two hydrogen bonds between each recognition unit and the groove of the G-quartet were then constrained to 2Å. Finally, the PNA backbone was constructed and the four recognition units were attached to the backbone. An MMFF calculation was performed on the G-quadruplex with the PNA oligomer in the groove. It was clear from the result that the Molecular Mechanics force field was not able to maintain the aromaticity in the recognition units. Despite many attempts to define planes along the axis of the ring, the parameterization in Molecular Mechanics was not sufficient for this system. The semi-empirical AM1 force field was then used.

The quadruplex from the crystal structure was subjected to an AM1 calculation with no constraints. The quadruplexes in the resulting structure were almost perfectly planar, and the hydrogen bond length between the guanine nucleotides was between 1.8 and 2.0 Å. As before, a PNA oligomer was constructed, however, there were no constraints placed on the hydrogen bond length. After an AM1 calculation, the hydrogen bond lengths were appropriate and the aromatic rings of the recognition unit were planar. The results of these calculations are illustrated in Figures 34 and 35.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

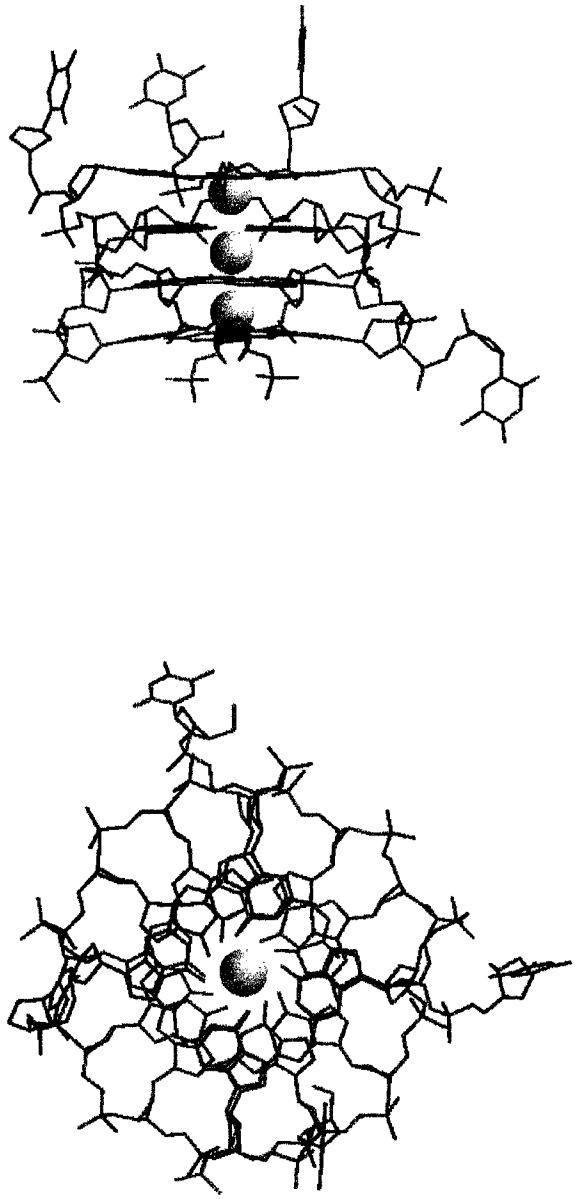


Figure 34: G-Quadruplex after AM1 Calculation, Side View (top) Top View (Bottom)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

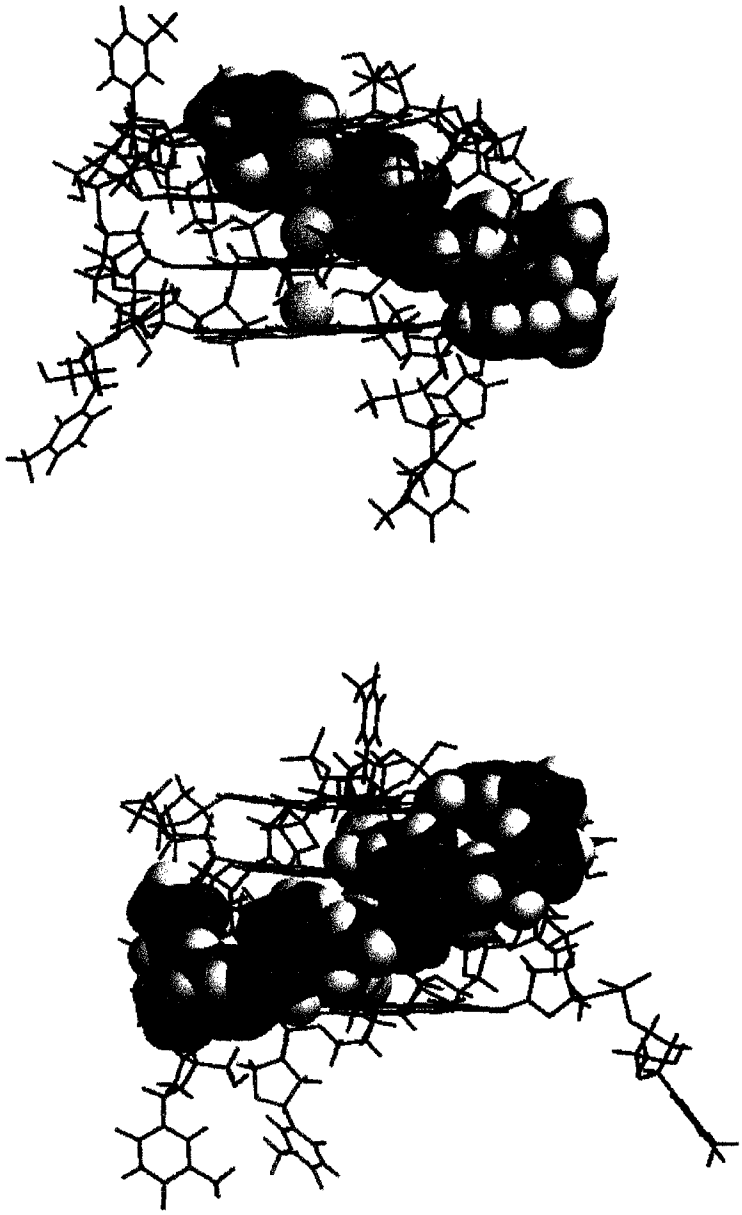


Figure 35: G-Quadruplex (tubes) & PNA Oligomer (space filling)

In Figures 34 and 35, note the groove that occurs between the phosphodiester backbone of the G-quadruplex. The calculations show that sterically, there is sufficient room for the oligomer to lie in the groove. Also significant was that the oligomer was not repelled by the G-quadruplex.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

To determine if there were any differences in the position of the phosphodiester backbone between the G-quadruplex structure with and without the PNA oligomer, several inter-atomic distances were measured. The proton on the amino group of the guanine that hydrogen bonds to the PNA oligomer was used as a reference point. The distance between that atom and the phosphorus atom of the phosphodiester that was attached to that same guanine was measured. Also measured was the distance between the amino proton and the oxygen of the sugar moiety. These distances are illustrated in Figure 36. The atoms under consideration are denoted by*.

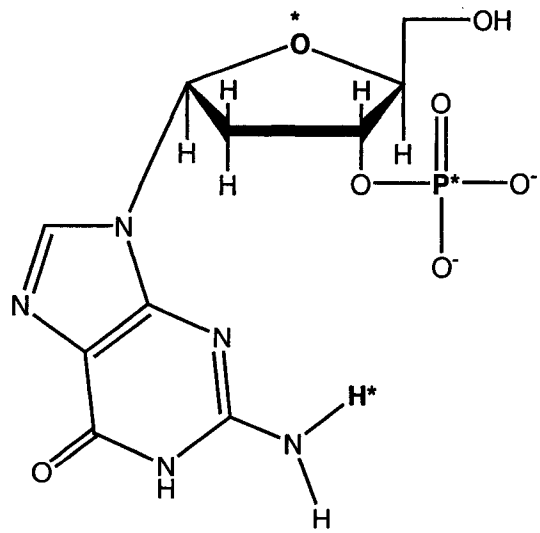


Figure 36: Illustrates Inter-Atomic Distances Measured

The results are summarized in the following tables.

Level	Distance *O-*H (Å) [no PNA]	Distance *O-*H (Å) [PNA]
1	6.005	6.036
2	6.830	6.649
3	5.805	6.064
4	6.606	6.613

Table 1: Distances between deoxyribose oxygen and amino proton

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Level	Distance *P-*H (Å) [no PNA]	Distance *P-*H (Å) [PNA]
1	9.812	9.700
2	9.627	10.152
3	9.826	8.383
4	10.173	9.845

Table 2: Distances between phosphorus and amino proton

The distance between the amino proton and the sugar oxygen remained quite stable, with each of the values to within 0.2Å. This is a semi-quantitative indication that the sugar portion of the backbone does not have to make major adjustments to accommodate the recognition unit. There are, however, larger changes in the distance between the amino proton and the phosphorus atom of the phosphodiester bond, indicating a slight change in conformation.

Finally, to qualitatively examine the interaction between the recognition unit and the G-quadruplex, a cross-section of the G-quadruplex was taken and shown in Figures 37 and 38.

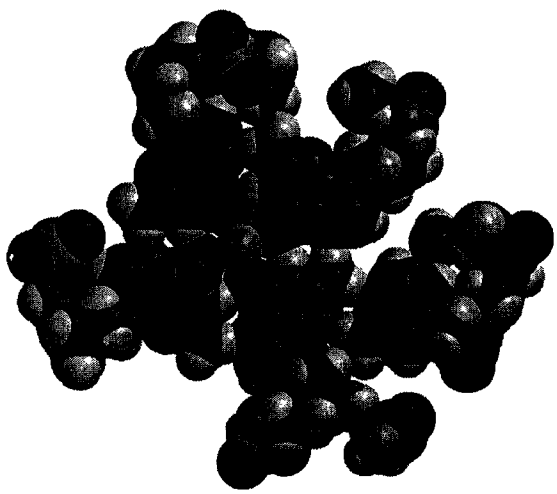


Figure 37: Cross-section of G-quadruplex (space filling)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

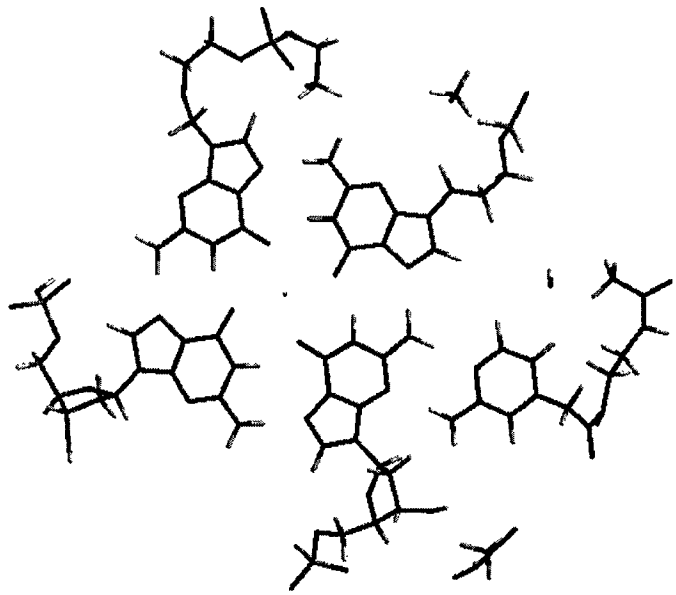


Figure 38: Cross-Section of G-Quadruplex with PNA (tubes)

Upon examination of the Figure 37 and 38, it was observed that there appears to be ample room in the groove for the recognition unit, and that the PNA backbone lies outside of the groove.

Computational calculations have found that there is not a significant disruption in the deoxyribose phosphodiester backbone of the G-quadruplex with insertion of the PNA oligomer. The recognition unit would insert planar to the G-quartet with the PNA backbone outside of the groove. For the PNA oligomer to insert into the groove, water molecules would have to be expelled. This process is anticipated to be entropically favorable.

There are major limitations to the analysis of these data. Primarily, the crystal structure used may not be representative of an actual G-quadruplex. It may just be the conformation adopted when a particular ligand is present. Secondly, these calculations were carried out in a "virtual vacuum", with no solvent molecules surrounding the G-

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

quadruplex. Solvent is a crucial consideration that cannot simply be left out if highly accurate results are desired. As an approximation only, however, solvent molecules may be omitted.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3. Monomer Synthesis

3.1 Choice of Monomer

In Chapter 2, it was established that the desired synthetic target was a PNA oligomer with an aminopyridine recognition unit. If the target oligomer is analyzed in a retrosynthetic fashion, a suitable monomer may be envisioned, as in Figure 39.

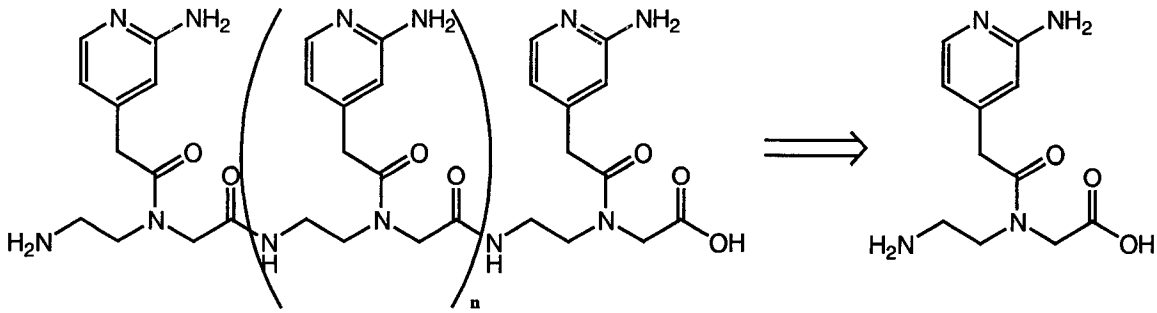


Figure 39: Retrosynthetic Analysis of Desired PNA Oligomer

The oligomer will be synthesized *via* solid-phase synthesis. The amide bond connecting each of the monomers would be able to be formed through well-known peptide coupling reactions involving activation of the carboxylic acid, followed by nucleophilic attack of amine to form the peptide bond.

3.2 Retrosynthetic Analysis of Monomer

Having chosen an appropriate monomer for the synthesis of the desired oligomer, attention was turned to synthesis of the monomer. Retrosynthetic analysis of the monomer yields the following disconnection, as illustrated in Figure 40.

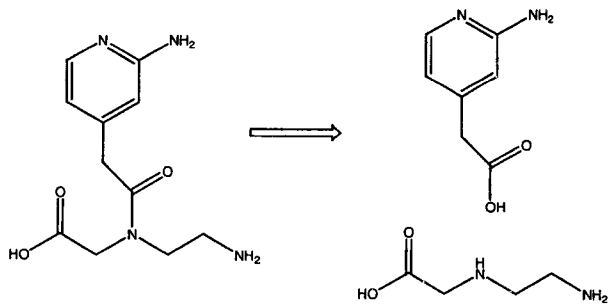


Figure 40: Retrosynthetic Analysis of Desired Monomer

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The connection between the backbone and the recognition unit is a peptide bond. Therefore, a peptide formation reaction was also to be used to synthesize the monomer. The following sections will focus on the synthesis of the 2-aminopyridine unit with the acetic acid tail, termed the 'recognition unit', the PNA backbone synthesis, as well as the peptide coupling to produce the monomer.

3.3 Synthesis of Recognition Unit

A number of different synthetic routes were envisioned for the 2-aminopyridine unit with the acetic acid tail, as described in the following sections.

3.3.1 Construction of Pyridine Ring from Acyclic Precursors

There has been a great deal of research dealing with the *de novo* synthesis of pyridine rings. One of the earliest examples of construction of a pyridine ring from acyclic precursors was reported by Chichibabin in 1905.[43] It consisted of the pressurized condensation of aldehydes with ammonia or amines, as outlined in Figure 41.

