4-27-2011 12:00 AM

Design and Syntheses of Fluorescent Cytosine Analogues

David W. Dodd
The University of Western Ontario

Supervisor
Dr. Robert H.E. Hudson
The University of Western Ontario

Graduate Program in Chemistry
A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy

DESIGN AND SYNTHESES OF FLUORESCENT CYTOSINE ANALOGUES

(Spine title: Design and Syntheses of Fluorescent Cytosine Analogues)

(Thesis format: Monograph)

by

David W. Dodd

Graduate Program in Chemistry

A thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

THE UNIVERSITY OF WESTERN ONTARIO
School of Graduate and Postdoctoral Studies

CERTIFICATE OF EXAMINATION

Supervisor

Dr. Robert H.E. Hudson

Supervisory Committee

Dr. Richard Pudephatt

Dr. Lars Konerman

Examiners

Dr. Leonard G. Luyt

Dr. Brian L. Pagenkopf

Dr. Masad J. Damha

Dr. Gilles Lajoie

The thesis by

David William Dodd

entitled:

Design and Syntheses of Fluorescent Cytosine Analogues

is accepted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Chair of the Thesis Examination Board

Abstract

The avid hybridization of peptide nucleic acid (PNA) to DNA and RNA along with the molecule's biological stability has led it to be used in both antisense and antigene capacities. PNA acts against translation *via* a steric blockade mechanism. It is therefore reasonable to assume that increased heteroduplex stability could lead to increased potency. Two ways of doing this were explored. I) *N*-Terminal attachment of a platinous chloride chelating moiety to PNA complementary to *Xenopus* noggin was synthesized with the objective of selective, covalent platination of the target transcript *in vivo*. Phenotypes consistent with knockdown of the selected gene product were observed; however, it remains unclear whether this is solely due to site-specific platination. II) The use of the modified cytosine nucleobase [bis-*o*-(aminoethoxy)phenyl]pyrrolocytosine (boPhpC) was previously shown to increase heteroduplex stability, putatively by interaction with guanine at both the Watson-Crick and Hoogsteen faces. PNAs containing this base were synthesized to target mutant huntingtin mRNA - of which the product is the causative agent of Huntington's disease – and tested in patient derived fibroblasts where selective inhibition of mutant huntintin was observed with concomitant fluorescence imaging.

Modified nucleobases find use in fields ranging from materials science to cytogenetics and has been an area of much endeavour over the past years. Modifications at the 5-position of uracil abound but examples of similarly modified cytosine are lacking. Rapidly developing the inventory of cytosine analogues is a primary goal of ours. We aim to synthesize compounds that may be used as base discriminating fluorophores, created pre- or post-synthetically. Through the reaction of 5-ethynyldeoxycytidine with 1,3-dipoles such as nitrile oxides or azides a series of heterocycle-appended cytosine analogues have been prepared and their fluorescence properties studied. They exhibit moderate to good quantum yields with high sensitivity to their environment and are considered good candidates for further use as base-discriminating fluorophores. Based on the known pyrrolocytosine scaffold, Indole-3-acetamide substituted deoxypyrrolocytidine (IAMpC) has shown the highest degree of solvatochromism for any pyrrolocytosine analogue known to date. The synthesis of oligodeoxynucleotides containing this base was carried out; however, final deprotection of the base does not proceed smoothly and the modification was found to be

slightly destabilizing towards duplex formation.

Keywords

DNA, PNA, mRNA, Platinum (II), Cisplatin, PAGE, Base-discriminating Fluorophore, Copper-catalyzed Azide/Alkyne Cycloaddition, Antisense, Huntington's Disease, Nucleoside Analogue, Hydrogelator.

Co-Authorship Statement

Contained in the following texts are notable contributions from the following individuals. Dr. Robert H.E. Hudson was behind the design of all projects and edited writing. Dr. Nathan D. Jones contributed to the design of projects contained in chapters 1 and 5. Former members of the Jones group, including Kalen N. Swanick, Jaquelyn T. Price and Alison Brazeau, contributed to work outlined in chapter 5. Dr. David R. Corey and Dr. Jiaxin Hu of UT Southwestern Medical Center and Dr. Filip Wojciechowski, formerly of the Hudson Group, contributed to chapter 7. Also, Jacquelyn Schulman contributed to the synthesis of molecules described in chapter 7. André St. Amant contributed to synthethic work towards the IAMpC monomer described in chapter 7.

Acknowledgments

This thesis would certainly not have been possible without Dr. Robert Hudson's patient guidance. I have often said that one could not wish for a better supervisor. He encouraged me to go to several conferences over the years and kept me going during the difficult times. He took a gamble by hiring a probationary student, an inorganic chemist with a biology problem, and I hope he has as few regrets as I do. I found a great home in nucleic acid chemistry - one where I can play with biologists and organic chemists alike. You have been an amazing mentor and a true friend. Thanks.

Dr. Nathan Jones played a pivotal role in my development as an undergraduate chemist without him I may not have pursued chemistry at the graduate level. I certainly owe him a debt for honing my writing skills, despite how frustrating his incessant corrections to manuscripts were. Also a good friend, I think we can all say that he is missed at Western and I wish him all the best in the future.

I would also like to mention early educators for turning me on to science in general. Dr. Kay Stevenson and Mr. Tim Lawrence of Felsted School deserve much credit. Incredible educators, they were tough and unforgiving but they were also affable and available at all hours for discussions. Three hour morning lectures would fly by. I have not seen them in many years but I hope they are both faring well.

Of course, I would like to thank my committee members for taking the time ro read this thesis and their helpful discussions and guidance along the way.

I would like to thank lab members past and present for their insight and their friendship. Particularly, I would like to thank Dr. Filip Wojciechowski, he was very much more than a co-worker to me but also taught me a lot of techniques that got me through graduate school. Fil always unselfishly helped me in my research and He and Kathy were there for me during some rough times, for this I thank them both. André, oh! André, you have always amused me. Thanks for everything. I hope California is as sunny as they say it is for you. Kirby, Mark, Mohamed, Mojmir, Melissa and McKenry all deserve my thanks for their friendship. You have all been great labmates and some of you as good friends as I will ever have. There

is nothing like the force of mutual experience of terrible results over and over again to draw people close. Keep in touch.

My mother's continuing love and reluctance to give up on me in the early years is what saw me here today and I owe to my father so much more than the skill of critical evaluation he gave me that lends itself so well to science. I thank and love them both for giving me opportunities throughout my life and making sure I turned out to be a 'balanced' person. They always put the kids first. Of course, Caz and Mark deserve thanks for being them and always being available to rant at, love you guys. Most of the people I mentioned here won't read this but I'm sure Aimée will. You pulled me out of a hole and into another one and I love you for it. Thanks for the support… everyone.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

A	adenine
Ac	acetyl
Ac_2O	acetic anhydride
AcCl	acetyl chloride
aeg	aminoethylglycine
ag	antigene
APS	ammonium persulfate
aq.	aqueous
as	antisense
ASO	antisense oligomer
BDF	base-discriminating fluorophore
Bn	benzyl
Boc	*tert*-butyloxycarbonyl
bp	base pairs
br	broad (spectral)
Bz	benzoyl
˚C	degrees Celcius
C	cytosine
Calcd.	calculated
cDNA	complementary deoxyribonucleic acid
d	doublet
δ	chemical shift in parts per million downfield from tetramethylsilane
DCC	dicyclohexylcorbodiimide
DCU	dicyclohexylurea
DIPEA	diisopropylethylamine
DMF	dimethylformamide
DMSO	dimethylsulfoxide
DMT	4,4'-dimethoxytrityl
DNA	deoxyribonucleic acid
EDTA	ethylenediaminetetraacetic acid
EtOAC	ethylacetate
ESI	electrospray ionization
Et	ethyl
equiv.	equivalents
EtOH	ethanol
FBS	fetal bovine serum
Fmoc	9-fluorenylmethoxycarbonyl
Fmoc-OSu	9-(fluorenylmethoxycarbonyloxy) succinimide
FCC	flash column chromatography
g	grams
G	guanine
hν	light
h	hours

HBTU	*O*-benzotriazolyl-1-yl-*N,N,N,N*-tetramethyluronium hexafluorophosphate
HOBt	1-hydroxybenzotriazole
HPLC	high-performance liquid chromatography
HRMS	high-resolution mass spectrometry
Hz	Hertz
J	coupling constant in Hz
λ	wavelength
K	lysine
L	litre
Lys	lysine
μ	micro
m	multiplet
m	meta
ma	major
MeCN	acetonitrile
MeOH	methanol
MEM	modified eagle medium
mi	minor
min	minutes
NEt_3	triethylamine
NMP	*N*-methylpyrrolidone
NMR	nuclear magnetic resonance
o	ortho
OD	optical density
p	para
PAGE	polyacrylamide gel elctrophoresis
PCR	polymerase chain reaction
PNA	peptide nucleic acid
ppm	parts per million
q	quartet
R_F	retention factor
RP	reversed phase
RNA	ribonucleic acid
RNAi	RNA interference
rt	room temperature
s	singlet
T	thymine
t	triplet
TBAF	tetrabutylammonium fluoride
TBDMS	tert-butyldimethylsilyl
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TLC	thin-layer chromatography
T_M	nucleic acid duplex melting temperature
TMED	tetramethylethylenediamine
TMS	trimethylsilyl
UV	ultraviolet

UTR	untranslated region
vis	visible
Φ_F	fluorescence quantum yield

Preface

During my undergraduate degree, the sequence of the human genome, consisting of approximately three *billion* base pairs, was published simultaneously by two groups. This occasion fell close to the 50th anniversary of Watson and Crick's discovery of the double helix. Now, as I come to the end of my Ph.D., we celebrate the 50th anniversary of the beginning of the determination of the genetic code. In approximately a decade the world went from not knowing the structure of DNA to knowing the gist of the language it spoke. The astounding pace of DNA related research continues unabated. DNA is the most fundamental of biological macromolecules and consequently, the most fascinating. Because of this, DNA attracts biologists, physicists, and chemists alike. However, chemists are perhaps the luckiest of these three categories of people: we can grasp the biology with some determined reading and (although I personally don't care to do so) many of us can approach the physics. Moreover, it is only chemists that can see the true beauty and evolutionary genius of the molecule in its organic chemistry. It is also only us that can rationally modify the molecule to make it do weirder and more wonderful things. We can change the base, the phosphate, or the sugar; tagging almost whatever we want to it, mixing it up with different monomers, activating it with light, forming scaffolds and, make it glow. I count myself lucky to have been introduced to nucleic acids, and having met them, I forsee spending the rest of my life in their intimate company.

Chapter 1

1 Introduction to selected aspects of nucleic acid chemistry

Over half a century since the elucidation of DNA's molecular architecture and a decade since the completion of the human genome project, nucleic acids still hold many questions. The deliberate application of synthetic organic chemistries towards nucleosides, DNA and their analogues has a rich history. The development of high-yielding oligomerization chemistries opened many doors for the study of nucleobase and backbone modifications, of which there is a seemingly endless array. These modifications can skew the basic properties of DNA to tune hybridization efficiency, enzymatic stability/biological recognition and photophysical properties. In this chapter selected historical and recent literature on the chemical aspects of nucleic acids and peptides will be discussed in order to put the reader 'in frame' for the present work. Excerpts of this chapter have been taken from Dodd, D.W.; Hudson, R.H.E. *Mini-Rev. Org. Chem.*, **2009**, *6*, 378-391.