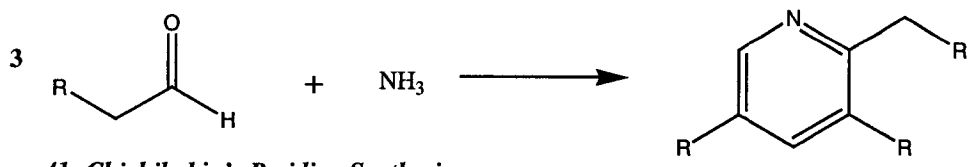


Figure 41: Chichibabin's Pyridine Synthesis

This method did not appear to be suitable for recognition unit synthesis due to lack of functionality at the 4-position.

A more versatile synthesis of pyridine rings was reported in 1961 by Kröhnke.[44] This synthesis employed a 1,4-Michael addition of pyridinium methyl ketone salts to α,β-unsaturated ketones, as illustrated in Figure 42.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

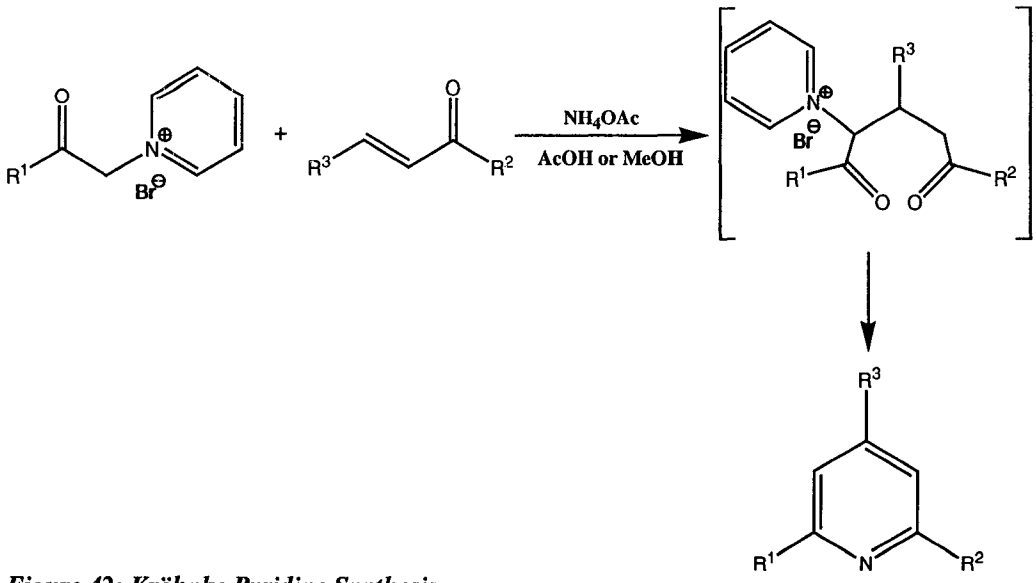


Figure 42: Kröhnke Pyridine Synthesis

The 1,5-dicarbonyl compound then experiences ammonium acetate-promoted ring closure to give the substituted pyridine. To use the Kröhnke Pyridine synthesis for the recognition unit synthesis, one could consider two possibilities of pairs of starting materials. While R^3 must equal $-CH_2COOH$ in this case, R^1 and R^2 are interchangeable. Figure 44 shows the pairs of possible starting materials.

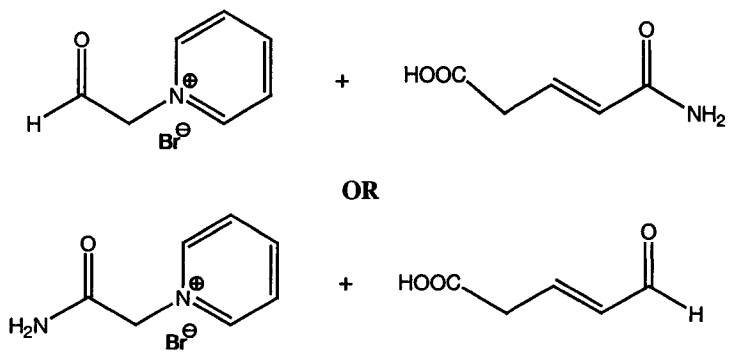


Figure 43: Possible Starting Materials for Kröhnke Pyridine Synthesis

Unfortunately, these starting materials are not suitable for the Kröhnke Pyridine synthesis. R^1 is generally an alkyl or an aryl group, as is R^2 and R^3. [45]

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

A synthesis of a pyridine ring from acyclic precursors that incorporated an amino group in the 2-position was reported in 1993.[46] It was thought that the synthesis could be modified to include a substituent on the 4-position of the pyridine ring. Upon examination

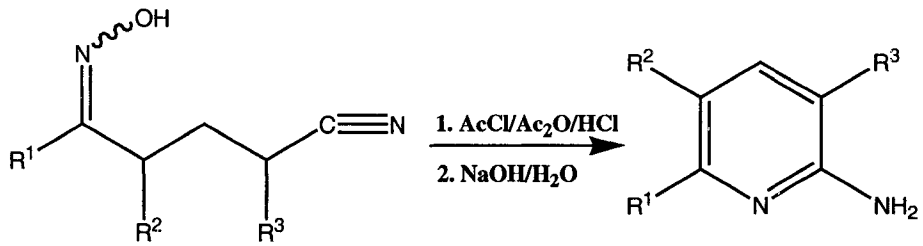


Figure 44: Pyridine Synthesis Incorporating Amino Group in 2-position

of the proposed mechanism though, it was seen that there must be no steric bulk on the carbon that becomes the 4-carbon in the pyridine ring.

There has also been a great deal of research done on the synthesis of pyridine rings mediated by transition metals. Takahashi and co-workers reported a method [47] for the regioselective synthesis of pentasubstituted pyridines with two unlike unsymmetrical alkynes as starting material.

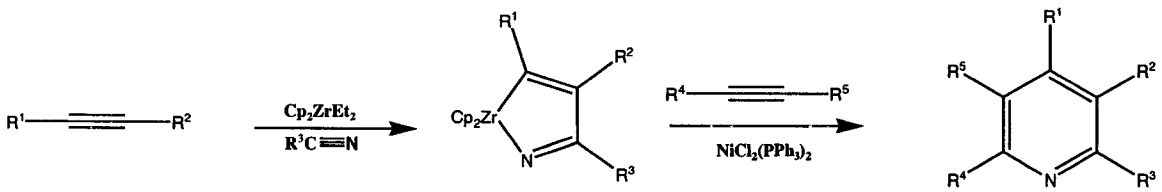


Figure 45: Transition-Metal Catalyzed Synthesis of Substituted Pyridine Rings

At initial glance, this reaction scheme would seem to be suitable for recognition unit synthesis. In this type of synthesis, however, regioselectivity is key. The authors report complete regioselectivity when one of R^1 or R^2 is a silyl group. For the second alkyne, regioselectivity was maintained throught the use of phenyl and alkyl substituents. Clearly, this type of method was not suitable for the recognition unit synthesis.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3.3.2 Chichibabin Reaction

A second route that was envisioned for recognition unit synthesis was to employ the Chichibabin amination of a pyridine ring. It was thought that the reaction of commercially available 4-pyridyl acetic acid hydrochloride and sodium

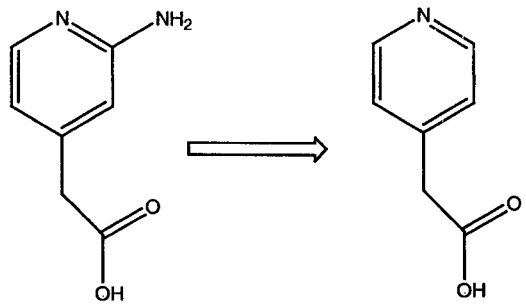


Figure 46: Retrosynthetic Analysis of Recognition Unit using Chichibabin Reaction

amide ($NaNH_2$) would aminate the pyridine ring in the 2-position. The Chichibabin reaction is known to selectively aminate the 2-position of the pyridine ring, as well as a small percentage at the 4-position[48] (Note the difference between the Chichibabin Pyridine Synthesis and the Chichibabin Reaction). It was thought that since the 4-position in the pyridine ring is blocked by the acetic acid group that the reaction would be entirely selective.

One possible problem with the reaction is that *ortho/para* directing groups in the 4-position of the pyridine ring are known to decrease the likelihood of a successful reaction, while *meta* directing groups increase the rate of success.[48] It was reasoned that since the acetic acid group could be considered an *ortho/para* director, then perhaps the reaction would not proceed as desired.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

A second disadvantage to this reaction is that it does not allow diversity in the recognition unit synthesis. It might be advantageous to be able to synthesize the recognition unit with the acetic acid tail in the 3- AND 4-position. This would be useful in the future if the attachment of the recognition unit to the backbone was changed. If the starting material were 3-pyridyl acetic acid, the product would be a mixture of regioisomers. Since the starting material in this case would lack a plane of symmetry, there would be two conceivable places to aminate: on either α-carbon adjacent to the pyridine ring nitrogen. The products would be 2-amino-3-pyridyl acetic acid (undesired product) and 2-amino-5-pyridyl acetic acid (the desired product).

Conceivably, the plane of symmetry and the blocked 4-position in 4-pyridyl acetic acid would allow this reaction to be completely regioselective. However, due to the necessary *ortho/para* group in the 4-position and the harsh conditions of the Chichibabin reaction, research continued to find the optimal method for synthesis of the recognition unit.

3.3.3 Electrophilic Addition of Carbon Dioxide to Lithiated Picoline

In their 1996 paper, Ihle and Krause[49] reported a method for the addition of electrophiles to the methyl group of 2-amino-4-picolines (or 2-amino-4-methylpyridines). Their method was to use the Boc protecting group on the amino function, then use

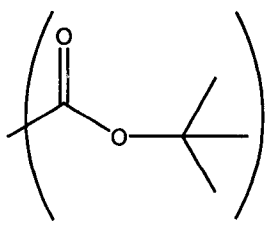


Figure 47: Boc Protecting Group

n-butyl lithium to deprotonate the methyl in the 4-position. A variety of electrophiles, such as allyl and benzyl bromide, as well as propylene oxide and benzaldehyde, were

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

used to quench the anion to produce a variety of products. It was reasoned that if CO_2 were used as the electrophile, the desired recognition unit would be produced.

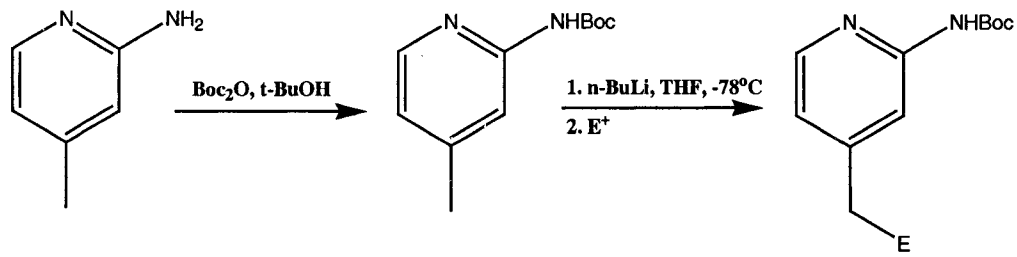


Figure 48: Method Reported by Ihle & Krause

This hypothesis was supported by the communication published by Beyeler, Belser and De Cola in 1997 [50]. The synthesis of their critical intermediate is illustrated in Figure 49.

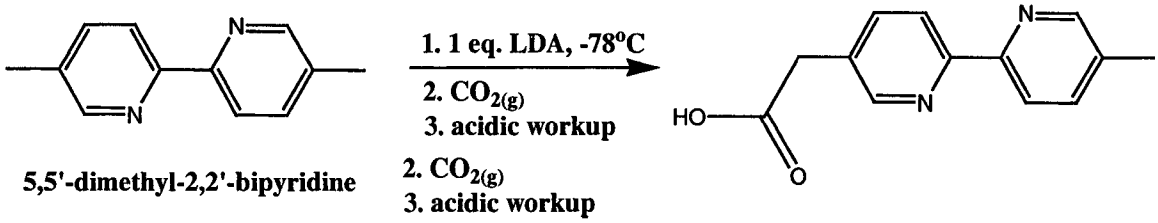


Figure 49: Method Reported by Beyeler, Belser & DeCola

It was decided to adapt a procedure from the carboxylation of 2-amino-4-picoline and from the carboxylation of 5,5'-dimethyl-2,2'-bipyridine. A major advantage of this approach to synthesis of the recognition unit is the potential versatility of the reaction. In theory, both 2-amino-4-picoline and 2-amino-5-picoline could be used as starting material. The reaction is specific for the position of the methyl group on the pyridine ring.

3.3.3.1 Protection of 2-amino-4-picoline

Ihle and Krause's report [49] calls for a Boc protecting group for the amino group. The protecting group is required because the n-butyl lithium will deprotonate the more

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

acidic amino group first before the highly basic methyl group. If there were no protecting group, the -NH group would be able to attack various electrophiles to produce undesired products. The fact that the Boc protecting group was used in this report was particularly advantageous for the research reported in this thesis. In peptide coupling, any amino groups that are not intended to be coupled to form a peptide bond must be protected. In the case at hand, without a protecting group on the amino group of the pyridine ring, the monomer would be able to potentially couple at two places. That would produce a branched oligomeric structure, rather than one that is linear, as desired.

Boc-2-amino-4-pyridine was synthesized according to Ihle and Krause [49]. The symmetrical anhydride of Boc, di-tert-butoxy dicarbonate, was employed. The solvent used was *tert*-butanol. This choice of solvent was interesting since the by-product of the reaction is also *tert*-butanol. The anhydride is attacked by the primary amino group then collapse of the tetrahedral intermediate eliminates one half of the anhydride, which then decomposes to carbon dioxide and *tert*-butanol. The reaction, however, was plagued with poor yields despite a reported 94% yield in the literature. It is believed that the yield is diminished due to the required use of column chromatography to separate starting material from product.

Ihle and Krause's paper gave the chemical shifts observed for Boc-protected 2-amino-4-picoline, but did not assign the heteroaromatic protons specifically. Since it was planned to further derivitize this product, it was thought useful to know the chemical shift of each particular proton on the pyridine ring.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Peak	δ (ppm)	J (Hz)	Integration	Multiplicity
A	6.79	5.1	0.98	d
B	7.83	n/a	0.94	s
C	8.13	5.1	1.00	d

Table 3: Aromatic Peaks in ^1H NMR of Boc-2-amino-4-picoline

The assignment of the heteroaromatic peaks was based on peak multiplicity and on coupling constants. From the above data, it is clear that doublets A and C are coupled to each other since the coupling constant for each peak was identical, and there were no other peaks in the spectrum that were coupled. Based on the value of the coupling constant, it also appears that the protons are separated by two atoms. It was therefore concluded that peaks A and C were due to the protons in the 5- and 6-position of the pyridine ring. It was also concluded that peak C was due to the proton in the 3-position since there are no adjacent protons with which to couple.

To unambiguously assign peaks A and C to a particular proton, the deshielding effect of the pyridine-ring nitrogen was considered. Since the proton in the 6-position of the pyridine ring is much closer to the nitrogen of the ring than is the proton in the 5-position, it would experience a much greater deshielding effect than would the other. It was therefore concluded that peak C was due to the proton in the 6-position of the ring, since it is found furthest downfield. The final assignment of these protons is shown in Figure 50.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

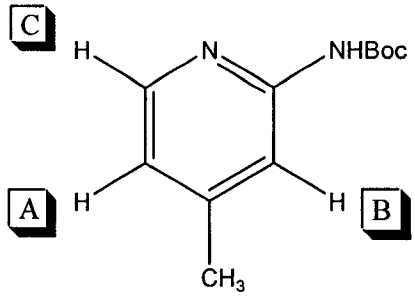


Figure 50: Assignment of Aromatic Protons in Boc-2-amino-4-picoline (in reference to Table 2)

3.3.3.2 Carboxylation of Boc-2-amino-4-picoline

The procedure for the carboxylation reaction was developed from the two references, as previously mentioned. The general reaction scheme is shown in Figure 52.