1.1 Intrinsically fluorescent base discriminating nucleoside analogues

Modified nucleosides possessing intrinsically fluorescent heterocycles as base surrogates that are capable of canonical base-pairing are becoming important biomolecular tools. These "base-discriminating fluorophores" (BDFs) find use in hybridization-based mismatch detection, elucidation of DNA damage and the study of localized structural phenomena as exemplified in the study of ribozymes and nucleic acid/ligand interactions.

Classically, the central dogma of biology purports that the characteristics which make us unique are peculiar to the heritable material within almost every cell of our body: our DNA. It is therefore of no surprise that many disease states, both acquired and heritable, are derived from permutations within the genome. Mutations can make one more susceptible to heart disease, various carcinoma and are also the cause of diseases such as phenylketonurea, autism and cystic fibrosis, to name only a few. Diagnosis of mutations

that make one, or one's progeny, susceptible to disease is a very desirable advancement of modern medicine and many research groups are currently working on making this process more rapid and economical. A few existing techniques for sequence analysis are outlined in the following paragraphs.

One of the most frequently used methods for the detection of permutations within a polynucleotide sequence employs molecular beacons based on fluorescent resonant energy transfer (FRET).[1] The quintessential morphology of a molecular beacon consists of an oligonucleotide labeled at opposing termini with a fluorophore and a quenching moiety. Complementary sequences are placed at each terminus such that a hairpin structure is formed and the loop region is chosen to be complementary to the target oligonucleotide sequence to be analyzed. Upon hybridization of the interrogated sequence to the loop region, the stem opens and the fluorophore is no longer quenched as this quenching is distance dependent (Fig. 1-1). This basic technique can be applied in a variety of different ways, the fluorophore and quencher may be appended to different sequences and FRET can be used to obtain emission at differing wavelengths.[2]

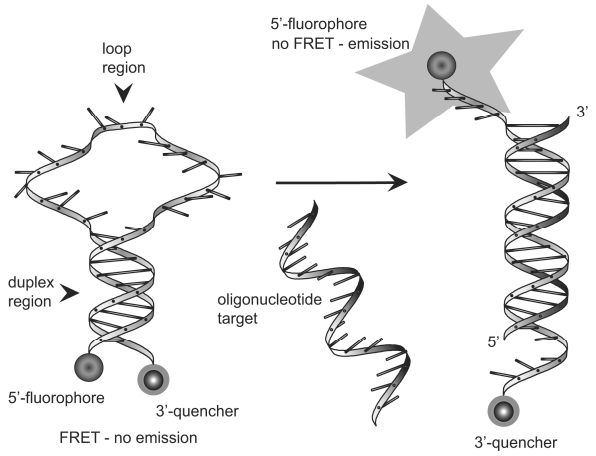


Figure 1-1: Illustration of the underlying mechanism of molecular beacon operation based on FRET.

Detection of single nucleotide polymorphisms (SNPs) can oftentimes be difficult due to the sheer size of the genome. A point mutation may occur only in low abundance creating the problem of interference from the wild-type gene. In order to increase sensitivity,

sequences can be examined and amplified through PCR-clamping using oligonucleotide analogues that will suppress the association of a primer to the template by steric hindrance. PCR clamping is most often peptide nucleic acid (PNA) mediated. Due to the unnatural *N*-(2-aminoethyl)glycine derived backbone, PNA is not recognized enzymatically and therefore does not act as a primer for DNA polymerase.[3] When a sequence of PNA is complementary to a PCR primer site it will virtually eliminate the formation of PCR product. PNA can be used in this manner to selectively silence or amplify genes differing by only one base. This technique, in conjunction with hybridization probes, has been used by several groups to great effect. The detection of point mutations using a wild-type-specific PNA and a mutant-specific PCR primer has been carried out for oncogenes[4,5] and even on mitochondrial DNA.[6]

Full sequencing of a PCR amplified gene is the archetypal method of SNP analysis; this technique is currently undesirable as it is costly and time consuming despite the significant advances that continue to be made in this area.

Electrochemical means may also be used for SNP detection. Many different approaches have been used, although most involve oligonucleotide modified electrodes that result in differing charge transfer rates upon complementary sequence binding. As an example of this, a recent paper has illustrated that a mismatch can be discriminated through labeling of an oligonucleotide immobilized on a gold surface with anthraquinone at the 2'-position, as the anthraquinone moiety intercalates the charge transfer rate decreases allowing for mismatch detection.[7] Electrochemical mismatch detection offers the advantages of simple readout and amenability to digitization. Microarray analysis of SNPs by electrochemical means remains limited as a mechanical connection is required to each electrode pad. This would make the manufacture of larger arrays, required for full sequence-screening, a difficult challenge.[8]

The aforementioned techniques, molecular beacons in particular, are limited as they rely in differences in hybridization efficiency. These differences vary greatly with sequence context and with the ionic strength of the medium and strict hybridization conditions must be met.[9] In the context of oligonucleotide arrays, variations in packing density can

affect hybridization efficiency therefore making the manufacture of high-throughput screening devices technically challenging.[10]

Fluorescent base-pairing nucleoside analogues which may fluorometrically respond to hybridization when placed within an oligonucleotide sequence are of high interest as they overcome the reliance on hybridization efficiency and offer a conveniently measured parameter in the wavelength or intensity of light emission (Fig. 1-2). These base-discriminating fluorophores (BDFs) may be used for a wide range of sensing applications including the detection of SNPs which may be used for the diagnosis of heritable diseases in a less costly and more expedient manner than full sequencing.

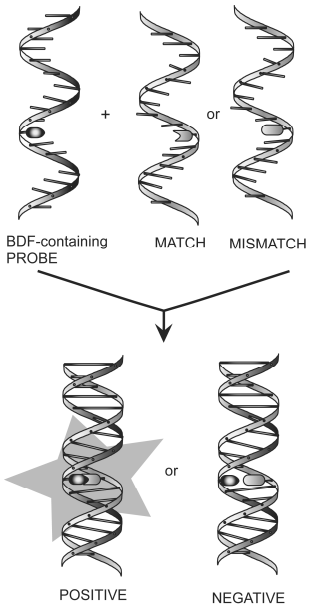


Figure 1-2: Illustration of the principle of sequence readout by a base discriminating fluorophore (BDF) containing probe. The interaction of a BDF-containing probe with a complementary sequence leads to an observable property.

The use of BDFs allows for the elucidation of localized events as opposed to other methodologies which monitor global structural phenomena (UV-Vis, CD spectroscopy). Other methods for detecting localized structural events do exist, such as NMR spectroscopy and gel electrophoresis of kinetically trapped radiolabeled products;

however, these techniques are time consuming and limited with respect to reaction time scale.

As the natural nucleobases are essentially non-emissive,[11] much work has been put towards the synthesis of base analogues which are luminescent. In the design of novel base-discriminating fluorophores several basic criteria should be met. The fluorophores should maintain a high degree of structural similarity with the natural nucleobases so as not to affect the hybridization efficiency, the compounds must obviously be fluorescent, their excitation wavelengths should not be within the absorption range of biological macromolecules, and the fluorescence should be sensitive to microenvironmental conditions. The fluorescence response may be a change in emission wavelength, an increase or decrease in quantum efficiency or length of the excited state lifetime; however, dramatic changes in emission wavelength or intensity are the most conveniently measured parameters.

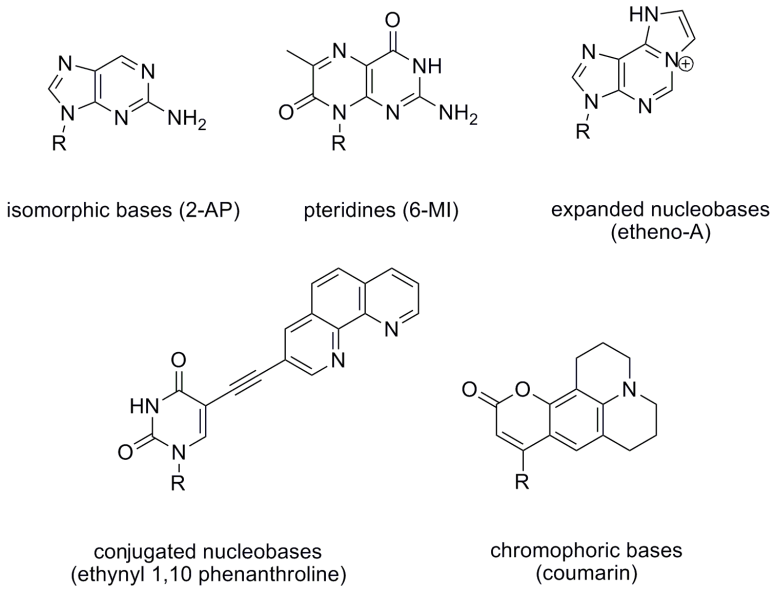


Figure 1-3: Representative molecules of the different classes of BDF as defined by Tor.[12]

Tor has divided fluorescent base analogues into five basic categories: isomorphic base analogues, pteridines as purine analogues, extended nucleobases, conjugated base analogues, and aromatic hydrocarbon and chromophoric base analogues (Fig. 1-3).[12]

Fluorescent base analogues may be more selectively divided into two categories which are capable of canonical base pairing: those possessing pendant fluorophores and intrinsically fluorescent nucleoside analogues. Below are pictured representative molecules of these classes (Fig. 1-4).[13,14] Both of these molecules are able to hydrogen bond with natural bases and respond fluorometrically to hybridization. The former class often has an advantage in that higher overall brightness (brightness is defined by the quantum yield multiplied by the extinction coefficient at the excitation maximum) and greater quantum yields are often achieved through the attachment of a traditional chromophore. However, fluorescence spectroscopy is a fantastically sensitive technique and oftentimes highly luminescent molecules are not required to be a practical and useful reporter group.

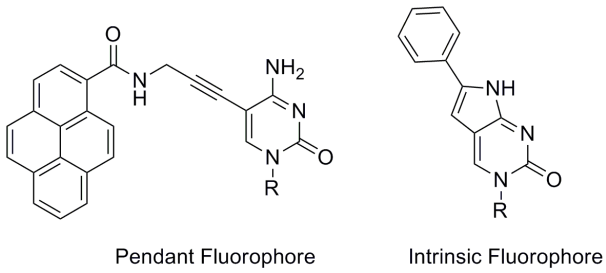


Figure 1-4: Representative molecules of the two basic fluorescent nucleobase analogue designs: those containing a pendant fluorophore, and the intrinsically fluorescent 6-phenylpyrrolocytosine.[13,14]

The use of conventional fluorescent moieties, attached to the base *via* a flexible linker, can allow independent movement of the fluorophore, this can make interpretation of results complicated as the fluorescence response is not necessarily originating from the area of interest. The intrinsically fluorescent BDF category has become attractive as it is the nucleobase itself whose environment is being monitored, they are synthetically accessible and also have potential for biological incorporation. Such intrinsically fluorescent, complementary base-discriminating/pairing nucleobase

analogues are the subject of this sub-chapter. Whenever possible the photophysical parameters characterizing the BDFs ablility to act as a reporter group are given; however, for some analogues this information is not yet available.

Intrinsically Fluorescent Purine Nucleobase Analogues: Complementarity and Response to Pyrimidines

Isomorphic bases: 2-aminopurine and 8-azaguanine

The base analogues 2-aminopurine (2-AP) and 8-azaguanine (8-azaG) are historically two of the most often used isomorphic, fluorescent nucleobases. Their overall size matches that of the natural bases and they are able to form isostructural Watson-Crick base pairs (Fig. 1-5). 2-AP, a constitutional isomer of adenine, has been used for almost 40 years.[15] Owing to its high quantum yield as the free nucleoside in aqueous solution (Φ = 0.68) and extreme sensitivity to microenvironmental changes with drastically lower emission observed in single-stranded (ss) and double stranded (ds) oligonucleotides (as much as 100 fold decrease in emission in dsDNA),[16] 2-AP has been employed in many studies of nucleic acid structure and dynamics and in a wide array of biochemical assays. Examples include: real-time monitoring of hammerhead ribozyme folding, oligonucleotide cleavage and inhibition and nucleic acid/protein and nucleic acid/drug interactions.[17] 2-AP can form Watson-Crick type base pairs with thymine, uracil or cytosine, this can either be a benefit or a drawback, depending on the application. It also has a red shifted absorption spectrum which allows differential excitation in the presence of biological macromolecules.

8-azaG is another isomorphic purine analogue which can be enzymatically incorporated into oligonucleotides. It is relatively emissive, exhibits a high degree of solvatofluorochromism and is quenched by adjacent nucleobases (Fig. 1-5). 8-azaguan(os)ine has a high quantum yield when *N*1 is deprotonated at high pH (Φ = 0.55) but much lower fluorescence when the natural Watson-Crick base-pairing face is presented. One of the more interesting and recent advances made with this BDF has been the development of a highly efficient enzymatic synthesis of the triphosphate from 8-azaguanine and the establishment of pH dependent fluorescence within a structured

oligoribonucleotide (with implications towards ribozyme mechanism elucidation).[18]

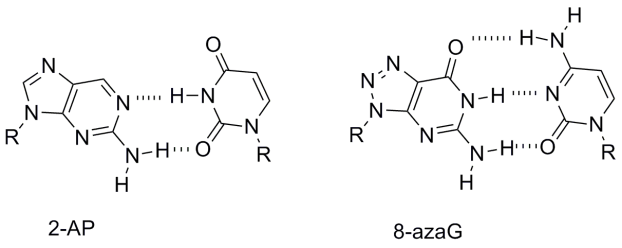


Figure 1-5: The isosteric base analogues 2-aminopurine and 8-azaguanine in hybridization with their natural base complements uracil and cytosine, respectively.

Purines modified at the 8-position

The 8-position of purines lends itself well to derivatization; 8-bromoadenosine is remarkably easy to prepare in high yield and can be used in a variety of transition metal-catalyzed cross-couplings. Deoxyadenosine and guanine derivatives with phenol directly attached at the 8-position can be accessed through the coupling of hydroxyphenylboronic acid with the appropriate 8-bromonucleoside in the palladium catalyzed Suzuki-Miyaura reaction.[19] These compounds are of interest for several reasons: the compounds are biomarkers for phenol exposure and are therefore of use for the study of mechanisms of carcinogenicity, the compounds exhibit pH dependent fluorescence [20] and the phenolate, anionic form of the modified base can act as a quencher through photoinduced electron transfer (PET).[21] Although these nucleosides have not yet been incorporated into oligonucleotides, their high quantum yields at neutral pH (Φ = 0.25–0.56) could make them quite useful as BDFs.

Various other moieties have been attached to purines in the same manner including

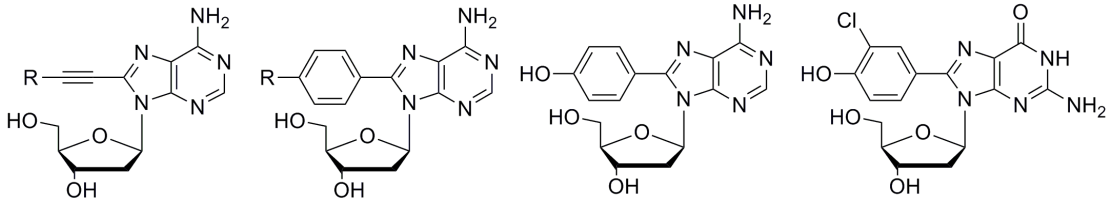


Figure 1-6: Purine nucleosides derivatized at C-8 (R = H, bpy).

bipyridine and phenanthroline for use as metal chelators (demonstrated for Ru^{II}) and various aryl substituents to tune fluorescence properties (Fig. **6**).[22]

Triazolyl adenosines have been made and characterized through a collaborative effort and have been found to be highly fluorescent (Fig. 1-7 left). The first report by Dyrager et al. showed that quantum yields as high as 0.62 are achieved.[23] A follow up study also showed that the base (when incorporated into oligomers) was sensitive to its microenvironment, caused only mild structural perturbations to B-form DNA and formed base pairs with thymine and also, more surprisingly, adenine.[24] The high fluorescence observed was surprising to us as we had synthesized geometric isomers of the compounds reported (Fig. 1-7, right) and found quantum yields to be very low (Φ_F ~ 0.02) (unpublished results).

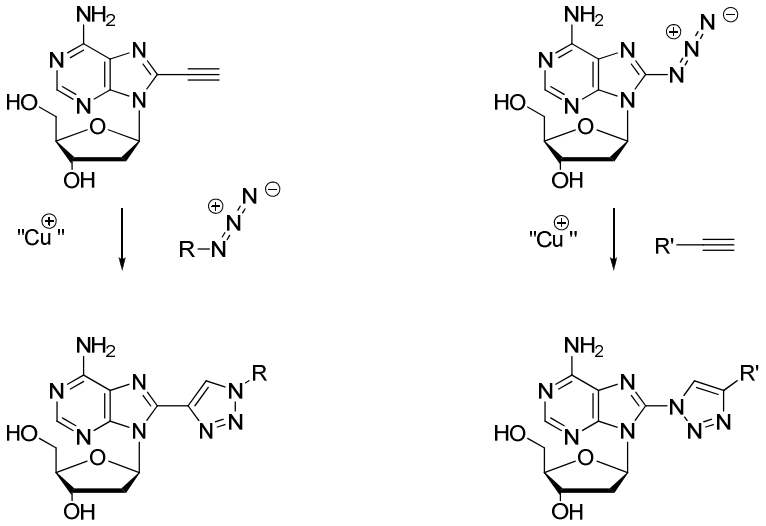


Figure 1-7: Geometric isomers of 8-triazolyldeoxyadenosines for use as BDFs, R=Alkyl.

Both syntheses started from 8-bromoadenosine, the former work then used a Sonogoashira coupling to introduce the alkyne while we took the opposite approach of introducing an azide *via* a simple S_NAr with sodium azide. We found that 8-azidoadenosine was very unreactive towards cycloaddition with alkynes and were only able to produce two compounds in low yield (R = Ph, R = EtOMe).

Purines modified at the 6-position: 2-Amino-6-(2-thienyl)purine and 2-amino-6-(2-thiazolyl)purine

The compounds shown in figure 1-8 have been known for some time and applications for the highly emissive base analogues continue to be found.[25-27] The unnatural base pairs of 2-amino-6-(2-thienyl)purine with 2-oxopyridine and 2-amino-6-(2-thiazolyl)purine with imidazolin-2-one remarkably function in transcription with good fidelity and hence may be placed specifically within an RNA transcript with ease. Both the thienyl and thiazolylpurines have quantum yields from 0.41-0.46 as the free 5'-monophosphates. These quantum yields are roughly halved upon incorporation into a ss oligonucleotide and halved again when in ds DNA/RNA. Used alone or in FRET experiments, with fluorescein derivatives (FAM) as an acceptor moiety, accurate descriptions of hybridization events have been obtained. When placed adjacent to one another, the modified bases exhibit self-quenching behaviour but may still allow for the excitation of a nearby FAM moiety. The authors suggest that this may be desirable as it would enable the use of excess probe to test sequence in biological assays.[27]

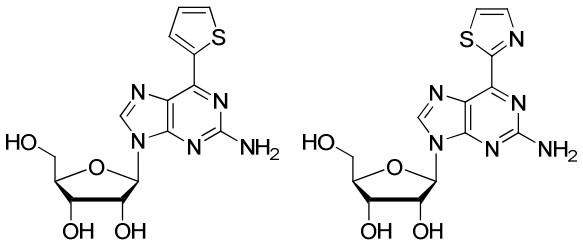


Figure 1-8: 2-amino-6-(2-thienyl)purine and 2-amino-6-(2-thiazolyl)purine used as guanosine mimics despite not presenting an isostructural Watson-Crick face.

Pteridine nucleoside analogues

Pteridines are naturally occurring compounds, originally isolated from the wings of pierid butterflies by the pioneering biochemist F.G. Hopkins in 1895.[28] Hopkins would later receive a Nobel prize for his discovery of vitamins (not that this is of any particular

relevance to us here). Pteridine derivatives have had many pharmaceutical applications, particularly as anti-cancer drugs. Pteridines have more recently been utilized as nucleoside analogues; they are highly fluorescent and are synthetically accessible through the condensation of the appropriate triaminopyrimidine with the desired pyruvate derivative (Fig. 1-9).[29]

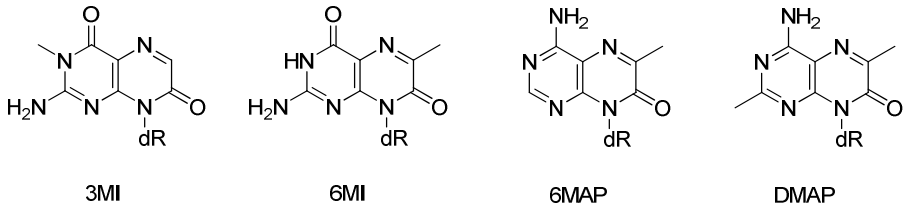


Figure 1-9: Pteridine purine analogues.

Once the base analogue has been synthesized, *N*-glycosylation can be carried out to yield the desired nucleoside, this in turn can be converted to the corresponding phosphoramidite by well established methods.

The pteridines described by the Hawkins group are structurally similar to the natural purine bases, containing the appropriate hydrogen bond donor and acceptor components, with one exception (*vide infra*). The compounds 3-methylisoxanthopterin (3MI) and 6-methylisoxanthopterin (6MI) have been employed as guanosine analogues while 4-amino-6-methyl-8-(2'-deoxyribofuranosyl)-7(8*H*)-pteridone (6MAP) and 4-amino-2,6-dimethyl-(2'-deoxyribofuranosyl)-7(8*H*)-pteridone (with the unforntunate acronym DMAP) have been used as adenosine mimics.[30] While other derivatives have been synthesized, it is the aforementioned compounds that have found the greatest success in the elucidation of DNA structural events.