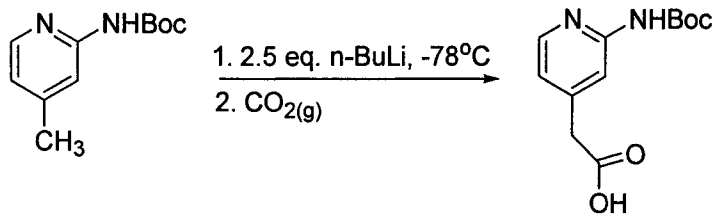


Figure 51: Reaction Scheme for Carboxylation of Boc-2-amino-4-picoline

The general procedure was to dissolve the starting material in dry THF in an oven-dried round bottom flask. After the flask was flushed with N_2, the flask contents were cooled to -78°C using a mixture of dry ice and acetone. The strong base was added to the cooled flask in a 2.5 mole excess. Upon addition of base, the solution became bright yellow unless the dryness of the THF had not been carefully controlled, in which case the solution would turn lime green. In these instances, no reaction would occur with addition of carbon dioxide.

In Ihle and Krause's paper, the cooling bath was removed after complete addition of the base and the solution was allowed to warm to room temperature over two hours with stirring. During this time, the solution changed in color to dark orange. The reaction mixture was then returned to the cooling bath for addition of the electrophile. This step was omitted for some of the trials, and product was obtained in one of those trials, while

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the step was performed in other trials and no product was obtained. The ability to obtain product does not seem dependent on warming the anion solution to room temperature.

In addition, the extremely reactive electrophiles reported in Ihle's report, such as allyl bromide, could be added to the reaction mixture at -78°C and product could be successfully obtained. However, the less reactive electrophiles required a higher temperature for the reaction to be successful. The reaction using CO_2 was performed a number of times with addition of the electrophile at -78°C, but never with a percent yield higher than 8%. It was reasoned that since carbon dioxide is not a *great* electrophile, perhaps a greater temperature would be required for its reaction as well.

The product was successfully characterized by NMR. The ^{1}H NMR spectrum as well as the ^{13}C NMR are found in Figures 52 and 53.

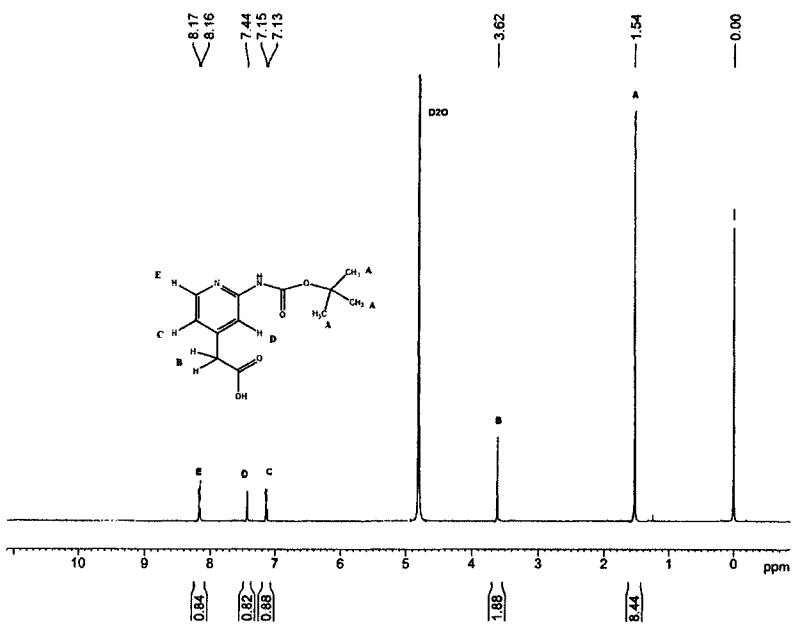


Figure 52: ^{1}H NMR (300MHz, D_2O) of Carboxylated Boc-2-amino-4-picoline

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

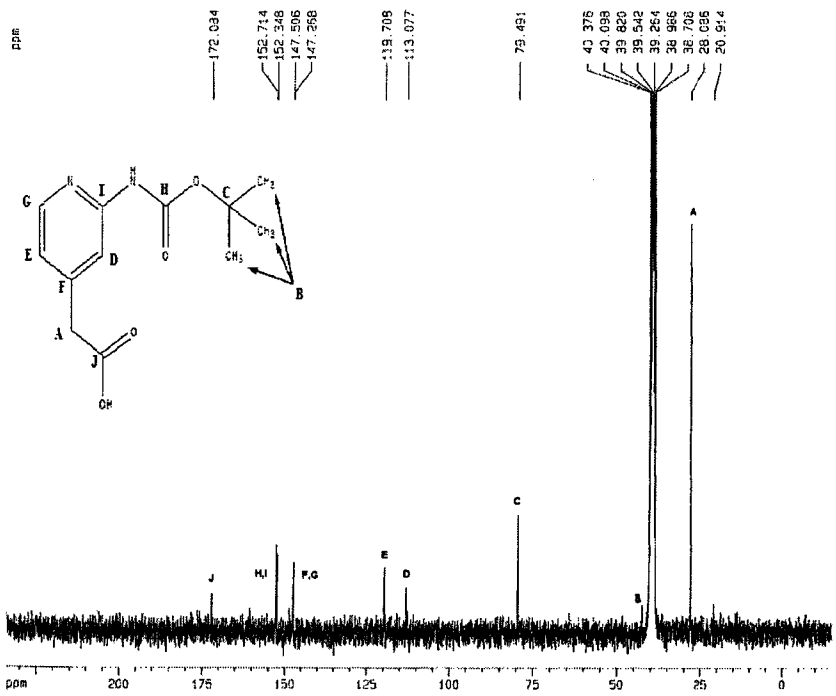


Figure 53: ^{13}C NMR (75MHz, d6-DMSO) of Carboxylated Boc-2-amino-4-picoline

From the above spectra, it was concluded that the desired product had indeed been synthesized. A crucial piece of evidence supporting this conclusion came from each of these spectra. When comparing the ^{1}H NMR of the product and the starting material, one can see a shift of a singlet to a more deshielded position, from 2.35ppm to 3.62ppm. Also significant was the change in integration from 3 to 2. This provided early evidence for the presence of a methylene peak in the product. The second piece of crucial information came from the ^{13}C NMR spectrum. The spectrum was nearly identical to the starting material except for one peak at 172.08 ppm. This peak is highly indicative of a carbonyl in a carboxylic acid [51]. A final piece of evidence for the formation of the desired product was the presence of an –OH peak in ^{1}H NMR spectrum with the sample dissolved in d6-DMSO, a non-exchangeable solvent.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In the final trial of the carboxylation reaction, the temperature of the reaction mixture was warmed to room temperature after the addition of n-butyl lithium. During the warming process, the solution changed color from yellow to bright orange. Carbon dioxide was then added at room temperature. Precipitation was observed instantly and within 30 seconds, all of the orange color had been completely consumed and replaced with a white, milky mixture. The regular work-up procedure was employed and the pH of the water layer was adjusted from 12 to 3.5 using 4M HCl. Partial concentration of the water layer under vacuum caused precipitation of a white solid, which was filtered out. The solid was not soluble in any organic solvents but was soluble in D_2O. The ^1H NMR spectrum showed only the solvent peak. It was concluded that the solid was only a mixture of inorganic salts. The organic layer was evaporated to dryness to yield a white solid, which was determined to be starting material by ^1H NMR. The water layer was then evaporated to dryness and the residue was extracted with absolute ethanol. After filtration of the mixture, and evaporating the ethanol solution to dryness, a sticky solid was obtained. It was not soluble in $CDCl_3$ but was soluble in D_2O. The ^1H NMR revealed an unexpected result. At 3.76ppm, there was a singlet integrating to 2, which is indicative of the desired product, however there was no trace of a singlet at ~ 1.5ppm for the Boc protecting group.

Detailed correlation charts of protective groups' reactivities under various conditions were located.[52] It was noted that at pH 3-4, the Boc protecting group has a "moderate" reactivity. Therefore, it was deduced that the acidification step had in fact deprotected the amino group. The acid used had been changed to 4.0M HCl for the last few trials to minimize the extra volume of solution which would need to be concentrated.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Regardless of the concentration of acid used, the pH was brought to approximately 4.5. It is believed that initial addition of strong acid created a local region of high acid concentration. It was concluded that despite the increase in volume from acidification with 2.0M or 3.0M HCl, these concentrations of acid would be used in the future to prevent deprotection. It was, however, promising to have again isolated a carboxylated product.

A German patent from 1979 [53] was recently located in which the unprotected 2-amino-4-picoline was used as starting material in a successful synthesis of the carboxylated product. The reported method that was used was seemingly identical to that which is described in Section 3.3.3.2. It was reasoned that the product could be isolated due to electron delocalization from the $^{-}$NH group towards the carbanion in the 4-position. The starting material used in this research, however, had the Boc protecting group on the amino and therefore was withdrawing electron density from the amino group, away from the carbanion. This electron withdrawal would make the carbanion less nucleophilic and therefore less reactive.

Literature review revealed that α-aryl carboxylic acids also easily undergo decarboxylation.[54] This is an example of reversible carbon-carbon bond formation. It is therefore proposed that the carboxylated product *is* being formed upon addition of carbon dioxide, however, it is not stable and can either decarboxylate spontaneously, or in the presence of water added to quench the reaction. This theory is evidenced by massive foaming that occurs when the THF is removed by rotary evaporation immediately following quenching. It was therefore concluded that the desired recognition unit was not stable and the synthesis of significant quantities *via* this method would not be possible.It

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

was therefore concluded that the presence of the Boc protecting group destabilizes the already unstable carboxylic acid recognition unit. This conclusion is backed up by a great increase in the % yield of the product, albeit unprotected.

3.3.3.3 Possible Alternatives to Carboxylation Reaction

It was thought that perhaps a functional group other than a carboxylic acid could be employed. It was conceivable to still use the same starting material and deprotonate using n-butyl lithium, however, a different electrophile could be used to generate in a pre-activated recognition unit. For instance, as illustrated in the following scheme, phosgene could be used as the electrophile to give an acyl chloride.

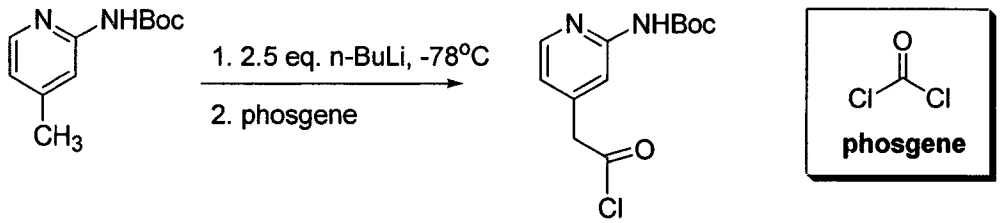


Figure 54: Alternative Synthetic Route to Carboxylated Boc-2-amino-4-picoline

This reaction could be performed *in situ* and then immediately added to the backbone for peptide coupling. However, the major toxicity of phosgene is a definite drawback to this proposal. Also, since the acyl chloride is so reactive, it could potentially dimerize two carboxylic acids into a symmetrical anhydride.

An alternative route that was envisioned for the future was to generate the carbanion as per usual, and react it with carbon dioxide but then convert the carboxylic acid to an acyl chloride *in situ*. The de-carboxylation appears to occur primarily during rotary evaporation of THF and during acidification. Therefore, by avoiding these steps, decarboxylation could conceivably also be avoided. The carboxylic acid could be converted to an acyl chloride through conventional means such as thionyl chloride ($SOCl_2$) or phosphorus chloride (PCl_3).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3.3.4 Bromination of 2-amino-4-picoline

With Boc- 2-amino-4-picoline as the starting material, an alternative synthesis of the recognition unit was designed. It was thought that the methyl group in the 4-position could be brominated with N-bromosuccinimide (NBS) in an analogous fashion to the

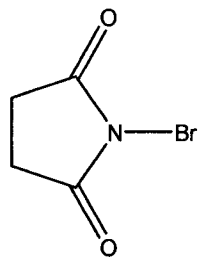


Figure 55: N-Bromosuccinimide (NBS)

bromination of the benzylic position of a substituted benzene ring. The bromide could then be converted to a nitrile by use of sodium cyanide in an S_N2 reaction. Finally, the nitrile could be hydrolyzed, catalyzed by acid, to produce the carboxylic acid, as shown in Figure 56.

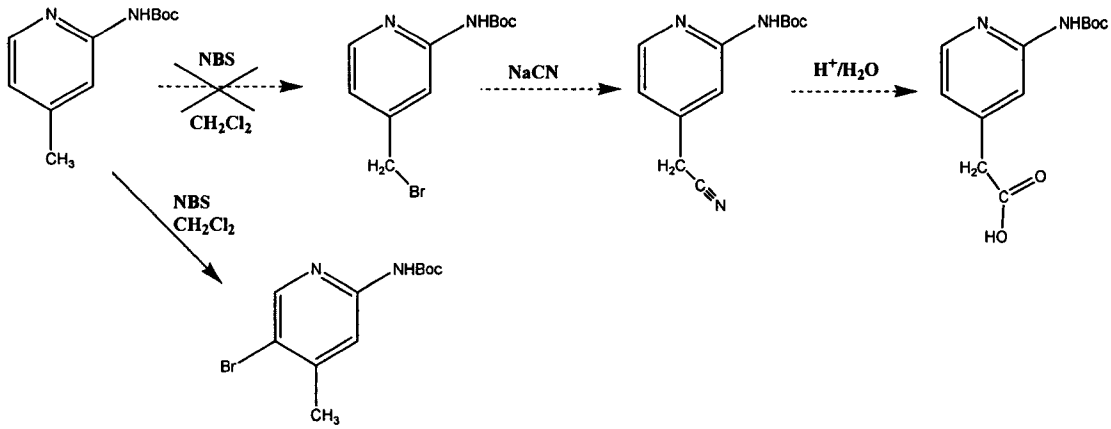


Figure 56: Proposed Synthetic Route to Recognition Unit via Brominated Intermediate

The bromination reaction was performed with Boc-2-amino-4-picoline as starting material and using a slight excess of NBS (1.1 equivalents)[55]. The reaction mixture was refluxed for 4 hours and monitored by TLC. After four hours, TLC of the reaction

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

mixture revealed three spots (R_f= 0, 0.11 and 0.32, 95% hexane/ethyl acetate). By performing a TLC with the starting material, NBS and the final reaction mixture on the same plate simultaneously, it was determined that NBS had an R_f= 0, while the starting material had an R_f = 0.11. It was therefore concluded that the product of the reaction was the least polar of the three and had an R_f = 0.32.

Column chromatography was employed to separate these compounds, starting with the eluting solvent at 97% hexane/ethyl acetate. Elutions were not showing UV activity, even after 150 mL, therefore the solvent system was changed to 95% hexane/ethyl acetate. Four 50mL fractions were sufficient to elute the product off the column. ^{1}H NMR ($CDCl_3$) data of the product are given in Table 4.

Peak	δ (ppm)	Integration	Multiplicity	Assignment
A	**1.54**	**9**	**s**	**-C(CH_3)$_3$**
B	**2.40**	**3**	**s**	**-CH_3**
C	**7.92**	**1**	**s**	**Heteroaromatic proton**
D	**8.25**	**1**	**s**	**Heteroaromatic proton**

Table 4: ^{1}H NMR chemical shifts for brominated picoline

Two distinct observations were made with reference to the above data to deduce the structure of the product. First, the singlet at 2.40 ppm was only very slightly shifted from the singlet observed in the starting material at 2.36 ppm. Also, the singlet integrated for three, indicating a methyl group instead of the desired methylene group. Secondly, upon comparison of the ^{1}H NMR of the starting material and the brominated product, it was observed that one of the heteroaromatic protons was no longer present.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In the aromatic region, the starting material displayed two doublets (6.79 and 8.13 ppm) as well as a singlet (7.83 ppm). This pattern of peaks was consistent with a 2,4-disubstituted pyridine ring. The product of bromination had only two singlets (7.92 and 8.25 ppm) in the aromatic region. It was easily deduced that one of protons A or C (see Figure 51) had been replaced, since there was no longer any coupling in the aromatic region.