The applications of the pteridines as BDFs are diverse as the compounds are highly fluorescent (Φ = 0.77-0.88 for the adenosine analogues and 0.39-0.48 for the guanosine) and highly sensitive to microenvironmental conditions. All compounds are well tolerated within the duplex excepting 3MI due to the methyl group at the 3-position interfering with hydrogen bonding. This methylation destabilizes the duplex to a similar extent as would a single canonical mismatch.

Pteridines have been used in the following applications: HIV-1 integrase activity assays, alkyl transferase coupled assays, use as hydridization probes, HU protein binding detection, 'A-tract' structure detection, RNA polymerase activity and intracellular oligonucleotide transport analysis.[30]

Intrinscially Fluorescent Pyrimidine Nucleobase Analogues: Complementarity and Response to Purines

5-Ethynyluridine

Variously substituted 5-ethynyluridine derivatives have been exploited in our laboratory for mismatch detection. These structurally simple and compact fluorophores are able to report hybridization events by "turning on" in the presence of a match sequence when placed internally within the modified strand. The greatest fluorescence response, out of a small selection of compounds, was observed for a *p*-methoxyphenylethynyluracil containing probe (Fig. 1-10) giving a six-fold increase in fluorescence upon encountering a match sequence.[31]

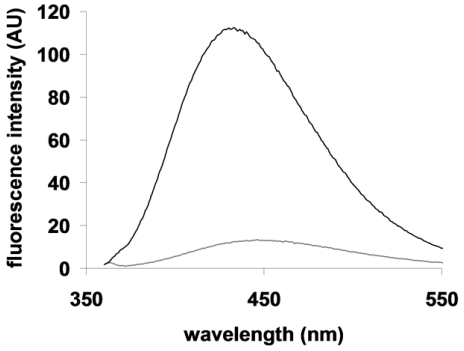


Figure 1-10: 5-(p-methoxyphenylethynyl)uracil containing oligonucleotide in the presence of a match sequence (black) as compared to ss (grey).[31]

The mismatch duplexes offer surprising results: when the mismatch is guanine the fluorophore is quenched considerably; when it is thymine, the chromophore suffers an intermediate degree of quenching and; when it is cytosine, mismatch duplexes are not quenched significantly. These data imply that the fluorophores are exquisitely responsive to their local environment and may therefore find further use in determining DNA

structural characteristics and ligand/DNA interactions.

Related alkynyluracils have been exploited for mismatch detection by Brown and co-workers. Again using Sonogashira cross-couplings, anthracenyl-ethynyl (Ae) and naphthalenyl-ethynyl (Ne) moieties have been appended to deoxyuridine at the 5-position

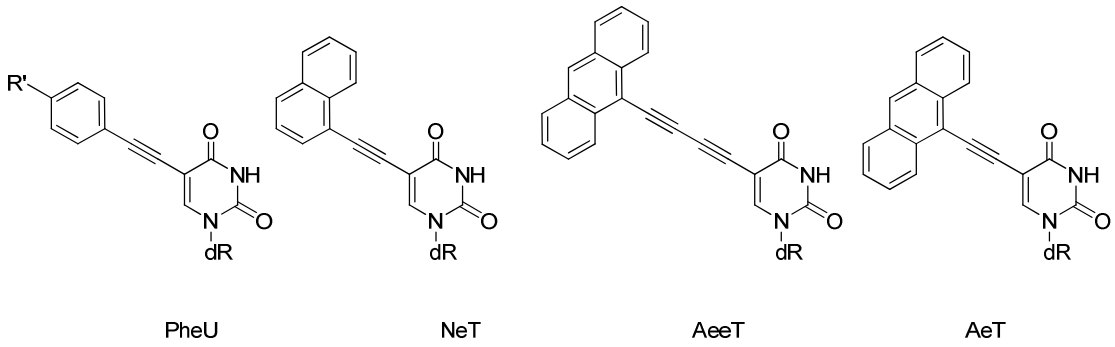


Figure 1-12: Phenyl, naphthalenyl and anthracenyl-alkyne derivatized deoxyuridines used for mismatch detection (R = H, OMe).[31,33]

(Fig. 1-11).[32,33] The corresponding anthracenyl -diyne (Aee) was also synthesized and these compounds were found to be only marginally destabilizing towards duplex formation. As is typically the case with fluorophores appended to pyrimidines at the 5-position, an increase in fluorescence was observed on duplex formation due to a change in the polarity of the microenvironment when the chromophore is projected into the aquated major groove of dsDNA *versus* its position in ss form.[32,33]

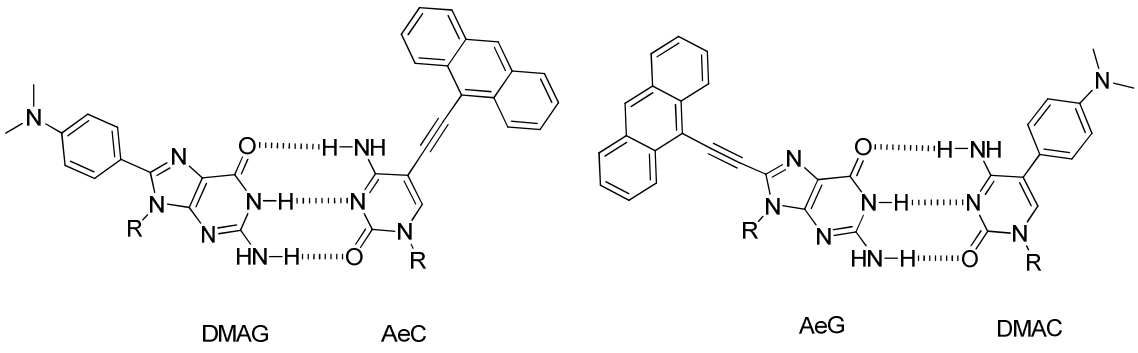


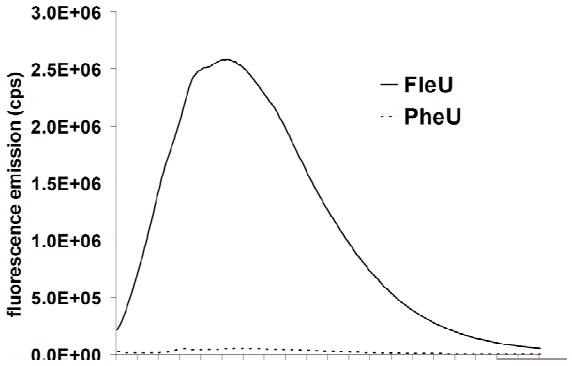
Figure 1-11: Anthracenyl-ethynyl (Ae) derivatized riboguanosine (G) and ribocytidine (C) and dimethylaniline-(DMA) modified bases.[34]

Alkynylnaphthalene analogue of deoxthymidine (AeT) containing oligomers showed the greatest increase in emission upon binding to a match sequence; however, when presented with a C mismatch the fluorescence was also much increased (quantum yields not reported). The results were very similar with the anthracene containing oligomer only with overall lower fluorescence. The diyne compound gave surprising results, the fluorescence increased greatest in the case of a C or T mismatch, the T mismatch increase in emission intensity was also accompanied by a bathochromic shift of 12 nm. It can be said that all three of the compounds tested could potentially be used to effect in A/G SNP typing.[33] A similar modification, but for cytidine, was reported earlier by Sessler and coworkers.[34] They prepared a pair of base-modified ribonucleosides to examine photoinduced electron transfer in a hydrogen bonded ensemble. They chose to prepare a *C*8-modified guanosine and a *C*5-modified cytidine (Fig. 1-12).

Although the purpose of the study was not base-discrimination, the fluorescence spectra of AeC and AeG were reported. The fluorescence was mostly characteristic of anthracene, yet each modified base had slightly different photophysical properties. Given the conjugated nature of the linkage of the luminophore to the nucleobase, it is reasonable to suspect that they may be responsive to base pairing in the context of an oligonucleotide. Likwise, the dimethylaniline-modified bases, especially DMAG, is structurally related to modified purines already presented (Fig. 1-6). A detailed characterization of DMAG and DMAC were not reported in this work.

A fluorene derivatized deoxyuridine[35] nucleoside and the analogous PNA monomer (Fig.1-13) have both been synthesized and were found to be moderately fluorescent (Φ = 0.14 for the free nucleoside).[36] Oligonucleotides containing the nucleoside analogue were amenable to SNP typing, selectively fluorescing in the presence of a match sequence in a quencher-free molecular beacon construct.[35]

Although the fluorescence of the PNA monomer was found to be *ca.* 50 times greater than the structurally similar 5-phenylethynyluracil derivative,[36] (Fig. 1-14) the fluorescence response on duplex formation was very modest. This highlights the principle of making only minor structural changes to the nucleobases to maintain sensitivity to the microenvironment. It also highlights the differences in the structures on ssPNA versus ssDNA and the challenges associated with portability of modifications/technologies to oligonucleotide analogues.



fluorophore undergoing charge redistribution dependent on the polarity of the environment. DAN has been appended to guanosine and cytidine *via* the exocyclic amino group in order to monitor the polarity of the minor and major grooves of DNA (in both B and Z-form) respectively.[38,39] This same fluorophore has recently been appended to 2'-deoxyuridine by Saito's group in different ways as shown in figure 1-15.[40]

The fluorescence of these nucleoside analogues was measured with respect to solvent polarity. The fully conjugated system (Fig. 1-15) was accessed from a Suzuki-Miyaura coupling of the DAN borate ester with 5'-*O*-DMT-5-iodouridine and showed the greatest

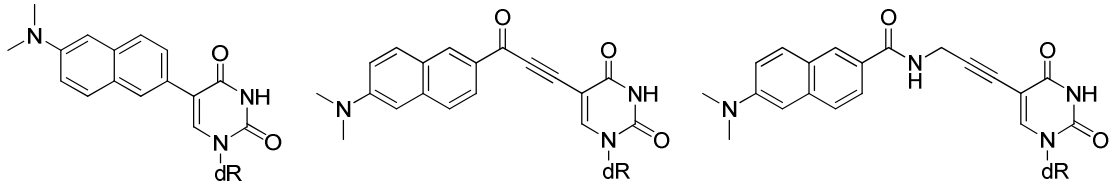


Figure 1-15: DAN appended nucleosides differing by linkage strategy: directly attached, propanoyl, and carboxamide linkages.[37]

degree of solvatofluorochromicity. Surprisingly, the pendant chromophore (electronically separated from the base) and cross-conjugated compounds (Fig. 1-15) showed very different fluorescence properties with the being the least fluorescent of the three and the propynoyl linked compound being the most, the highest quantum yield being 0.26. The high degree of solvatofluorochromicity displayed by these compounds implies that there is strong charge transfer character to the relevant excited state.[40]