If one considers an electrophilic aromatic substitution of a proton in the 5 or 6-position for a bromine, the answer becomes clear. When the aromatic electrons in the bond between carbons 5 and 6 attack NBS, the bromide can be added to either the 5 or 6 carbon, leaving a carbocation on the adjacent carbon. When resonance is considered, it is found that one of the resonance structures derived from having the carbocation in the 5-position results in a highly unstable intermediate in which the pyridine-ring nitrogen has an incomplete octet and is positively charged. This situation does not arise when the carbocation is in the 6-position and the bromine is in the 5-position. This method for recognition unit synthesis was abandoned since the 4-methyl group was not brominated.

3.4 Synthesis of PNA Backbone

With synthesis of the recognition unit complete, attention was turned to the PNA backbone. PNA is composed of an N-(2-aminoethyl)glycine backbone. There are a variety of methods available to prepare these backbones [56-58]. It was desired that there be a trityl (triphenylmethyl) group on the terminal amine. This was due to the intent to use the Trityl-ON purification method (to be described later in Discussion). The references that were located gave many different possibilities of terminal amine protecting groups

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

and carboxylic acid protecting groups. The desired protected backbone is as illustrated in Figure 58.

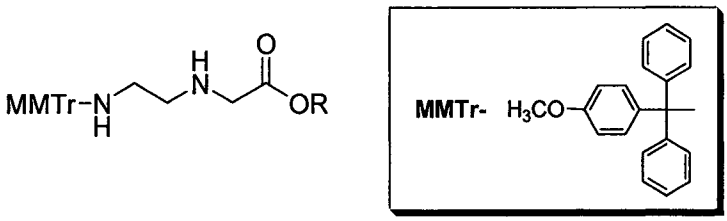


Figure 57: Structure of Desired PNA Backbone for Monomer

3.4.1 Retrosynthetic Analysis of Backbone

The desired backbone could be easily envisioned as being derived from three commercially available components, as illustrated in the following scheme:

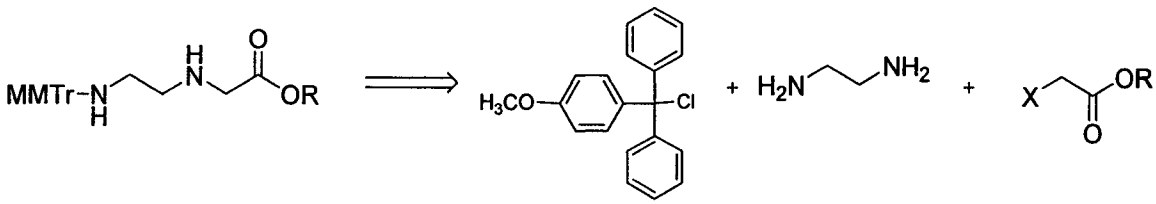


Figure 58: Retrosynthetic Analysis of Desired Backbone

To synthesize the backbone, the necessary components would be ethylenediamine, a halogenated methyl ester of acetic acid and monomethoxytrityl chloride.

3.4.1.1 First Synthetic Attempts at Backbone

Initially, the method chosen was that reported by Breipohl *et al* in 1997 [58]. The final backbone would have a *p*-monomethoxytrityl group on the terminal amine and a *tert*-butyl ester at the opposite terminus. The reaction scheme is found in Figure 59. First, ethylenediamine was protected using di-*tert*-butyl dicarbonate to produce a mixture of mono- and bis- protected product. These are easily separated, as reported by Krapcho and Kuell[59]. Due to their aqueous solubility, excess ethylenediamine and bis-protected ethylenediamine can be removed by washing with water.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The glycine portion of the backbone was introduced by adding *tert*-butyl chloroacetate with a catalytic amount of potassium iodide and an equivalent of base to absorb the hydrochloric acid by-product. The synthesis was completed successfully up until this point. At this point, the authors report that the backbone could be coupled with various nucleotide bases, and the resulting product could be treated with

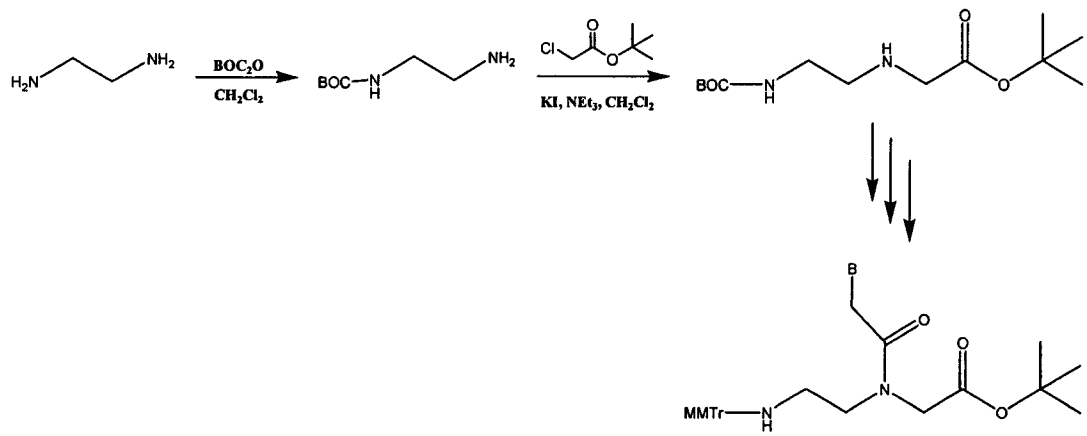


Figure 59: Synthetic Route for First Attempt at Backbone Synthesis

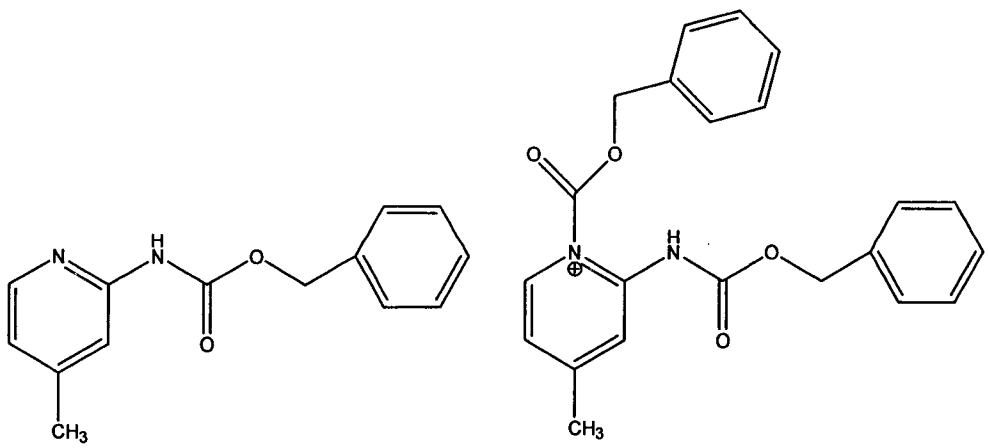
trifluoroacetic acid to simultaneously deprotect both the terminal and secondary amines. Only at this point is the monomethoxy trityl group introduced. The article boasts that this is a major improvement over other syntheses because the trityl group is highly acid-labile and introduction at the last step prevents premature de-tritylation.

The use of the Boc protecting group on the backbone prevented the use of Boc on the pyridine ring amine of the recognition unit. In this particular synthesis, the introduction of the MMTr- group was dependent on the BOC-deprotection of the terminal amine, which would also deprotect the pyridine ring amine. It was therefore decided that a different protecting group would be used for the recognition unit. The Cbz- or Z- (benzyl carbamate) protecting group was chosen. Like Boc, Cbz- is a carbamate protecting group, but unlike Boc, it is not susceptible to acid cleavage.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3.4.1.2 Change in Protecting Group for Recognition Unit

To produce the Cbz-protected recognition unit, an analogous method to producing the Boc-protected recognition unit was employed. 2-amino-4-picoline was treated with benzyl chloroformate to yield Cbz-2-amino-4-picoline. The TLC of the crude reaction mixture revealed two spots and these were separated using column chromatography with a 70% hexane/ethyl acetate solvent system. Two different compounds were isolated from the column. The identity of the top spot on the TLC was the desired product. The bottom spot was believed to be bis-protected 2-amino-4-picoline, i.e. the pyridine nitrogen had also been protected. This conclusion was made from the fact that the integration in the aromatic region was too large for only the pyridine ring and a phenyl ring, and also from the presence of two methylene peaks.



mono-protected **bis-protected**

Figure 60: Structure of Products Isolated from Column after Reaction of Cbz-protection 2-amino-4-picoline

Unfortunately, there was no success when an attempt was made to carboxylate the Cbz-protected molecule in an analogous manner to the Boc-protected molecule. Careful examination of the literature describing the properties of the Cbz- protecting group revealed that it is not expected to be stable in the presence of alkyl lithiums.[52] The

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

expected reaction is a return to the starting material. Ironically, that same literature revealed that the Boc-protecting group, as well as all carbamate protecting groups, are not expected to be stable in the presence of alkyl lithiums.

3.4.1.3 Improved Synthesis of PNA Backbone

With the inability to produce a Cbz-protected recognition unit in this manner, attention was turned to modifying the method of backbone synthesis to accommodate the Boc-protecting group on the recognition unit. A reference that described a much simpler and versatile synthesis of a PNA backbone was suggested. [60-62] In three conceptually simple steps, a PNA backbone with a monomethoxy trityl group on the terminal amine and a methyl ester flanking the other terminus could be synthesized.

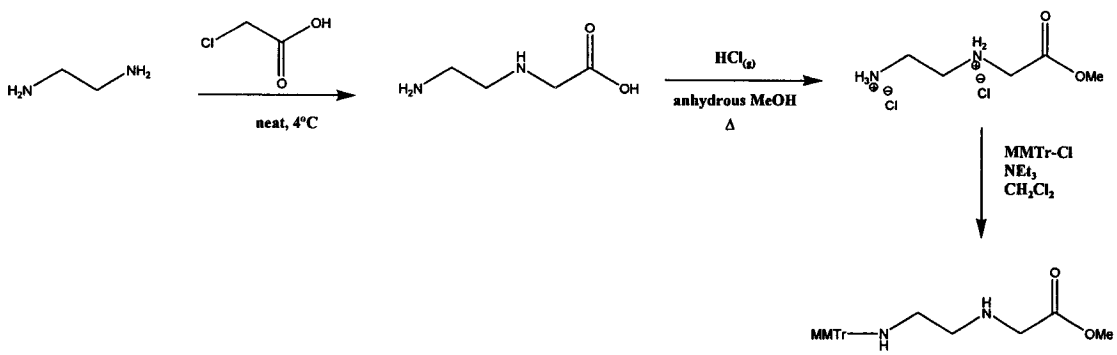


Figure 61: Synthetic Route Followed for Second Attempt at Backbone Synthesis

A ten fold excess of ethylenediamine was combined with chloroacetic acid in a neat reaction at 0°C. The large excess of ethylenediamine combined with temperature control promoted reaction at just one of the amino groups of the ethylenediamine unit. The reaction was stirred overnight and the ethylenediamine was then evaporated *in vacuo*. Trituation with DMSO yielded white solid after standing. Bromoacetic acid could not be successfully used in this manner. For bromoacetic acid to be soluble in ethylenediamine, the temperature must be approximately 4°C (cold water bath). However, this temperature is too high to restrict the addition to one amino group. When

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the reaction was attempted using bromoacetic acid, no product could not be isolated presumably due to the solubility of bis-protected ethylenediamine in DMSO.

The ^{1}H NMR of the product, N-(2-aminoethyl)glycine, was clean and showed two triplets for the methylene groups between the nitrogens and a singlet for the methylene in the acetic acid unit. Each peak integrated equally. Since the product was dissolved in D_2O for the NMR, the amino protons and hydroxyl protons are exchanged away and therefore not visible in the spectrum.

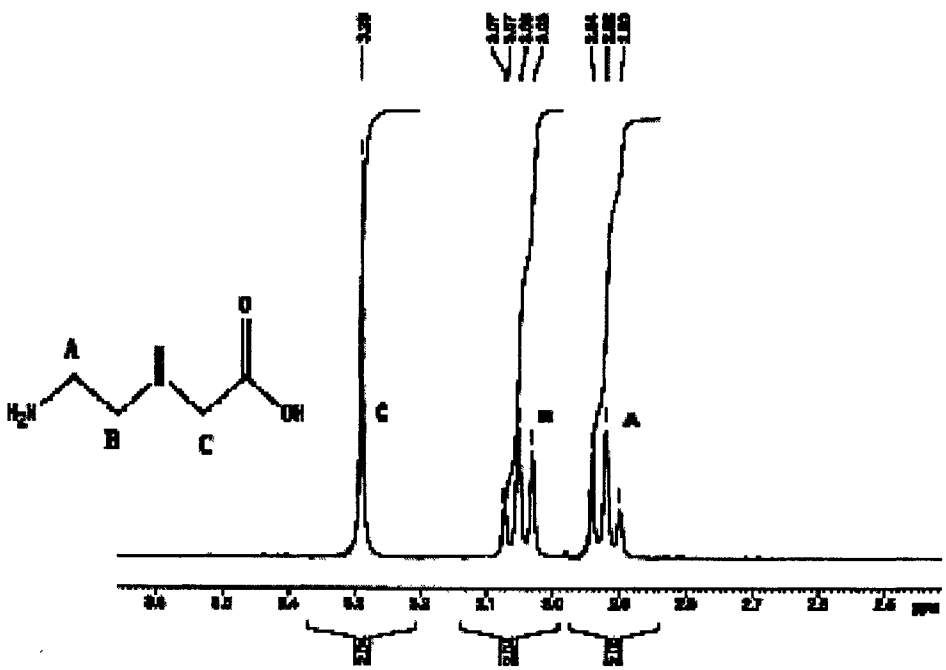


Figure 62: ^{1}H NMR (300MHz, D_2O) of N-(2-aminoethyl)glycine (2.5-3.6ppm)

Chemical reasoning would imply that the futher away from the carboxylic acid group a methylene group was, the further upfield its peak would be. To gain insight into

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the 3-dimensional shape of N-(2-aminoethyl)glycine and to confirm which triplet was due to which methylene, a NOESY NMR experiment was performed. NOESY (nuclear overhauser effect spectroscopy) is a 2-dimensional NMR experiment that examines the spatial relationships between protons.

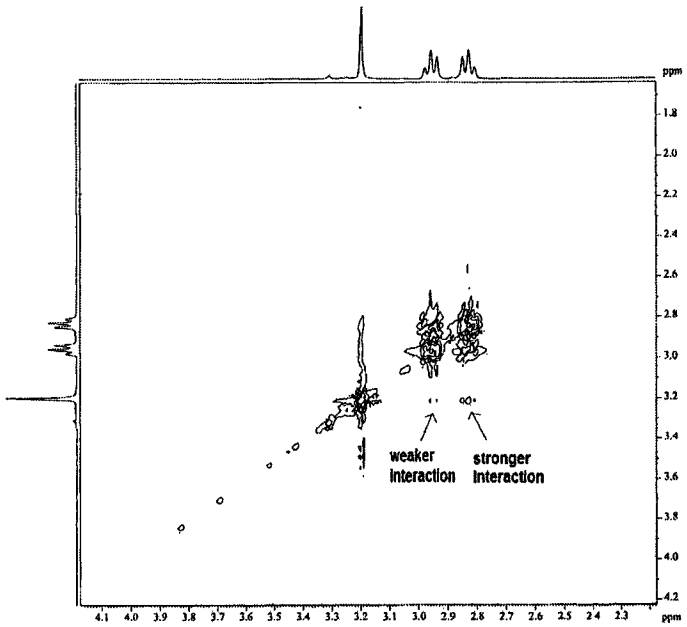


Figure 63: NOE Spectrum of N-(2-aminoethyl)glycine

In Figure 63, the ^{1}H NMR is found on both the x and y axes. To determine which protons are close to one another in space, a line was drawn vertically down from each triplet to the level of the singlet along the horizontal. It is clearly seen that there is a stronger interaction between the more shielded triplet and the singlet. Combining the data from the ^{1}H NMR and the NOSEY spectrum, one can infer that despite the protons due to peak A and peak C are further apart along the chain of the molecule, they are actually closer together in space. This indicates that the actual structure must be somewhat folded.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

This is not inconceivable owing to N-(2-aminoethyl)glycine being set up perfectly to cyclize to form a six-membered ring.

N-(2-aminoethyl)glycine was then esterified using anhydrous methanol and hydrogen chloride gas. In the first attempts at this reaction, compressed hydrogen chloride gas from a cylinder was used, as described in the reference. However, in later attempts, a procedure modification was made. It is possible to generate anhydrous hydrogen chloride gas *in situ* through the reaction of acetyl chloride and anhydrous methanol.[52] The product of this reaction is anhydrous $HCl_{(g)}$ and acetic acid methyl ester.

One hour of reflux is required, and upon cooling to room temperature, the product begins to crystallize. Addition of ether is required for crystallization to occur if HCl has been generated *in situ*. A simple filtration yields the double hydrochloride salt of N-(2-aminoethyl)glycine methyl ester. It can be observed from the NMR spectra of the products that resulted from *in situ* generation of HCl that despite pumping down on vacuum for several hours, there remains traces of acetic acid methyl ester. Future attempts may include a recrystallization step to complete this synthesis.

The final step of the backbone synthesis is a highly versatile protection step based on the research of Di Giorgio *et al* [62]. They report protection of N-(2-aminoethyl)glycine methyl ester · 2HCl with monomethoxytrityl (MMTr-), Boc- and Cbz-. This diversity is highly advantageous for future variations of the monomer. The salt was combined with MMTr-Cl in dichloromethane with three equivalents of triethylamine (NEt_3) in CH_2Cl_2 at 0°C. After stirring for two hours, a standard work-up is completed to yield a yellow oil. Chromatography conditions are reported within the reference, however it was learned [60] that the crude product was suitable for coupling procedures.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

It should be noted that this reaction will not yield the desired product if performed at a higher temperature. The identical reaction was performed at 4°C and was found to yield largely bis-tritylated product. In previous reactions, as well as in the literature reference, the ^{1}H NMR displays two singlets at 3.72 and 3.78 ppm that integrated to approximately three each. These peaks are due to the two methoxy groups in the backbone: one on the trityl group and one as a carboxylic acid protecting group. In the ^{1}H NMR of this product, there are three methoxy peaks are shifted slightly upfield. A third singlet can be observed within 6 Hz of the more downfield peak. The peak position of the methylene protons of the acetic acid unit also provide evidence for bis-tritylation. With the mono-tritylated product, there occurs a singlet integrating for 2 at 3.35 ppm. However, in the NMR of this product the singlet has shifted upfield to 3.04 ppm. This observation would be consistent with a trityl group present on the center nitrogen of the backbone.

3.5 Peptide Coupling

With the two components of the monomer successfully synthesized, a DCC/DMAP coupling was chosen for initial studies since these reagents are commonly available and there are well known methods for their use. With respect to the carboxylic acid to be activated, the coupling reaction generally required one equivalent of DCC and a catalytic amount of DMAP. The polar, aprotic DMF is generally used as solvent. The general procedure was adapted from Yuan *et al.*[63]

3.5.1 Synthesis of Control Monomer

To begin solid-phase studies, a model compound was initially studied (see Figure 65). The commercially available 4-pyridyl acetic acid hydrochloride was stirred with dry

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

DMF and neutralized with one equivalent of triethylamine. Complete dissolution occurred

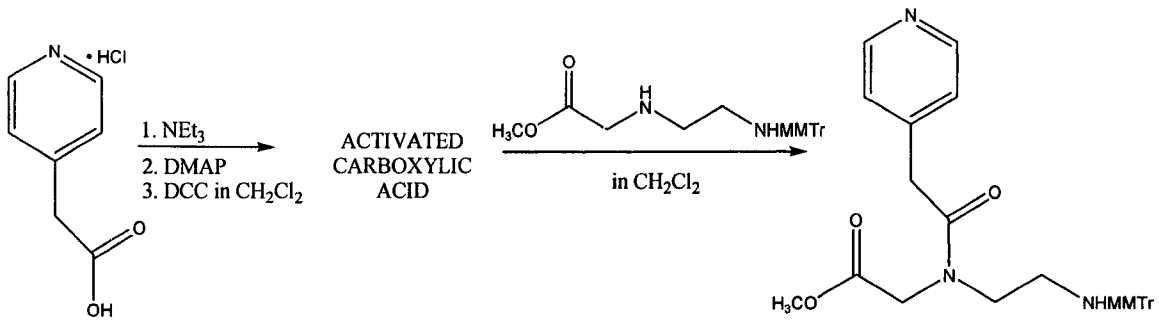


Figure 64: Synthesis of Control Monomer

after neutralization and approximately five minutes of stirring. The catalytic DMAP was then added, followed by the addition of DCC dissolved in CH_2Cl_2. This final addition caused the solution to turn bright yellow, but after overnight stirring, the solution returned to colorless. An appropriate amount of protected backbone, dissolved in CH_2Cl_2, was then added to the reaction mixture and stirred for four hours.

The first attempt at this synthesis used one equivalent of both recognition unit and backbone. It was found, however, that the reaction did not go to completion and there was a great deal of backbone left over, as observed by TLC. Since both the coupled product and the backbone have trityl groups, they are both extremely soluble in organic solvents. This precluded using a water extraction to separate the two. A second attempt at the reaction was made using 1.5 equivalents of recognition unit and 1 equivalent of backbone. It was hoped that this mole ratio would help consume the backbone completely. Any leftover recognition unit was anticipated to be removed by a water extraction. TLC analysis revealed an apparently smaller amount of backbone relative to other bands, but unfortunately the excess recognition unit was not soluble in water.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The work-up consisted of three separate washings. The first was with water, the second with saturated potassium bicarbonate and the third with brine. The standard acid wash was omitted due to the presence of the highly acid sensitive trityl group. It was crucial also to add ~0.5% NEt_3 to each of the washing solutions, as trace amounts of acid and even water might removed the trityl group. The yellow organic layer was then simply evaporated to dryness to leave a yellow oil with some solid suspended within it. TLC of this oil showed several bands. When the oil was stirred with 80% hexane/ethyl acetate/0.5% NEt_3, a great deal of white solid crashed out and could be removed by filtration. TLC of the filtrate revealed that one of the bands had disappeared. The solid that crashed out was presumably DCU, since it is insoluble in most solvents.

Column chromatography was performed on the semi-purified product. The results of TLC analysis of the fractions revealed that the bottom band that hardly moved off the baseline was actually two spots superimposed. On the column, a dark yellow band began to form at the top of the column, but as it eluted, it separated into two bands. The lower band was light yellow while the upper band, which did not move off the top under these conditions, was orange in color. After the yellow band had completely eluted as observed by disappearance of color at the end of the column and by UV monitoring, the solvent system was changed to 10% methanol/ethyl acetate. The top band eluted quickly under these conditions. These fractions were evaporated to dryness and analyzed by ^1H NMR. It was determined that the yellow band was the desired product. While this mixture requires further purification, it is felt that the desired control monomer was synthesized successfully. Figures 65 and 66 compare the backbone starting material and the product.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

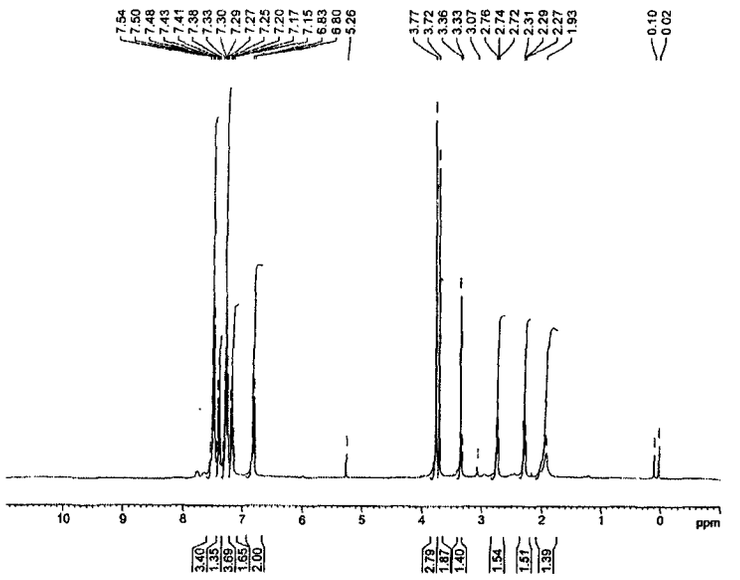


Figure 65: ^1H NMR (300MHz, $CDCl_3$) of protected backbone

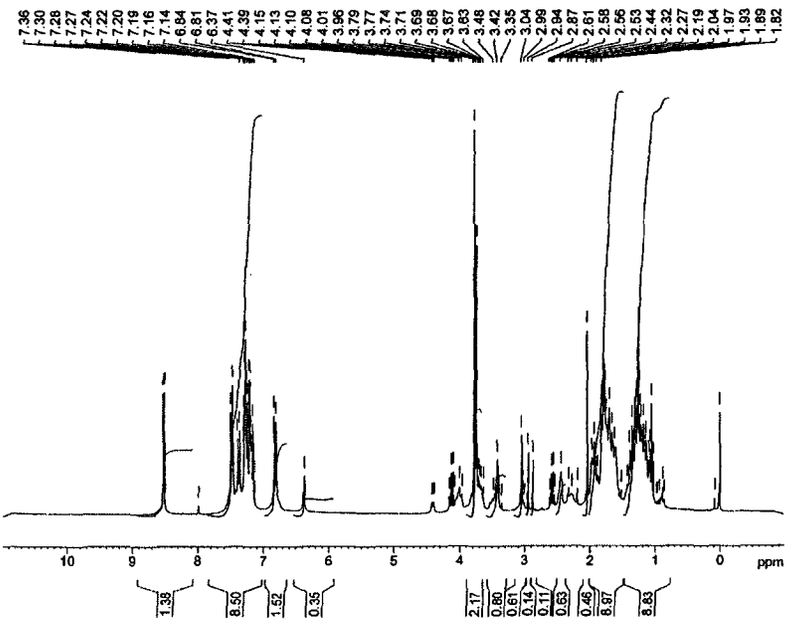


Figure 66: ^1H NMR (300MHx, $CDCl_3$) of crude reaction mixture for control monomer

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3.5.2 Synthesis of Monomer

In an analogous fashion to the peptide coupling for the control monomer, the Boc-protected carboxylated picoline was coupled to the backbone.

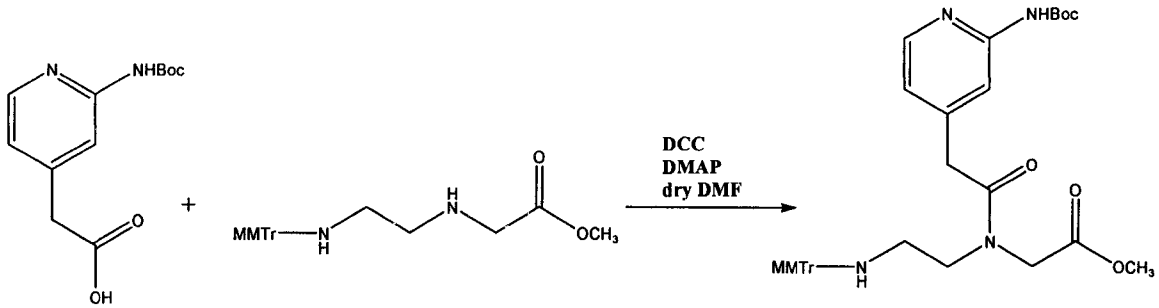


Figure 67: Reaction Scheme for Synthesis of Monomer

The same synthetic method was used except that there was no need to add one equivalent of triethylamine since the recognition unit was not in the form of a hydrochloride salt. The crude mixture revealed three bands by TLC analysis, but the ^{1}H NMR indicated possible presence of the desired monomer. The peak at 7.99 ppm in Figure 68 is not present in either starting material. It could be a peak for one of the heteroaromatic protons

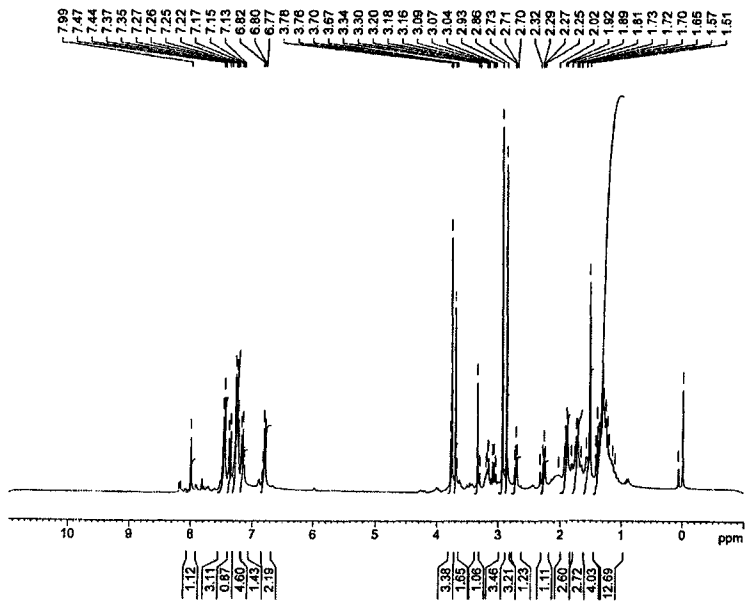


Figure 68: ^{1}H NMR (300MHz, $CDCl_3$) of crude mixture from synthesis of monomer

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

when coupled to PNA backbone.

3.5.3 HPLC of Crude Control Monomer

A sample of the crude control monomer was dissolved in 60% trimethylammonium acetate buffer (pH 8.4)/40% acetonitrile. The starting conditions for the HPLC were the same concentration as the solvent in which the control monomer was dissolved. A 10μL sample was injected and after one minute, the % of acetonitrile was increased in a gradient fashion to 70% after twenty minutes. The elution continued at that solvent ratio until 30 minutes had passed. The chromatogram revealed well-separated peaks.

For comparison purposes, a sample of backbone starting material and each recognition unit was run. The silica in the HPLC column used was coated with C18 chains making the column environment highly non-polar. Since the recognition units are carboxylic acids, their retention time was very short. The retention time for 4-pyridyl acetic acid (control recognition unit) was 1.2387 minutes, while for the carboxylated Boc-2-amino-4-picoline, retention time was 1.4107 minutes.

To obtain purified samples for mass spectrometry, preparative HPLC was performed. Concentrated samples of the crude control monomer were prepared in the initial solvent conditions. Fractions were collected while observing the increase and subsequent decrease in UV activity. Despite raising the concentration dramatically, good separation between peaks remained. Figure 69 shows the chromatogram from the preparative HPLC of the control monomer on top and the chromatogram from backbone starting material on the bottom.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The first large peak in the top chromatogram of Figure 69 appeared to be a completely new peak, different from any of the starting materials. The retention time of the second large peak was thought to be significantly different (0.3 minutes) from the similar peak observed in the chromatogram for the backbone. Fractions that were collected during the development of peaks 1 and 2 were evaporated to dryness using a vacuum centrifuge. The residue was sent for mass spectrometric analysis. The remaining peaks were assigned to backbone or were too small to warrant mass spectrometric analysis.