C5-Furan substituted uridine

Purines, typically modified at *C*8, and pyrimidines modified at *C*5 are prominent in the literature due to a minimal thermodynamic penalty upon hybridization and ease of synthesis. A particularly rewarding modification at *C*5 of uridine was the appending of a furan ring *via* the 2-position of furan. This was done by the Stille coupling of 2-(tributylstannyl)furan with 5-iodouracil and was first reported in 1991.[12,41] The furanyluracil was found to have a relatively low quantum yield (Φ=0.03 in H_2O) but demonstrated interesting bathochromic and hyperchromic shifts upon increase in solvent

polarity (ether→water). Despite the low luminescence, the sensitivity to microenvironment implies that the compound may be useful in mismatch detection; indeed, the compound has been used for the successful detection of abasic sites in DNA.[32] In the same paper, nucleosides modified with oxazole, thiazole and thiophene showed less favourable photophysical properties. In addition to the favourable photophysical properties displayed by 5-furanyluridine it was also later found that the 5'-triphosphate was able to be incorporated by T7 RNA polymerase in the place of thymidine with good fidelity.[42] The fluorescence of the monomer has been shown to be dependent on microscopic polarity; a well definined linear plot of fluorescence intensity versus experimentally determined $E_T(30)$ values [43] has been obtained. The use of microscopic polarity scales, such as $E_T(30)$ values, has often been a more accurate parameter for such research than the use of dielectric constants.[44] The furan modified nucleobase has been incorporated into the potential pharmaceutical target: the bacterial decoding site known as the A-site and was used to screen aminoglycoside antibiotics. Although the fluorescent response on binding of the aminoglycoside was less than that exhibited by 2-aminopurine, when used in conjunction with 2-aminopurine an effective assay was developed. The need for new RNA structural probes still exists as potent inhibitors of this site do not induce a proportional change in fluorescence when using the currently available fluorophores. This fluorophore has also been used to examine the polarity of the major grooves of A and B-form nucleic acids due to its strong solvatochromic behaviour.[45] The *C*5-furan modified deoxycytidine has also been prepared, but it is approximately 3-fold less emissive than furanyl-U (Fig. 1-16). Despite this lower emission the base has been employed in the successful discrimination between 8-Oxoguanosine, G and T. 8-oxoguanosine is a biomarker for oxidative stress and has been found to cause transversion mutations during DNA replication; therefore, its detection/quantification within the genome is desirable.[46]

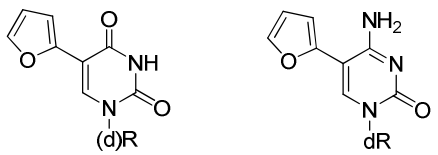


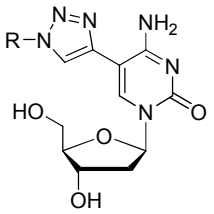
Figure 1-16: Furan-decorated nucleosides.

Due to the success found with furan-modified uridine, Tor and co-workers prepared the *C*8-furan modified adenosine and guanosine. Even though the nucleosides were highly emissive with quantum yields of 0.57 and 0.69, respectively, they possessed maximal emission in the near UV (375 nm) which displayed little solvent polarity dependence. Likely because of these suboptimal properties, investigation of their use as BDFs in oligonucleotides has not yet appeared.[12]

Triazole appended nucleosides

Although there are many examples utilizing the Huisgen 3+2 azide/alkyne cycloaddition[47] for the modification of pre- and post-oligomerized nucleic acids, little attention has been paid towards the fluorescent properties of the resultant molecules. The most popular substrate for this copper(I) catalyzed "click" reaction has been 5-alkynyluracil as it can be accessed readily from the Sonogashira cross-coupling of 5-iodouridine with trimethylsilylacetylene, followed by desilylation. Various azides have been used and it has recently been found that stacking between adjacent triazole modified bases can be stabilizing towards the duplex.[48] A full set of fluorescent nucleobase analogues have also been developed through the "clicking" of azidocoumarin with 7-deaza-7-alkynyldeoxyguanosine, 7-deaza-7-alkynyldeoxyadenosine, 5-alkynyl-deoxyuridine and 5-alkynyldeoxycytidine to form 1,4-disubstituted triazoles.[49] These compounds are not "inherently fluorescent" as categorized by the present review as they possess an independently fluorescent moiety attached to the nucleoside. Carell and co-workers have also used this reaction as a means of modifying DNA post-synthetically; the applications of this strategy abound.[50]

Benzyltriazolyluridines have also been synthesized for use as small molecule hydrogelators [51,52] and our group has synthesized the corresponding cytidine analogues that were found to be moderately fluorescent (Fig. 1-17).[53] A variety of azides have been utilized including benzyl, phenyl, ethylphenyl, thiophenyl (connected *via* the 3-position of thiophene) and bithiophenyl azide. These monomers have been studied as potential base-discriminating fluorophores and are described in greater detail in chapter 5.



R = Ph, Bn, EtPh, thiophenyl,
bithiophenyl, fluorenyl.

Figure 1-17: 1,4-disubstituted triazolylcytidines synthesized for use as potential BDFs (further described in chapter 5).

Thiophene appended nucleosides

5-(Thiophen-2-yl)uridine was first synthesized in 1994 with the aim of increasing duplex stability through stacking interactions.[54] However, only a modest increase in melting temperature resulted. The fluorescent properties of this and related molecules have been since exploited for mismatch detection. Bi- and terthiophene moieties appended to deoxyuridine with or without an ethynyl spacer (Fig. 1-18) all bestow fluorescent properties on the nucleosides with varying suitability towards practical application.[55]

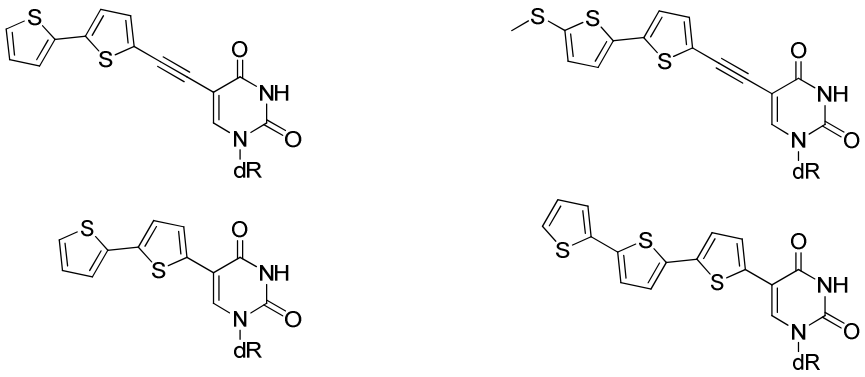


Figure 1-18: : Bi- and terthiophene derivatized uridines synthesized for use as BDFs.

It was found that there was only a small thermodynamic penalty associated with duplex formation; however, high ss emission and complex fluorometric data make the nucleoside analogues of limited use for mismatch detection.

Fused ring systems: Fluorescent analogues of purines and pyrimidines

Benzo- and Naphtho-pyridopyrimidines

Some base discriminating nucleobases of particular interest have been described by Saito and co-workers. Benzopyridopyrimidine (BPP) is quenched effectively by G but fluoresces in the presence of A; however, low overall quantum yield and high ss emission prompted the researchers to further elaborate on this scaffold thus giving rise to naphthopyridopyrimidine (NPP) (Fig. 1-19).[9] The latter modified base was found to have a markedly higher quantum yield and less intense ss emission and has therefore found use in A/G SNP typing. Two other fluorescent bases of importance devised by the same group are methoxybenzodeazaadenosine (^{MD}A) and methoxybenzodeazainosine (^{MD}I) (Fig. 1-19).[9] Probes containing these bases are quenched in the presence of T and C complementary oligonucleotides respectively.

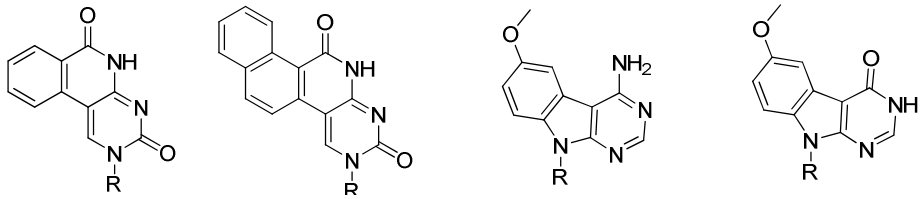


Figure 1-19: Expanded-ring-system fluorescent nucleobases developed by Saito and coworkers.

Size expanded nucleobases

A discussion of fluorescent, modified nucleobases would not be complete without mention of the pioneering work by Leonard and subsequent elaboration by Kool on size expanded nucleobases.[66,57] Leonard's *lin*-benzoadenosine was one of the first examples of size expanded nucleobases which retain the natural hydrogen-bonding characteristics of the parent base (Fig. 1-20). Through the insertion of benzene into the purine ring, desirable photopysical properties may be achieved. Kool then applied this ideology to the

remaining three natural bases and acquired a full set of benzo-expanded nucleosides and incorporated them into an oligomer which was coined "xDNA" (Fig.1-20).[59]

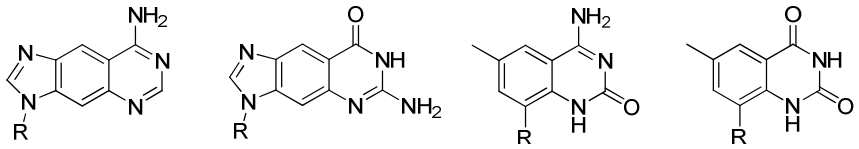


Figure 1-20: The full set of "xDNA" benzo-expanded fluorescent bases which maintain canonical Watson-Crick hydrogen bonding.[56-58]

This work has been developed further by the creation of "yDNA", which differs by the point of insertion of the benzene into the natural base, and naphtho-expanded nucleobases, both of which are highly fluorescent.[59] Although these fluorophores are not well tolerated with natural nucleotides in the same strand, full sequences of the modified bases show interesting fluorescent properties. New emission maxima appear within certain sequence contexts due to excited state interactions and these may be exploited in the future for SNP analysis or other sensing applications.

Thieno[3,2d]pyrimidine

A particularly elegant, divergent synthesis of isomorphic fluorescent base analogues was recently described by Tor which relies on modification of the thieno[3,2*d*]pyrimidine

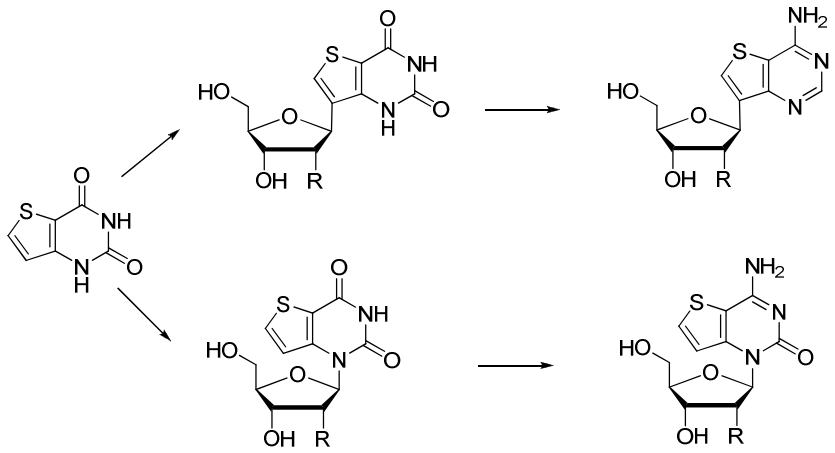


Figure 1-21: Alternative glycosylation of the thieno[3,2d] pyrimidine core with a protected ribofuranoside yields both purine (top) and pyrimidine (bottom) analogues.

core. Glycosylation at *N*-1 provides a pyrimidine analogue, while *C* glycosylation at the β position of thiophene yields the purine analogue (Fig. 1-21).[60]

The functional groups can then be manipulated to yield isosteric nucleobase analogues that maintain hydrogen bonding with the natural bases.