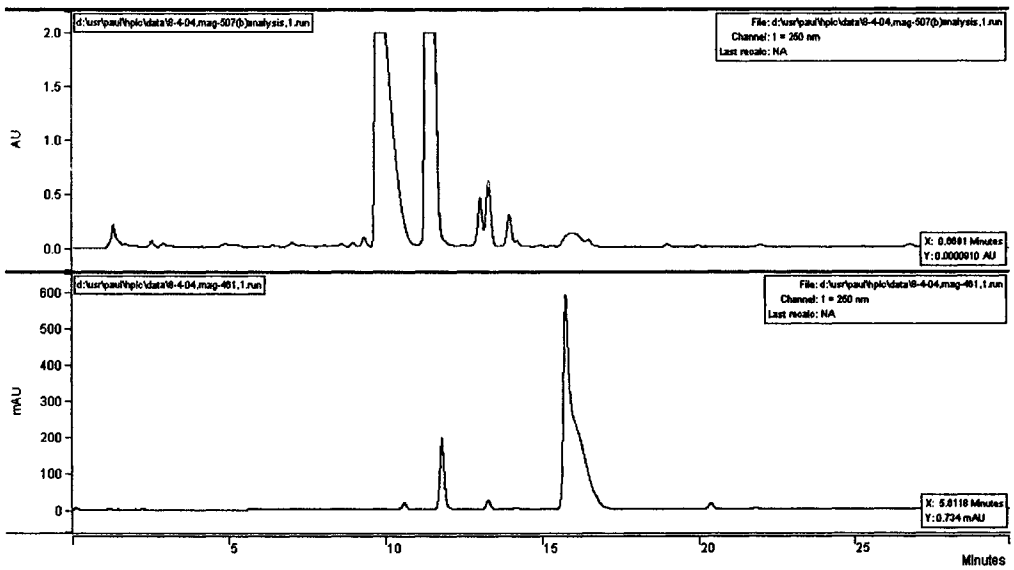


Figure 69: Comparison of chromatograms of crude control monomer (top) and backbone starting material (bottom)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3.5.4 HPLC of Crude Monomer

The same procedure was used to obtain a qualitative chromatogram of the crude monomer as was used for the crude control monomer. Good peak separation was observed using the same elution conditions.

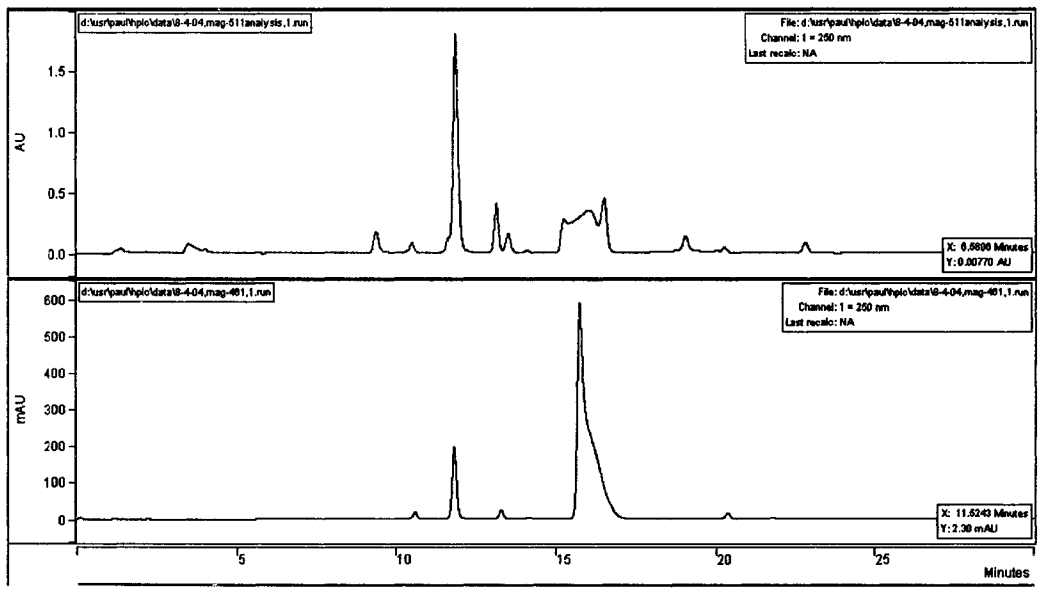


Figure 70: Comparison of chromatograms of crude monomer (top) and backbone starting material (bottom)

Preparative HPLC was performed on the crude monomer mixture. The first peak observed in the chromatogram for the crude monomer (9.374 minutes) appears to be a completely new peak, independent of all starting materials. Using the vacuum centrifuge, the solvent from the fraction was removed and the residue was sent for mass spectroscopic analysis. The peak at 11.8179 was possibly backbone but was collected for analysis none the less.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3.6 Conclusions & Future Work

By examining the groove of a G-quadruplex, a template has been designed to act as a recognition unit. Using PNA as a backbone to connect these recognition units, the designed oligomer was found to fit in the major groove of a G-quadruplex by semi-empirical AM1 calculations. A suitable monomer for the target oligomer was designed and synthetic routes towards those monomers have been explored. The recognition unit was synthesized *via* lithiation of Boc-2-amino-4-picoline and quenching the anion with carbon dioxide. A model protected monomer with no amino group has been successfully synthesized as well as the desired protected monomer through standard peptide coupling conditions.

The last step in the synthetic route to a protected monomer ready for solid-phase organic synthesis would have been to deprotect the carboxylic acid protecting group on the backbone terminus. This can be accomplished by using a variety of nucleophiles, however the use NH_4OH is particularly advantageous because the by-product of the reaction is ammonium hydroxide and excess ammonia will simply evaporate off. This greatly eases the process of purification in the step. The product can be extracted into an organic solvent, leaving the NH_4OH in the aqueous phase.

The monomer depicted in Figure 71 is the required monomer to begin solid-phase synthesis. Key features are the orthogonally protected amine groups. To orthogonally protect indicates that the cleavage conditions for one protecting group do not affect the other. In this case, the Boc protecting group is base labile while the MMTr protecting group is acid labile. When the first monomer is attached to a solid support, the MMTr group could be cleaved by 95% TFA in preparation for another coupling, while the Boc

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

group would remain unaffected. After the coupling of n units is complete, introduction of base into the system would simultaneously remove n Boc protecting groups.

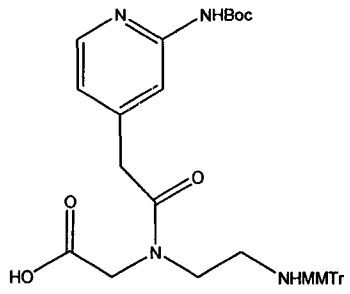
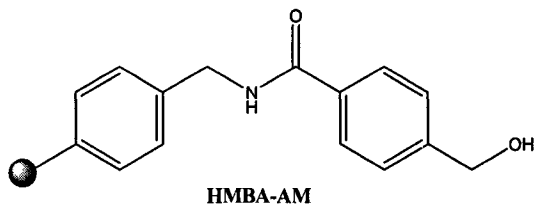


Figure 71: Structure of Required Monomer for Solid-Phase Synthesis

Since an acid labile protecting group had been chosen for the backbone amino group, a base labile resin must also be chosen. Otherwise, when one attempted to remove the MMTr group in order to couple again, the linker would be cleaved from the resin. In addition to being base labile, the resin must also be unaffected by the acid required to remove the Boc protecting group.

A suitable candidate for a resin was found to 4-hydroxymethyl benzoic acid AM polystyrene, or HMBA-AM, as shown in the following figure



HMBA-AM

Figure 72: Structure of Chosen Resin for Solid-Phase Synthesis

The grafted linker in this case is the 4-hydroxymethyl benzoic acid moiety. This linker is particularly versatile. By varying the type of nucleophile used for cleavage, one can obtain primary or secondary amines, alcohols and methyl esters.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

A common method for loading the resin with the first monomer is to form the symmetrical anhydride of the monomer from two equivalents of itself.

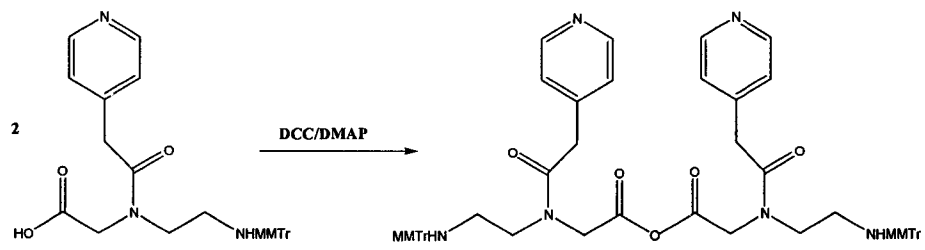


Figure 73: Synthesis of Symmetrical Anhydride for Solid-Phase Synthesis

This symmetrical anhydride is sufficiently active to couple to the resin through reaction of the terminal hydroxyl group to form an ester linkage. Treatment of the resin with 95% TFA removes the trityl group, preparing to couple another monomer to the growing chain. This coupling is done in the same manner as the coupling of the recognition unit to the backbone: with DCC/DMAP. After coupling to form a dimer, acetic anhydride is added to the resin. This is to cap any monomer that did not couple to another unit. Once a non-coupled terminus is capped, it will no longer be coupled to.

Once coupling has proceeded to the desired length, cleavage of the chain from the solid support is achieved through use of aqueous ammonia. Using ammonia as the nucleophile results in a terminal amide group. The final trityl group is not removed. Due

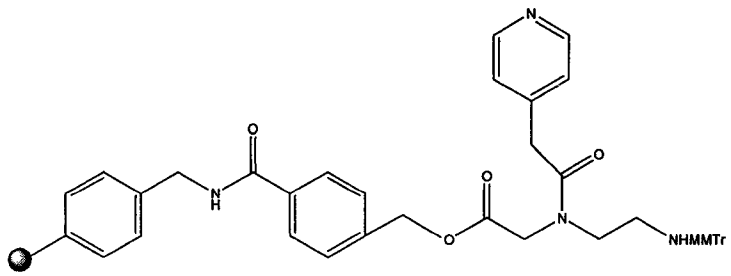


Figure 74: First Monomer attached to Solid-Phase Resin

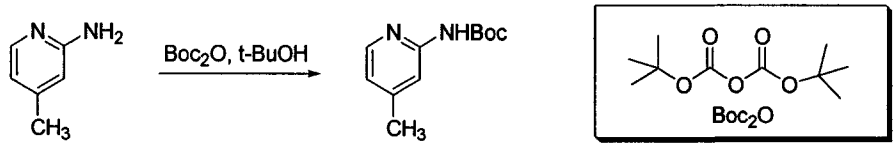
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

to incomplete coupling throughout the procedure, there will be chains of varying length that are cleaved from the resin, however only the chains that are the desired length will have the trityl group at their end. Purification by HPLC will be simple with this trityl group due to its high UV activity. Using the trityl protecting group in this way is known as the Trityl-ON purification.

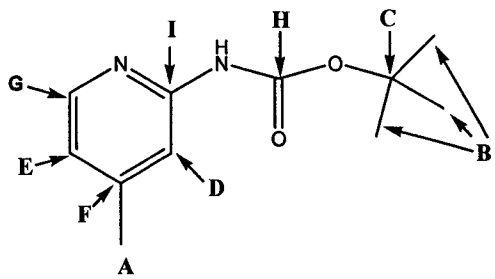
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

4. Experimental

4.1 2-[N-(*tert*-Butoxycarbonyl)amino]-4-picoline

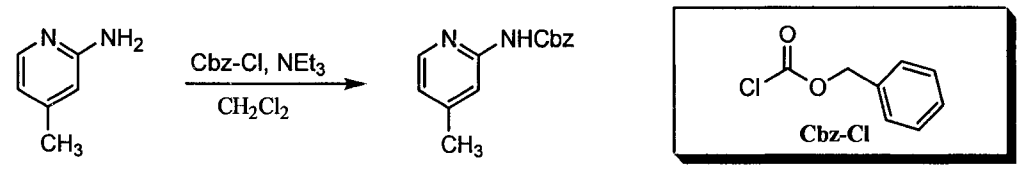


With good stirring, 2-amino-4-picoline (5.0363g, 46.57mmol) was added in small portions to 100mL of melted *tert*-butyl alcohol. After complete dissolution, the solution was pale yellow in color. Di-*tert*-butyl dicarbonate (11.0996g, 50.869mmol, 1.1eq.) was added to the solution, which was allowed to stir overnight. After 24 hours, the solution was bright yellow in color. Solvent was evaporated *in vacuo* yielding bright yellow solid. The residue was stable in the freezer for up to one month. Flash column chromatography (50 x 150mm, silica, $CHCl_3$) was performed. Two yellow bands were observed on the column: one dark yellow band did not move past the baseline, while the other lighter yellow band moved through the column. Rotary evaporation of the column product gave a yellow solid which was recrystallized from hot 10% ethyl acetate/hexane. Insoluble materials were removed by gravity filtration. After cooling at 4°C for 24 hours, filtration yielded 2.2449g (23.1%) of white crystals. ^1H NMR (300MHz, $CDCl_3$) δ 8.37 (br s, 1H), 8.11 (d, J = 5.1Hz, 1H), 7.83 (s, 1H), 6.79 (d, J = 5.1Hz, 1H), 2.35 (s, 3H), 1.54 (s, 9H). ^{13}C NMR (75MHz, $CDCl_3$) δ 21.62 (A), 28.58 (B), 80.93 (C), 113.04 (D), 119.68 (E), 147.56 (F), 149.56 (G), 152.67 (H), 153.05 (I). IR (KBr) 3206.9, 2985.6, 1723.3, 1576.8, 1292.8, 1231.4, 1158.3, 1120.0 cm^{-1}.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

4.2 Cbz-2-amino-4-picoline

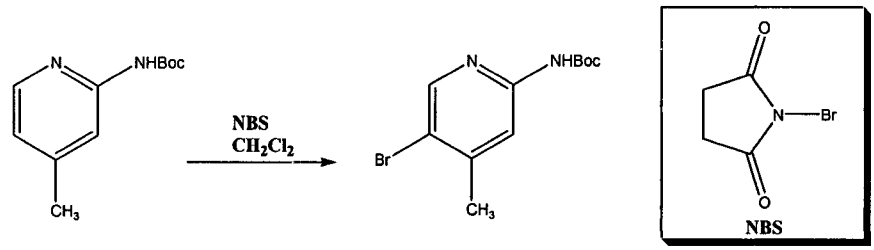


To approximately 20mL of stirred dichloromethane was added 2-amino-4-picoline (1.001g, 9.248mmol). Following complete dissolution, triethylamine (1.29mL, 9.248mmol) was added. The solution became pale yellow in color. The flask contents were cooled in an ice bath to 0°C. Using a syringe, benzyl chloroformate (1.3mL, 9.248mmol) was removed from a SureSeal© bottle and then added dropwise to the cooled solution. The flask contents were allowed to stir for one hour, at which point a white solid was observed floating on the top of a clear, yellow solution. This salt was removed (the product of triethylamine and the hydrochloric acid by-product) by vacuum filtration. The filtrate was concentrated *in vacuo* to give an oil initially, but yielded a yellow solid after twenty minutes. The crude material was found to be soluble in 5% CH_3OH/CH_2Cl_2 TLC analysis was performed and revealed six distinct spots. Methanol was used to recrystallize the crude material to yield a white powder. TLC analysis of the recrystallized material gave $R_f = 0.56$ and 0.24 (5% CH_3OH/CH_2Cl_2). Column chromatography (30 x 100mm, silica, 5% CH_3OH/CH_2Cl_2) was used to separate the two materials. After 120mL of solvent, the top peak was no longer visible and after an additional 60mL, the bottom peak could not be observed, as monitored by TLC. The column products in solution were concentrated *in vacuo* to yield two white solids. ^1H NMR Solid A (300MHz, $CDCl_3$) δ 8.30 (br s, 1H), 8.22 (d, $J = 6$ Hz, 1 H), 8.07 (d, $J = 6$ Hz, 1H), 7.84 (s, 1H), 7.43-7.33 (m, 5H), 6.82 (d, $J = 6$Hz, 1H), 6.78 (d, $J = 6$Hz, 1H),

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

5.31 (s, 1H), 5.23 (s, 2H), 2.36 (s, 7H), 2.18 (s, 1H). ^{1}H NMR Solid B (300MHz, $CDCl_3$) δ 8.08 (d, J = 6 Hz, 1H), 7.94 (br s, 1H), 7.83 (s, 1H) , 7.43-7.37 (m, 5H), 6.79 (d, J = 6 Hz, 1H), 5.23 (s, 2H), 2.36 (s, 3H).