Of the 4 isosteric base analogues synthesized, the thymine mimic showed the greatest fluorometric response to changes in solvent polarity. Upon titration with the four natural 5'-monophosphates this analogue showed quenching in the presence of all bases excepting the complementary nucleotide AMP.[60] Such quenching may also occur in ss form making SNP analysis troublesome; indeed, it is yet to be shown if this fluorescence response can be replicated in the context of oligomers.

More recently, a constitutional isomer of the aforementioned compound, based on the thieno[3,4-*d*]pyrimidine core, has been synthesized. Due to the apparent hypsochromic effects observed in lower polarity solvents, the triphosphate was incorporated enzymatically into an RNA transcript. This oligomer was shown to report a C mismatch by an increase in emission.[61] Introduction of this monomer into a specific RNA hairpin the detection of the heterodimeric protein ricin was possible. Ricin catalyses the de-purination of ribosomal RNA at specific sites and therefore causes a halt in translation and cell death. Upon cleavage of the unnatural nucleobase an increase in fluorescence is observed. This is a useful contribution to the field as a practical field test for a potential biological warfare agent can be envisaged in the near future.[62]

Benzoquinazolines

Benzoquinazoline derivatives have been used as probes for duplex and triplex formation and can be synthesized either by the condensation of an α-aminocyanonaphthalene or an α-aminonaphthanoic acid with urea (Fig. 1-22).[63,64]

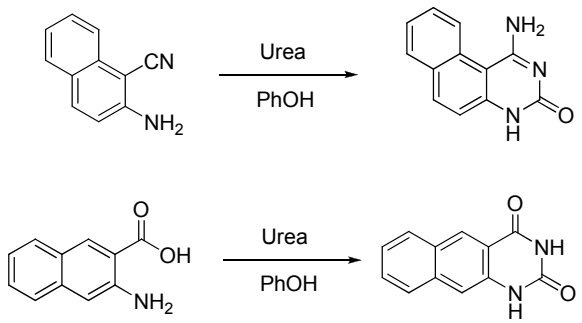


Figure 1-22: Synthesis of benzoquinazolines for use as fluorescent pyrimidine mimics.[63,63]

The reaction of α-aminocyanonaphthalene with urea results in a cytosine analogue which can be converted to the thymine mimic by treatment with conc. HCl in DMF.

Amongst other uses, the benzoquinazolines have been successfully employed as triple helix probes. Homopurine tracts are requisite for triple-helix formation whereby a third strand binds in the major groove of the duplex; in parallel triplexes the third strand, composed of pyrimidines, binds to the purine strand of the duplex in the major groove by Hoogsteen base-pairing.[65,66] Benzoquinazoline thymidine mimics (Fig. 1-22) were incorporated into an oligopyrimidine sequence, they were found to be stabilizing towards triple-helix formation (due to more favourable stacking interactions) and, more relevantly, they were also found to respond fluorometrically to hybridization, particularly when incorporated into the Hoogsteen strand. Fluorescence maxima were blue-shifted and decreased in intensity while the excited state lifetime increased upon triple-helix formation. As the formation of parallel triplexes are pH sensitive (relying in part on CG*C^+ triplets) the fluorescence observed was also subject to pH.[64]

Etheno(ε)-bridged bases

Although structurally similar to the natural bases, etheno(ε)-bridged bases possess strikingly different photophysical properties (Fig. 1-23).

These base analogues can be formed *in vivo* through the action of various carcinogens such as vinyl chloride and ethyl carbamate on the natural exocyclic amine-containing

bases.[67]

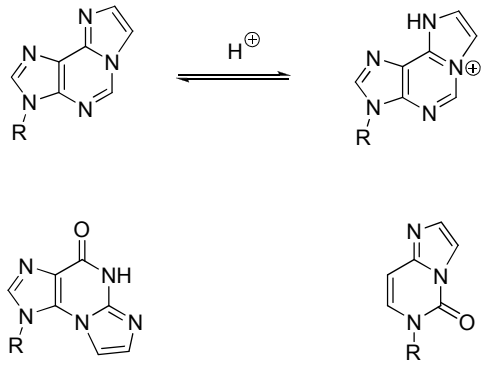


Figure 1-23: Etheno(ε)-bridged bases: ε-adenosine in equilibrium with the cationic form responsible for its desirable photophysical properties, ε-guanosineand ε-cytidine.

These analogues are highly fluorescent (quantum yields as high as 0.6) and also very pH sensitive, in the deprotonated state the compounds are only weakly fluorescent, this pH sensitivity has been exploited in related derivatives to gauge the acidity of nucleotide binding sites,[68] and with excitation wavelengths of 290-300 nm, their usage in the presence of absorbing biological macromolecules is permitted.

Described by Leonard in the early 1970s, ε-adenosine has been used in numerous applications. Its first biologically significant uses were as the bridged ε-adenine dinucleotide (εNAD+) and ε-adenosine triphosphate (εATP).[69,70] The synthetic NAD+ analogue was found to be an active coenzyme in 4 different dehydrogenase-catalyzed reactions and was also an effective substrate for NADase and phosphodiesterase I. Although the emission intensity of εNAD+ is only 8% of the parent monophosphate, full emission is restored upon phosphodiester cleavage allowing the real-time monitoring of such hydrolyses. εATP was found to be a versatile analogue of ATP, it is an allosteric effector of several enzymes and a co-substrate of kinases.[70]

It has been found recently that ε-7-deaza-2'-deoxyadenosine has a higher quantum yield than the corresponding ε-adenosine and is more stable to extreme pH.[71] Despite the solution of a crystal structure containing a putative ε-A – G base pair [72] evidence for

similar behaviour in solution is scant as both ε-A and its corresponding deaza-compound are destabilizing towards duplex formation when placed within an oligonucleotide sequence. This lack of base-pairing competency obviously limits their potential use as BDFs.

Tricyclic analogues of cytosine: 3,5-Diaza-4-oxophenothiazine (tC) and 3,5-diaza-4-phenoxazine (tC^O)

These base analogues have been known for over a decade and were the first example of tricyclic pyrimidine base analogues capable of complementary base-pairing and compounds of similar structures continue to be pursued (Fig. 1-24). These tricyclic cytidine analogues base-pair with guanine effectively and also impart greater stability towards the duplex, purportedly due to increased stacking interactions; indeed, when two phenoxazine or phenothiazine tricycles are neighbouring one another the duplex stability is further increased.[73] The compounds are highly fluorescent, tC (X=S) as the free nucleobase analogue (which can be prepared according to Roth and Schloemer [74]) is only very slightly more emissive than when incorporated into an oligonucleotide. A small increase in anistropy is found upon hybridization to a complementary strand, brought about by the increase in effective molecular mass. The long emission wavelength of tC means it may be used in FRET experiments coupled with chromophores such as rhodamine as the acceptor moiety.[75] Unfortunately, there is little change in the emission spectrum upon hybridization to a complementary strand and only a small increase in the fluorescence lifetime.[76] The phenoxazine (X=O) (tC^O) has very similar properties to the phenothiazine although it is slightly more emissive and has had the honour of "brightest DNA-incorporated base analogue" bestowed on it by the authors of a recent puclication, as judged by its quantum yield in the duplex state.[77] This striking luminescence was put to use in probing the conformational dynamics of DNA polymerase. Using an Alexa-555 labeled cysteine residue as the FRET acceptor, movement of a subdomain was measured using stopped-flow experiments on nucleotide binding.[78]

The phenoxazine derivative, first prepared by Matteucci and co-workers, has been used more recently by Sigurdsson as a structural scaffold for nitroxide spin labeling which is

useful for EPR studies (Fig.1-24).[79] The nitroxide, quite expectedly, quenched the

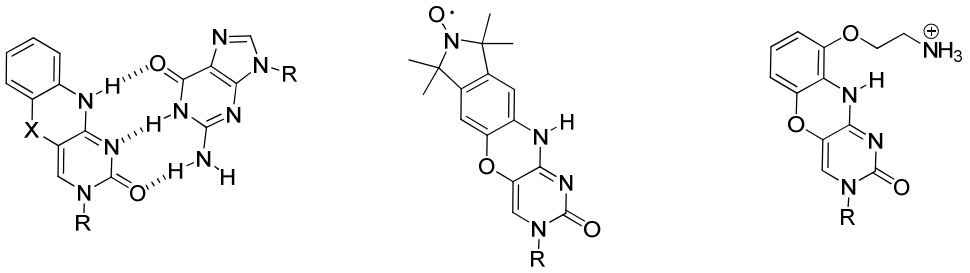


Figure 1-24: 3,5-diaza-4-oxophenothiazine (tC) (X = S) and 3,5-diaza-4-phenoxazine (tCO) (X = O) in hybridization with guanine. Elaboration of the tC(O) scaffold has yielded a nitroxide spin labeled compound that may be used for EPR measurements and the "G-clamp" which has increased binding affinity to guanine.[73-83]

fluorophore; however, fluorescence could be re-introduced by reduction of the nitroxide to the sulphonate ester or the hydroxylamine (the latter of which is fully reversible) while incorporated within an oligonucleotide. This switchable behaviour has obvious benefits, fluorescence and EPR measurements may be made within the same vessel. The use of this compound for SNP typing has been investigated in some simple sequences, it is possible to distinguish between single stranded, match sequence and various mismatches based on emission intensity. Matches with G result in a *ca.* 60% decrease in intensity from the single stranded form with mismatches being flanked by the two extremes of emission. When placed opposite A several maxima appear in the emission spectrum for reasons, as of yet, unclear.[80]

Another elaboration of the phenoxazine has been the introduction of an ethoxyamino group to the terminal aromatic ring. The primary ammonium is capable of forming an additional hydrogen bond to guanine on the Hoogsteen face, further increasing duplex stability while maintaining desirable photophysical properties [81] this additional hydrogen bonding inspired the moniker "G-clamp". The G-clamp has been introduced into both DNA and PNA and shows a high degree of sequence dependence with regards to binding affinity.[82,83] The increased binding affinity of the G-clamp has been exploited to sterically inhibit HIV-1 Tat-dependent *trans*-activation *in vivo* when incorporated into a 2'-*O*-Me oligoribonucleotide.[83] Replacing the primary amine of the "G-clamp" with a guanidine

functionality further stabilizes the duplex through an additional hydrogen bond with the *N*7 of guanine. These modified bases have been incorporated into PNA, 2'-*O*-Me oligoribonucleotides, 2'-*O*-methoxyethyl phosphorothioate oligoribonucleotides and DNA primarily for stability purposes and the fluorescence response data have not been hitherto reported.[84-86]

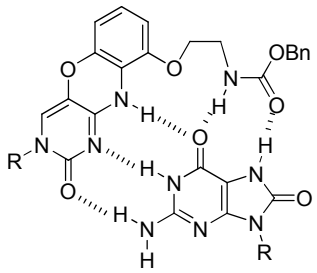


Figure 1-25: Purported hydrogen-bonding motif between Cbz-protected "G-clamp" and 8-oxoguanosine (R=2'-deoxyribose).