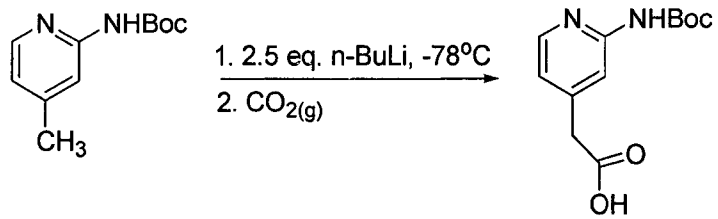
4.3 2-[N-(*tert*-Butoxycarbonyl)amino]-5-bromo-4-picoline



2-[N-(*tert*-Butoxycarbonyl)amino-4-picoline (0.1080g, 0.519mmol) was dissolved in 25mL of dry CH_2Cl_2. NBS (0.0933g, 0.52mmol) was added to the solution. The reaction mixture was refluxed for 4 hours with a $CaCl_2$ drying tube attached to the condenser. After 15 minutes of reflux, the reaction mixture was dark orange in color, but after an additional 15 minutes, the mixture had turned bright yellow. For the remainder of the reflux period, the yellow color decreased in intensity. TLC analysis of the final reaction mixture revealed three bands: R_f = 0, 0.11 and 0.32 (95% hexane/ethyl acetate). Comparative TLC analysis revealed that the starting material had R_f = 0.11 and NBS had R_f = 0. Column chromatography (30 x 100mm, silica, 95% hexane/ethyl acetate to 93% hexane/ethyl acetate) was performed. Approximately 200mL of 95% hexane/ethyl acetate was eluted through column to obtain the R_f = 0.32 band. The solvent was removed by rotary evaporation to yield a white powder. ^{1}H NMR (300MHz, $CDCl_3$) δ 8.25 (s, 1H), 7.92 (s, 1H), 2.40 (s, 3H), 1.54 (s, 9H).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

4.4 2-[N-(*tert*-butoxycarbonyl)amino-4-pyridyl acetic acid

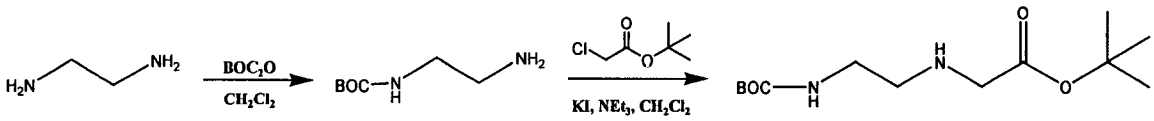


All glassware was dried overnight in a 110°C oven prior to the experiment. In 30 mL of dry THF, 2-[N-(*tert*-Butoxycarbonyl)amino]-4-picoline (1.0026g, 4.814mmol) was dissolved with stirring. The flask was flushed with nitrogen gas. The solution was cooled to -78°C through the use of a dry ice and acetone mixture. *n*-butyl lithium (1.6M in hexanes, 7.5 mL, 12.035mmol) was added *via* syringe over a two minute period. A change from colorless to yellow was observed during the addition. The flask contents were allowed to stir for two hours in a cooling bath. The contents of two large balloons filled with dry carbon dioxide were bubbled through the cooled solution *via* a long needle, where upon a milky precipitate was immediately formed. Thirty minutes after addition, the flask was allowed to warm to room temperature under an atmosphere of CO_2. The reaction was quenched using H_2O. The solvent was removed *in vacuo* to yield a chalky white residue. This residue was taken up in 40mL 1M NaOH. This mixture had a yellow hue. Hexane (40mL) was added and after vigorous shaking, the layers were separated. The organic layer was washed with 40mL of 1M NaOH. The aqueous layers were combined and the pH of the solution was adjusted from 12.5 to 4.5 using 2M HCl. The flask contents changed from cloudy yellow to clear and colorless. The solvent was concentrated *in vacuo* to approximately 75% of its original volume. A small amount of solid was observed in the solution. The solution was cooled to 4°C for 72 hours. After this time, floating balls of white solid were observed. The flask contents were filtered to

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

yield a white solid (0.1079g, 8.87%). ^{1}H NMR (300MHz, D_2O) δ 8.17 (d, J = 5.7 Hz, 1H), 7.44 (s, 1H), 7.14 (d, J = 5.7 Hz, 1H), 3.62 (s, 2H), 1.54 (s, 9H) ^{13}C NMR (75MHz, d6-DMSO) δ 152.69, 152.37, 147.31, 147.14, 119.66, 113.04, 79.49, 28.06. IR (KBr) 1727.8, 1700.0, 1577.8, 1538.9, 1433.3 cm^{-1}.

4.5 Boc protected backbone

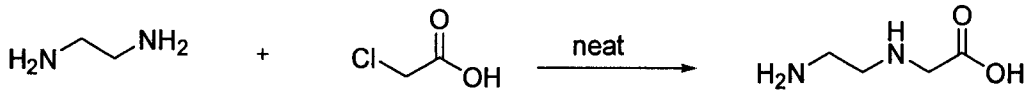


Ethylenediamine (5.6mL, 83.8mmol) was added to 20mL of dichloromethane. The solution was cooled in an ice bath. Di-tertbutoxydicarbonate (2.2713g, 10.4mmol) was dissolved in approximately 20mL of dichloromethane and was added dropwise to the reaction mixture over a period of 30 minutes. The cooling bath was removed after addition was complete and the reaction was left to stir for 48 hours. The reaction mixture was extracted with distilled water (10mL). The layers were separated and the aqueous layer was washed with dichloromethane (5mL). The organic layers were combined and dried using Na_2SO_4. The organic phase was then concentrated to one-third of its original volume. Triethylamine (1.3mL, 9.33mmol) and potassium iodide (0.0067g, 0.0404mmol) were added to the concentrated solution. In dichloromethane (5mL), *tert*-butyl chloroacetate (1.5mL, 11.7mmol) was dissolved and the solution was added dropwise to the reaction mixture over 5 minutes. The reaction mixture was refluxed for 48 hours. The reaction mixture became brown after approximately 24 hours, and was orange in color after 48 hours. The solution was extracted with distilled water (6mL). The organic phase was dried with Na_2SO_4 and was evaporated to dryness to yield a bright orange oil. Column chromatography (30 x 100mm, silica, 5% MeOH/EtOAc) removed polar

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

impurities and evaporation of the eluent gave a pale yellow oil. TLC analysis revealed two bands (R_f = 0.95 and 0.50). The TLC was visualized by treatment with 5% TFA/CH_2Cl_2 followed by ninhydrin and heating. Column chromatography (30 x 100mm, silica, 5% MeOH/EtOAc) was performed and the two bands were successfully separated. The band with R_f = 0.50 was found to be the desired product by NMR. ^{1}H NMR (300MHz, $CDCl_3$) δ 5.24 (br s, 1H), 3.32 (s, 2H), 3.12 (m, 2H), 2.64 (t, J = 6Hz, 2H), 1.88 (br s, 1H), 1.38 (s, 9H), 1.35 (s, 9H). ^{13}C NMR (75MHz, $CDCl_3$) δ 171.82, 156.18, 81.30, 79.0, 53.41, 51.65, 40.28, 28.48, 28.15.

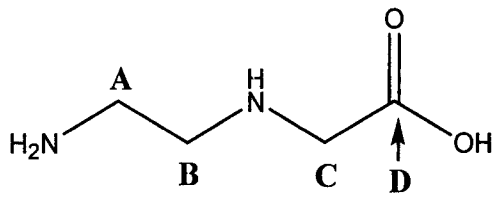
4.6 N-(2-aminoethyl)glycine



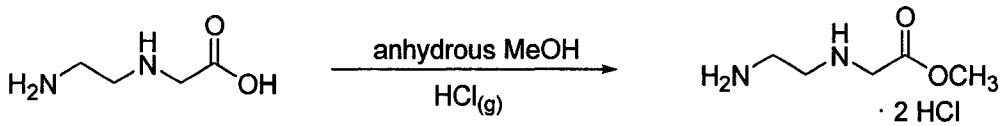
Ethylenediamine (70.1mL, 1.06mol) was cooled in a cold water bath. Chloroacetic acid (9.9936g, 0.106mol) was ground using a mortar and pestle, then was added in small portions to the ethylenediamine with good stirring. The reaction mixture was left to stir overnight. Excess ethylenediamine was removed under high vacuum with the aid of a 60°C water bath. The residue was thick, viscous and cloudy. Approximately 270mL of DMSO were added to the residue to triturate. The mixture was left undisturbed for 10 days in a fumehood. A large quantity of white solid had formed. The mixture was filtered to yield large, thick white flakes of solid. The solid was washed with ~30mL of DMSO, then with ~200mL of ether. The crude product had a mass of 8.3046g. The crude product was dissolved in approximately 30mL of 60°C distilled water. Approximately 100mL of absolute ethanol was added to precipitate the product. The mixture was cooled at 4°C overnight to produce crystalline white solid. The mixture was vacuum filtered and the crystals were dried *in vacuo* to yield 7.8064g (62.5%). ^{1}H NMR (300MHz, D_2O) δ 3.28

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(s, 2H), 3.04 (t, J = 6 Hz, 2H), 2.91 (t, J = 6 Hz, 2H) ^{13}C NMR (75MHz, D_2O) δ 38.54 (A), 46.49 (B), 51.52 (C), 178.26 (D).



4.7 N-(2-aminoethyl)glycine methyl ester · 2HCl



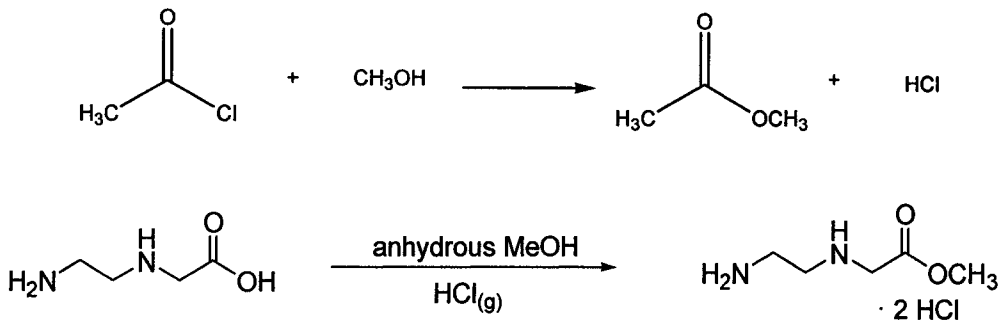
THIS EXPERIMENT MUST BE CONDUCTED IN AN EFFICIENT FUMEHOOD. Glassware was dried overnight in a 110°C oven prior to the experiment. Vacuum dried N-(2-aminoethyl)glycine (2.6362, 22.3mmol) was dissolved in 250mL of anhydrous methanol with stirring. The reaction flask was sealed with a rubber septum. Compressed hydrogen chloride gas was used in this experiment. The contents of five balloons full of HCl gas were bubbled through the reaction mixture *via* a long needle through a septum. A second, shorter needle was placed in the septum to equalize the pressure. Bubbling of the gas caused immediate formation of a precipitate, which then all re-dissolved. Precipitation was again observed approximately five minutes later, followed by the same re-dissolution. A final precipitation was observed after another five minutes however the precipitate did not re-dissolve. The flask contents were then refluxed for one hour under an atmosphere of HCl gas. The cloudiness diminished as time progressed; after 45 minutes, the solution was totally clear. The flask contents were allowed to cool to room temperature, then the flask was cooled to 4°C to allow crystallization of product. After 72

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

hours, the flask contents were vacuum filtered to yield a fluffy, white solid. The solid was washed with cold, anhydrous methanol, then with cold ether. The product was dried under vacuum to yield 2.3889g (55.6%) ^{1}H NMR (300MHz, D_2O) δ 4.11 (s, 2H), 3.84 (s, 3H), 3.54-3.40 (m, 4H) IR (KBr) 1747.9, 1558.7, 1247.3

4.8 N-(2-aminoethyl)glycine methyl ester · 2HCl (*in situ* generation of $HCl_{(g)}$)

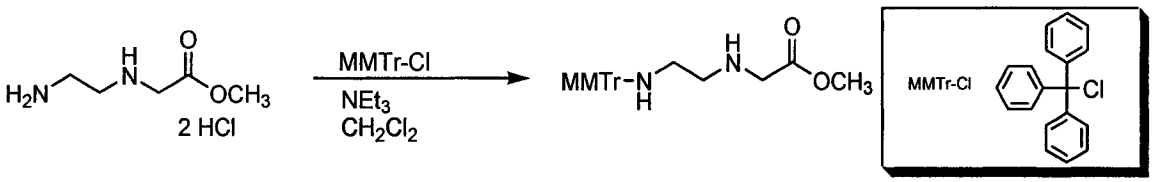
THIS EXPERIMENT MUST BE CONDUCTED IN AN EFFICIENT FUMEHOOD



Acetyl chloride (24.0mL, 338.4mmol) was added in small portions to stirred, anhydrous methanol (500mL). **CAUTION: HEAT EVOLVED!** The mixture was allowed to stir for 2 hours. Vacuum dried N-(2-aminoethyl)glycine (5.0027g, 42.3mmol) was dissolved in 100mL of anhydrous methanol and added to the reaction mixture. A milky, white suspension was formed. The mixture was refluxed for 1 hour, after which time the reaction mixture was clear and colorless. The reaction flask was cooled to room temperature and 200mL of ether was added, causing white solid to precipitate out. The mixture was stored at 4°C overnight, and was then filtered to yield 5.3400g (65.4%) of product. ^{1}H NMR (300MHz, D_2O) δ 4.14 (s, 2H), 3.87 (s, 3H), 3.77 (methyl acetate impurity), 3.56 – 3.47 (m, 2H), 3.44 (methyl acetate impurity). ^{13}C NMR (75MHz, D_2O) δ 170.79, 167.87 (methyl acetate impurity), 53.89, 49.52 (methyl acetate impurity), 47.95, 44.37, 35.79 (methyl acetate impurity), 35.66.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

4.9 Methyl N-[2-(N-monomethoxytrityl)aminoethyl]glycinate

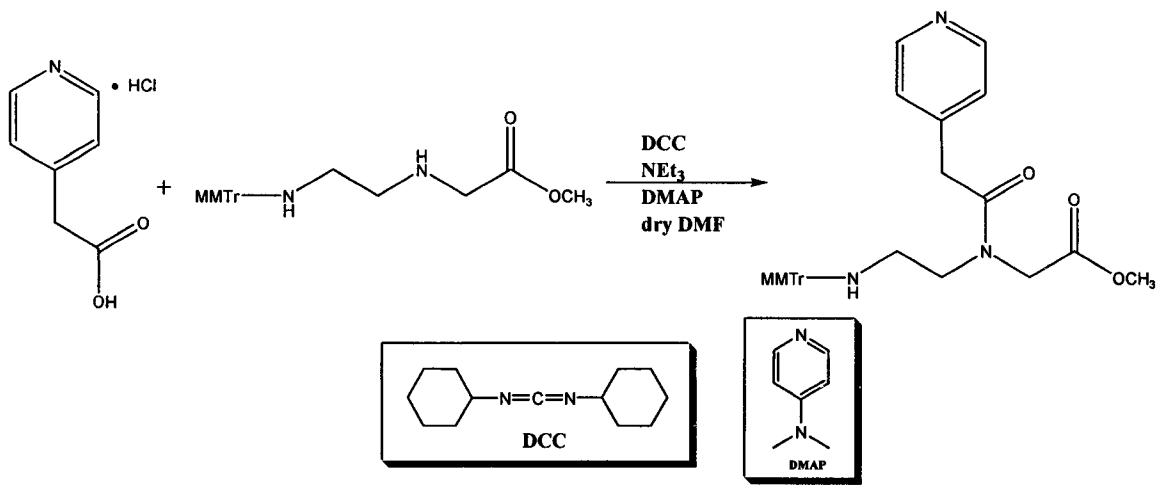


Vacuum dried N-(2-aminoethyl)glycine methyl ester · 2HCl (299mg, 1.549mmol) was added to ~25mL of CH_2Cl_2 and the mixture was cooled in an ice bath. Upon addition of triethylamine (650μL, 4.664mmol), the cloudy solution cleared. A solution of *p*-anisylchlorodiphenylmethane (MMTr-Cl) (0.4808g, 1.557mmol) in 10mL CH_2Cl_2 was added dropwise to the cooled reaction mixture with stirring. The reaction progress was monitored over two hours. MMTr-Cl was run as a comparison, $R_f = 0.81$ (50% hexane/ethyl acetate). Five minutes after MMTr-Cl addition was complete, two spots (R_f = 0.22 and 0.05) were observed, as well as the spot for MMTr-Cl. After two hours of stirring, the solvent was removed *in vacuo*. The residue was taken up in CH_2Cl_2 and washed three times with 10% potassium bicarbonate, then with brine three times. The organic layer was dried with sodium sulfate, then filtered by gravity. The solvent was removed by rotary evaporation to give the crude product. Column chromatography (30 x 100mm, silica, 50% hexane/ethyl acetate/0.5% TEA) was performed on the crude material. Fractions with a volume of 10mL were collected and monitored by UV. Approximately 50mL of solvent was sufficient to elute the top spot (MMTr-Cl), while another 150mL were required to elute the bottom spot. The solvent was evaporated from the second elution by rotary evaportation to yield an amorphous solid (0.1690g, 27.0%). ^{1}H NMR (300MHz, $CDCl_3$) δ 7.48 (d, J = 6 Hz, 4H), 7.39 (d, J = 9 Hz, 2H), 7.28 (d, J = 15 Hz, 4H), 7.18 (dd, J = 6 Hz, 2H), 6.82 (d, J = 9 Hz, 2H), 3.81 (s, 3H), 3.73 (s, 3H),

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3.36 (s, 2H), 2.74 (t, J = 6 Hz, 2H), 2.28 (t, J = 6 Hz, 2H), 1.87 (br s, 1H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 173.25, 158.01, 146.58, 138.48, 130.01, 129.64, 128.77, 126.34, 113.27, 70.41, 55.35, 51.90, 50.74, 49.97, 43.25. IR (cast) 3326.8, 3056.1, 2951.8, 2836.1, 1740.0, 1607.1, 1507.2.