G-clamp derivatives have been used for the fluorometric detection of 8-oxoguanosine.[87] Protection of the pendant amino functionality of the "G-clamp" with a Cbz group allowed for selective quenching in the presence of 8-oxoguanosine. This is purportedly due a hydrogen bonding interaction of the carbonyl oxygen of the Cbz group with the *N*7 hydrogen of 8-oxoguanosine (Fig. 1-25).

Quenching was found to occur only on titration of 8-oxoguanosine, the natural nucleosides, A, C and T, did not affect fluorescence appreciably.

CONCLUSIONS

Although much progress has been made, the repertoire of useful fluorescent nucleotides must continue to be expanded as no universal system exists. SNP analysis remains one of the primary goals in this field; however, new applications continue to be found. The ability to monitor gene regulation, transcription and protein/nucleic acid interactions in general is becoming increasingly accessible. Combinations of a BDF with an oligonucleotide analogue which provides high affinity and discriminating binding, such as LNA, PNA or morpholinos has yet to be fully explored.

Certain examples were highlighted in order to give a broad overview of the field. For a

more complete overview of the literature the reader is directed to several excellent reviews.[9,42,88-93]

1.2 'Chemical' synthesis of peptides and nucleic acids

DNA: An Historical Perspective 1871-1953.

As the story goes, on the 28th of February, 1953 James Watson and Francis Crick walked into the Eagle pub in Cambridge proudly declaring "We have found the secret of life." Worn-out controversies surrounding X-ray diffraction data aside, Watson and Crick were standing on tall shoulders indeed when they proposed the structure of the double helix.[94] The story of the elucidation of DNA's molecular architecture begins much earlier and what follows is a very brief tour of the early science which led Watson and Crick to the pub that day for a celebratory pint.

He did not know the gravity of his findings at the time but the Swiss scientist Friedrich Miescher discovered DNA in 1871, a substance he termed nuclein.[95] It was isolated from puss-soaked bandages he obtained from a nearby hospital and was not totally free of protein contaminants. Many at the time discredited his discovery thinking that it was simply phosphorus contaminated protein, but Meisher continued to work and found that nuclein could be isolated from other sources and had a high molecular weight (inferred from the viscosity of his preparations in solution). He was also able to conduct elemental analysis and found that it had a unique P/N ratio. In 1880 Kossel, embarked on the first real structural determination of DNA and found that a hydrolysate of nuclein contained four bases; adenine, cytosine, guanine and thymine.[96] Later, Altmann obtained protein-free nuclein and also found it to contain phosphorus and that it was acidic in character. He then introduced the term nucleic acid. It was now little doubted that nucleic acids contained bases and phosphorus but another component was missing. Phoebus Levene discovered that nucleic acids from isolated from some sources contained D-ribose while from other sources the hitherto unreported sugar 2-deoxyribose predominated.[97]

At this stage the components of DNA were known and now it was a matter of fitting them together. At the time, there was no reason to believe that DNA was a

macromolecule. People still held that proteins were the hereditary material accounting for Mendel's pea's behaviours. Nothing that has only four distinct units could ever be the basis for the astounding complexity of life on earth (*sic*). It was also thought that the bases were present in equal proportions. Due to these preconceptions and compounded by the lack of sufficient analytical techniques, a tetranucleotide hypothesis for the structure was put forward by both Levene and Takahashi. Levene proposed a linear tetranucleotide and Takahashi, a circular. Having correctly determined the ring structure of thymidine, Levene and Tipson were remarkably close with their second proposal[98] knowing that the phosphate must link 5' and 3' positions. However, over the next two decades, two major developments would reveal the inadequacies of Levene's tetrameric idea.

Remarkable work, first by Griffith,[99] then Avery, McCleod and McCarty,[100] culminating in the Hershey-Chase[101] experiments showed that DNA was the hereditary material. Hershey and Chase showed through radiolabeling that bacteria transformed by a T2 phage did not retain the tranfected protein and DNA was the molecule responsible.[101] Edwin Chargaff found that the proportions of nucleobases varied between samples but the molar ratios of A to T and G to C remained constant, regardless of the source of the DNA.[102,103]

The lack of analytical techniques in the past is really what spurred on natural product-driven chemistry. There was a time, and it is even so today (albeit to a lesser extent), where to prove the identity of a molecule it must be pepared at the bench and compared to the natural substance. Although we do not often think of it in these terms, DNA *is* a natural product and it was chemical synthesis of oligodeoxynucleotides that originally determined and confirmed its primary structure. Emil Fischer had used this method to unambiguously determine the structure of the purines around the turn of the century. He recognized that guanine, adenine, xanthine, uric acid, and caffeine were all derivatives of a parent bicyclic compound he named purine. Having synthesized purine in 1898 he then found the differences between the naturally occurring compounds were differential hydroxyl and amino substitutions. Lord Alexander Todd, just in time for Watson and Crick, unambiguously determined the β-anomeric configuration in nucleotides[104,105] and then set about to make a dinucleoside phosphate with the natural 3'→5' linkage.[106] In

doing so, he would lay the basis for further oligomerization chemistries.

Before the synthesis of the dinucleoside phosphate, as was suggested, Watson and Crick made their leap of thought to correctly guess the structure of DNA. Crystal data supplied by Maurice Wilkins and (perhaps unwillingly supplied by) Rosalind Franklin told them the structure was helical in nature and highly hydrated. In talking with chemists in neighbouring laboratories they were told the proper tautomeric forms of the bases. With Chargaff's rules in hand, the leap doesn't seem so far today. Their model correctly contradicted that of Linus Pauling.[107] Pauling had proposed a three stranded structure with many inherent problems. They proposed that there was hydrogen bonding between hydrogen phosphates that would be deprotonated at physiological pH and they placed the hydrophobic bases on the outside of the helix. There was no plausible mechanism for self directed replication. Watson and Crick's model accounted for self replication, a helical structure that seemed chemically sound. We will now focus back to chemistry directed towards the synthesis of oligomeric DNA.

DNA Synthesis.

Todd's contributions were significant as they created both the H-phosphonate and phosphotriester approaches (Fig. 1-26). The first step shown below is an H-phosphonate coupling giving the H-phosphonate diester that may be activated using an electrophilic chlorine source (in this case *N*-chlorosuccinimide), which then undergoes a phosphotriester coupling to give the desired linkage (Fig 1-26).[106] Deprotection thusly yielded the first synthetic dinucleoside phosphate. There were inherent drawbacks to this method. Most seriously there was no clear way to extend this method to trimers and beyond. Although it would later prove necessary to protect the phosphate, the protecting group choice of benzyl was not ideal as it was somewhat base-labile. Todd's approach was abandoned for a period as a new method was introduced by Khorana.

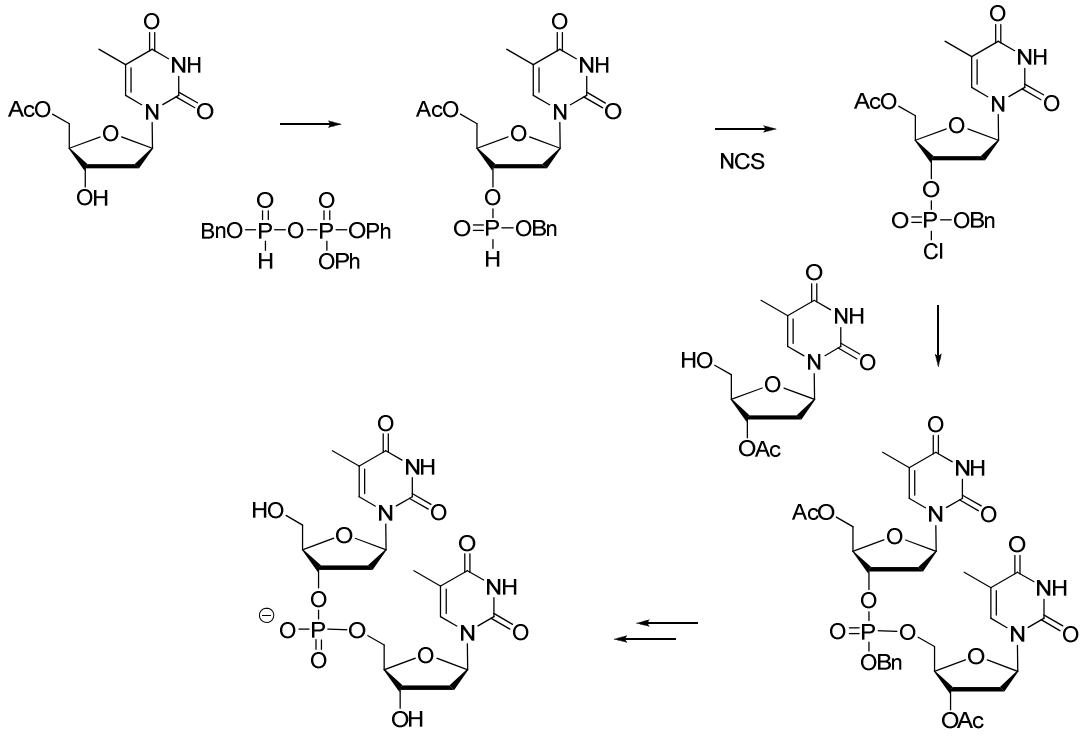


Figure 1-26: Lord Todd's synthesis of a dinucleoside phosphate using the phosphotriester approach.

Khorana's synthesis came later to be known as the phosphodiester method,[108-110] and is important for several reasons. The synthetic scheme is outlined in figure 1-27. It is perhaps the most simple of approaches; a condensation of a phosphate with an alcohol. At first, dicyclohexylcarbodiimide was used as a coupling reagent but was later replaced by arenesulfonyl chlorides (Fig.1-27).[111]

Perhaps the most important part of this synthesis is the choice of 5'-protection. 4,4'-dimethoxytrityl (DMT) is able to be introduced selectively to the 5'-OH due to its steric bulk and is labile under very mildly acidic conditions. The glycosidic bond of purine 2'-deoxynucleosides is susceptible to acid hydrolysis; however, not under conditions required for DMT removal. This choice is orthogonal to base protecting groups commonly used to date (*vide infra*) and allows for stepwise elongation of the oligomer. Downfalls of this approach include the long coupling times required (up to 2 days), and the accumulation of charge on the oligomer. The anionic backbone is able to act as a

nucleophile in subsequent coupling steps and therefore reduces yield considerably. Purification was also laborious due to the presence of many side-products.