4.10 Synthesis of Control Monomer



Dry DMF (~75mL) was added to 4-pyridyl acetic acid hydrochloride (524.6mg, 3.02mmol) with stirring. Triethylamine (420μL, 3.02mmol) was added to the mixture. DMAP (152mg, 0.302mmol) was added. DCC (622.6mg, 3.02mmol) was dissolved in CH_2Cl_2 and then added slowly to the reaction mixture. The color changed to pale yellow from colorless. The reaction mixture was allowed to stir overnight. Protected backbone (616.1mg, 1.52mmol) was dissolved in CH_2Cl_2 and added to the flask. The reaction mixture was allowed to stir for four hours. The CH_2Cl_2 was removed from the solution by rotary evaporation. The DMF solution was taken up in ethyl acetate and washed with water twice. The organic phase was then washed with saturated potassium bicarbonate twice and brine twice. IMPORTANT NOTE: all of the washing solvents contained 0.5% triethylamine for protection of the trityl group. The organic phase was dried with $MgSO_4$,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

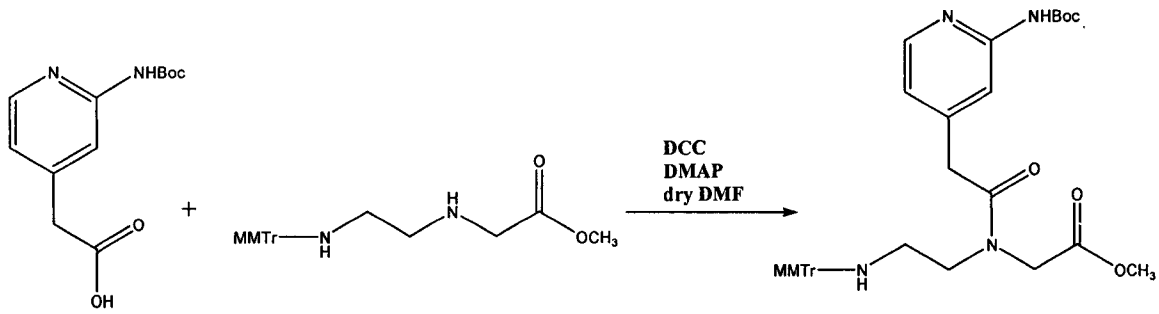
then was evaporated to dryness. This oily residue was taken up in 80% hexane/ethyl acetate/0.5% NEt_3 and a white solid crashed out. This was filtered to give the chromatography solution. Column chromatography (15 x 120mm, silica, 80% hexane/ethyl acetate/0.5% TEA) was performed and three separate UV active compounds were isolated. Each was characterized by ^1H NMR and the yellow band was found to contain the desired product. HPLC (150 x 4.60mm, Varian C18, 90Å, 99007-6) was performed on crude product. Solvent A was trimethylammonium acetate buffer, pH 8.4 and solvent B was HPLC-grade acetonitrile. Starting conditions were 60% A/ 40% B. After one minute, a gradient elution was begun over twenty minutes, finishing at 30% A/70% B. Elution was continued for 10 minutes after finishing the gradient elution.

Peak	Retention Time (mins)	Relative Absorbance (AU)
*1	9.8427	2+
*2	11.4831	2+
3	12.9896	0.430
4	13.2575	0.592
5	13.9605	0.289
6	16.0027	0.139

Table 5: Retention times and relative absorbances for crude control monomer (* sent for mass spectral analysis)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

4.11 Synthesis of Monomer



Dry DMF (~15mL) was used to dissolve 2-Boc-amino-4-pyridyl acetic acid (32.8mg, 0.130mmol) in an oven dried flask. DMAP (4.4mg, 0.0360mmol) was added to the solution. DCC (26.5mg, 0.130mmol) was dissolved in CH_2Cl_2 and was added slowly to the stirred solution. The solution was allowed to stir at room temperature overnight. Protected backbone (23.6mg, 0.065mmol) dissolved in CH_2Cl_2 was added and the solution was allowed to stir for 4 hours. The CH_2Cl_2 was removed from solution by rotary evaporation. The remaining solution was stored at 4°C overnight. The solution was then combined with an approximately equal volume of ethyl acetate. This mixture was washed twice with distilled water (0.5% NEt_3), twice with saturated potassium bicarbonate (0.5% NEt_3) and twice with brine (0.5% NEt_3). CAUTION: Addition of NEt_3 may reverse the relative densities of the organic and aqueous layers. The organic layer was dried with magnesium sulphate. The yellow solution was then evaporated to dryness to yield an oil. HPLC (150 x 4.60mm, Varian C18, 90Å, 99007-6) was performed on the crude product. Solvent A was trimethylammonium acetate buffer, pH 8.4 and solvent B was HPLC-grade acetonitrile. Starting conditions were 60% A/ 40% B. After one minute, a gradient elution was begun over twenty minutes, finishing at 30% A/70% B. Elution was continued for 10 minutes after finishing the gradient elution.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Peak	Retention Time (mins)	Relative Absorbance (AU)
*1	9.3740	0.177
2	10.4787	0.0866
*3	11.8179	1.80
4	13.0901	0.399
5	13.4981	0.158
6 (broad)	16.5049	0.441
7	19.057	0.142
8	22.8992	0.0538

Table 6: Retention times and relative absorbances for control monomer (* sent for mass spectral analysis)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

References

1. Gellman, S. *Chem. Rev.* **1997**, 97, 1231-1232.

2. Szejili, J. *Chem. Rev.* **1998**, *98*, 1743-1753.

3. Rebek, J.; Nemeth, D. *J. Am. Chem. Soc.* **1986**, *108*, 5637-5638.

4. Uhlmann, E.; Peyman, A. *Chem. Rev.* **1990**, *90*, 544-584.

5. http://www.liv.ac.uk/~giles/theory/rnaseh_cleavage.htm, May 25, 2004.

6. Saenger, W. (1984) *Principles of Nucleic Acid Structure*, John Wiley, New York.

7. De Mesmaeker, A.; Häner, R.; Martin, P.; Moser, H. *Acc. Chem. Res.* **1995**, *28*, 366-374.

8. De Mesmaeker, A.; Altmann, K-H.; Waldner, A.; Wendeborn, S. *Curr. Opin. Struct. Biol.* **1995**, *5*, 343-355.

9. Egholm, M.; Nielsen, P.; Buchardt, O.; Berg, R. *J. Am. Chem. Soc.* **1992**, *114*, 9677-9678.

10. Nielsen, P.; Egholm, M.; Berg, R.; Buchardt, O. *Science* **1991**, *254*, 1497-1500.

11. Ray, A.; Nordén, B. *FASEB.* **2000**, *14*, 1041-1060.

12. Nielsen, P. *Acc. Chem. Res.* **1999**, *32*, 624-630.

13. Garrett, R.; Grisham, C. (1999) *Biochemistry*, Saunders College Publishing, 2nd ed.

14. Gellert, M.; Lipsett, M.; Davies, D. *Proc. Natl. Acad. Sci. USA* **1962**, *48*, 2013-2018.

15. Donahue, J. *Proc. Natl. Acad. Sci. USA* **1956**, *42*, 60-65.

16. Pinnavaia, T.; Miles, H.; Becker, E. *J. Am. Chem. Soc.* **1975** *97*, 7198-7200.

17. Pinnavaia, T.; Marshall, C.; Mettler, C.; Fisk, C.; Miles, H.; Becker, E. *J. Am. Chem. Soc.* **1978**, *100*, 3625-3627.

18. Davis, J. *Angew. Chem. Int. Ed.* **2004**, *43*, 668-698.

19. Smith, F.; Feigon, J. *Biochemistry* **1993**, *32*, 8682-8692.

20. Patel, D.J.; Bouaziz, S.; Kettani, A.; Wang, Y. in Neidle, S., Ed. (1999) *Oxford Handbook of Nucleic Acid Structure*, Oxford Science Publications, New York.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

21. McEachern, M.J.; Krauskopf, A.; Blackburn, E.H. *Annu. Rev. Genet.* **2000**, *34*, 331-358.

22. Rezler, E.M.; Bearss, D.J.; Hurley, L.H. *Curr. Op. Pharm.* **2002**, *2*, 415-423.

23. Sun, D.; Thompson, B.; Cathers, B.; Salazar, M.; Kerwin, S.; Trent, J.; Jenkins, T.; Neidle, S.; Hurley, L. *J. Med. Chem.* **1997**, *40*, 2113-2116.

24. Perry, P.; Gowan, S.; Reszka, A.; Polucci, P.; Jenkins, T.; Kelland, L.; Neidle, S. *J. Med. Chem.* **1998**, *41*, 3253-3260.

25. Agbandje, M.; Jenkins, T.; McKenna, R.; Reszka, A.; Neidle, S. *J. Med. Chem.* **1992**, *35*, 1418-1429.

26. Tanious, F.; Jenkins, T.; Neidle, S.; Wilson, W. *Biochemistry*, **1992**, *31*, 11632-11640.

27. Perry, P. Reszka, A.; Wood, A.; Read, M.; Gowan, S.; Dosanjh, H. Trent, J.; Jenkins T.; Kelland, L.; Neidle, S. *J. Med. Chem.* **1998**, *41*, 4873-4884.

28. Harrison, R.; Gowan, S.; Kelland, L.; Neidle, S. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 2463-2468.

29. Read, M.; Harrison, J.; Romagnoli, B.; Tanious, F.; Gowan, S.; Reszka, A.; Wilson, W.; Kelland, L.; Neidle, S. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 4844-4849.

30. Haider, S.; Parkinson, G.; Neidle, S. *J. Mol. Biol.* **2003**, *326*, 117-125.

31. Anantha, N.; Azam, M.; Sheardy, R. *Biochemistry* **1998**, *37*, 2709-2714.

32. Wheelhouse, R.; Sun, D.; Han, H.; Han, F.; Hurley, L. *J. Am. Chem. Soc.* **1998**, *120*, 3261-3262.

33. Shin-ya, K.; Wierzba, K.; Matsuo, K.; Ohtani, T.; Yamada, Y.; Furihata, K.; Hayakawa, Y.; Seto, H. *J. Am. Chem. Soc.* **2001**, *123*, 1262-1263.

34. Kim, M-Y.; Vankayalapati, H.; Shin-ya, K.; Wierzba, K.; Hurley, L. *J. Am. Chem. Soc.* **2002**, *124*, 2098-2099.

35. Sheehan, J.; Hess, G. *J. Am. Chem. Soc.* **1955**, *77*, 1067-1068.

36. Han, S.; Kim, Y. *Tetrahedron* **2004**, *60*, 2447-2467.

37. Merrifield, R.B. *J. Am. Chem. Soc.* **1963**, *85*, 2149-2154.

38. Guillier, F.; Orain, D.; Bradley, M. *Chem. Rev.* **2000**, *100*, 2091-2157.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

39. http://www.ccl.net/cca/documents/dyoung/topics-orig/compchem.html, July 10, 2004.

40. http://www.netsci.org/Science/Compchem/feature01.html, July 10, 2004

41. Hehre, W. (2003) *A Guide to Molecular Mechanics and Quantum Chemical Calculations*. Wavefunction, Inc.

42. PDB Crystal Stucture.

43. Chichibabin, A.E. *J. Russ. Phys. Chem. Soc.* **1905**, *37*, 1229.

44. Kröhnke, F.; Zecher, W. *Angew. Chem. Int. Ed.* **1962**, *1*, 626-632.

45. Kröhnke, F. *Synthesis* **1976**, 1-24.

46. Vijn, R.; Arts, H.; Maas, P.; Castelijns, A. *J. Org. Chem.* **1993**, *58*, 887-891.

47. Takahashi, T.; Tsai, F.-Y.; Li, Y.; Wang, H.; Kondo, Y.; Yamanaka, M.; Nakajima,

K.; Kotora, M. *J. Am. Chem. Soc.* **2002**, *124*, 5059-5067.

48. Deasy, C. *J. Org. Chem.* **1945**, *10*, 141-144.

49. Ihle, N.; Krause, A. *J. Org. Chem.* **1996**, *61*, 4810-4811.

50. Beyeler, A.; Belser, P.; De Cola, L. *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 2779-2781.

51. Silverstein, R.; Webster, F. (1998) *Spectroscopic Identification of Organic Compounds*, 6th ed., John Wiley & Sons, Inc.

52. Greene, T.; Wuts, P. (1999) *Protective Groups in Organic Synthesis*, 3rd ed., John Wiley & Sons, Inc.

53. Kamiya, T.; Tsutomu, T.; Nakai, Y.; Sakane, K.; Goto, J. Patent DE 2848912, 1979.

54. March, J. (1992) *Advanced Organic Chemistry: Reactions, Mechanisms, and Structure*, 3rd ed., John Wiley & Sons, Inc.

55. Djerassi, C. *Chem. Rev.* **1948**, *43*, 271-317.

56. Stetsenko, D.; Lubyako, E.; Potapov, V.; Axhikina, T.; Sverdlov, E. *Tetrahedron Lett.* **1996**, *37*, 3571-3574.

57. Will, D.; Breipohl, G.; Langner, D.; Knolle, J.; Uhlmann, E. *Tetrahedron* **1995**, *51*, 12069-12082.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

58. Breipohl, G.; Will, D.; Peyman, A.; Uhlmann, E. *Tetrahedron* **1997**, *53*, 14671-14686.

59. Krapcho, A.; Kuell, C. *Synth. Commun.* **1990**, *20*, 2559-2564.

60. personal communication, Dr. Robert H.E. Hudson, University of Western Ontario, August 19, 2003.

61. Heimer, E; Gallo-Torres, H.; Felix, A.; Ahmad, M.; Lambros, T.; Scheidl, F.; Meienhofer, J. *J. Peptide Protein Res.* **1984**, *23*, 203-211.

62. Di Giorgio, C.; Pairot, S.; Schwergold, C.; Patino, N.; Condom, R.; Farese-Di Giorgio, A.; Guedj, R. *Tetrahedron* **1999**, *55*, 1937-1958.

63. Xiao, J.; Yuan, G.; Huang, W.; Chan, A.; Lee, K.-L. *J. Org. Chem.* **2000**, *65*, 5506-5513.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.