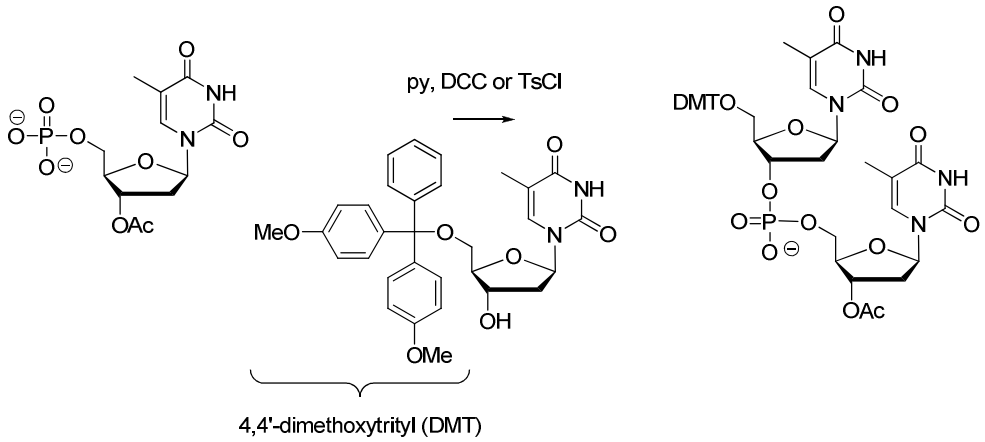


Figure 1-28: Khorana's phosphodiester approach sees the introduction of the important 4,4'-dimethoxytrityl protecting group.

In the late 1960s Letsinger began to revisit the phosphotriester approach of nucleotide synthesis as outlined in figure 1-28. We have already seen chemistry similar to this; however, here we are introducing a new protecting group. As mentioned earlier, protection of the phosphate was imperative; it cut down on side reactions and allowed easier purification. The choice of the base labile β-cyanoethyl group as a permanent protecting group was an important one. It is cleaved along with base protecting groups and is completely orthogonal to the DMT group which Letsinger had brought from Khorana's work.[112,113]

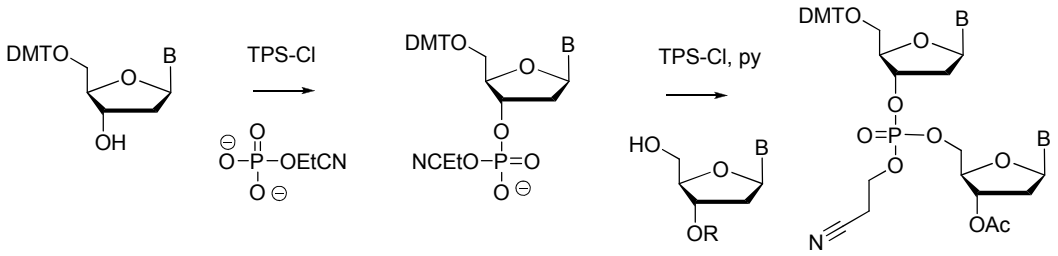


Figure 1-27: Letsinger's phosphotriester approach utilizing the base-labile 2-cyanoethyl protecting group.

Again, these coupling reactions were slow and prompted Letsinger to harness the inherently higher reactivity of P (III) phosphorylating agents thus giving rise to the

phosphite triester approach shown in figure 1-29.[114]

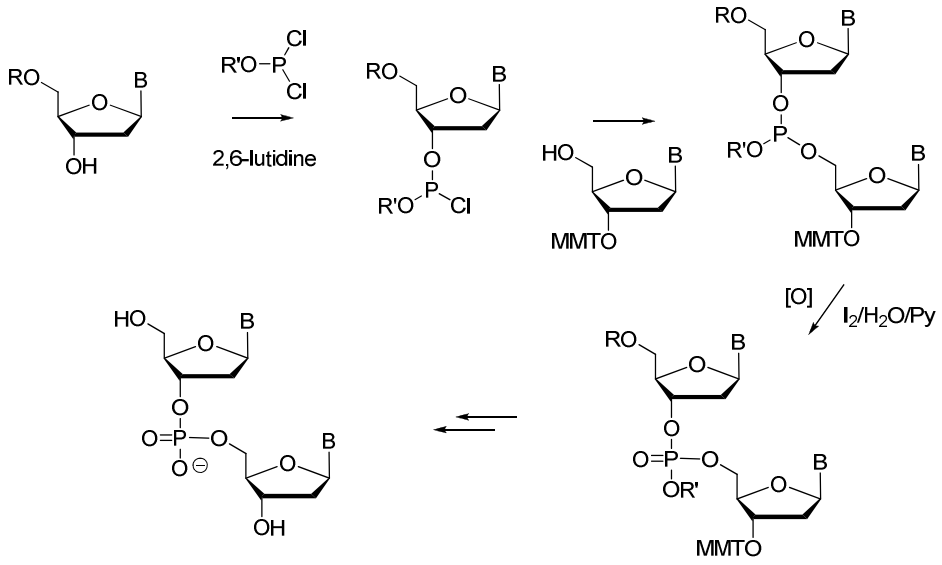


Figure 1-29: Letsinger's phosphite triester approach harnessing the high reactivity of phosphorus III. R = phenoxyacetyl, R' = Cl_3CCH_2

The first step of the reaction shown above (Fig. 1-29) is incredibly fast and is carried out at -78˚C in THF. The intermediate phosphorochloridite is not isolated, as it is extremely sensitive to water, and is reacted directly with the free 5'-hydroxyl of another nucleoside to give the phosphite triester which can be oxidized in situ with iodine and water. Although much faster, there were several disadvantages to this method; the high reactivity of the phosphodichloridite and its bifunctionality necessitated the use of low temperature and led to the possibility of 3'→3', 3'→5', and 5'→5' linkages occurring. Nevertheless, Letsinger was able to carry out the synthesis of a tetramer and laid in place the last of the important developments to allow the now gold standard phophoramidite coupling method to emerge.

In 1981 Beaucage and Caruthers developed the phosphoramidite technique which is what is almost exclusively used today.[115] Caruthers' technique along with the introduction a few minor modifications[116,117] is shown in figure 1-30 and borrows on many of the chemistries hitherto developed. The power of this technique lays in the caged reactivity of the monofunctional phosphoramidites used. Only in the presence of a weak acid are

they activated sufficiently and are stable to ambient conditions to a much larger degree than the previously used P(III) derivatives.

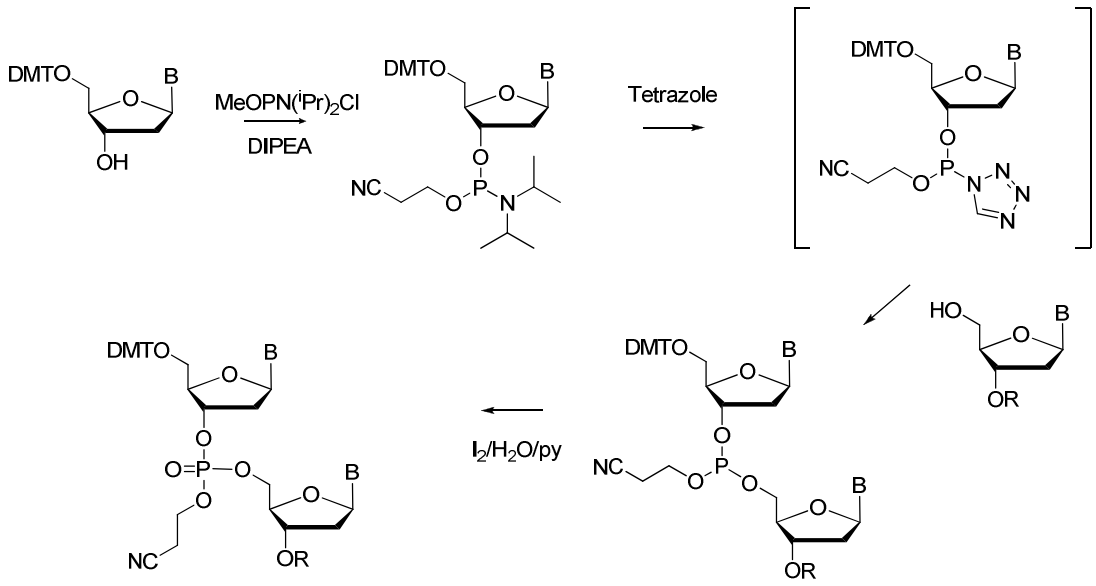


Figure 1-30: Caruthers' phosphoramidite approach. R = levulinyl.

Using tetrazole as a weak acid, the diisopropylamine moiety becomes protonated and displaced by tetrazolate which then acts as leaving group for the incoming 5'-OH. After this addition oxidation can occur in the same manner as previously described by Letsinger.

The phosphoramidite approach represents the highest yielding and most practical way of synthesizing oligonucleotides in the laboratory; however, it would be of limited use if it was not used in conjunction with the solid-supported chemistry that is outlined in the next segment of this sub-chapter.

Solid-Supported Oligomerization of Nucleic Acids and Peptides.

It is little doubted that few things have changed the face of biological polymer synthesis, and indeed oligonucleotide synthesis, to the extent that solid-phased chemical approaches have. There are three main benefits to solid-phase synthesis. 1) Separation of the resin bound product or intermediate products can be carried out simply by filtration and a few washing steps; this makes the process amenable to automation. 2) Minimized handling losses occur as the product remains attached to the resin. 3) Reagents can be used in

excess in order to drive the reaction to completion resulting in very high stepwise yields. The three preceeding points also make the technique particularly amenable to the synthesis of oligomers as repeated chemistries may be applied to lengthen a chain sequentially. As are all scientific discoveries today, solid-phased peptide synthesis (SPPS) was only possible due to the extant body of hard won research in related fields; however, this makes its advent no less significant. Proof-of-principle was illustrated in a pair of reports from R.B. Merrifield at The Rockefeller Institute in 1963 and 1964, respectively.[118,119] The second of these papers demonstrated a method that, with only minor modifications, is still in use today. Insoluble chloromethylated polystyrene was used and to it was attached the *C*-terminus of an (*N*-Boc protected) amino acid. Merrifield synthesized the nonapeptide bradykinin *via* dicyclohexylcarbodiimide (DCC) coupling in record yield. Synthesis, purification and characterization was complete in 8 days. Coupling efficiency was shown to be nearly quantitative and the final product obtained behaved identically, both biologically and chemically, to those prepared previously. Temporary protection of the terminal amino group was done using *tert*-butyloxycarbonyl (Boc) resulting in an acid labile carbamate that was readily cleaved with 1 M HCl in AcOH in iterative cycles while cleavage of the benzyl ester linkage to the resin could be carried out with HBr in trifluoroacetic acid (TFA). This was later optimized to give the current Boc chemistry protocol whereby the temporary protection is removed with TFA while the permanent protecting groups and the linkage to the resin are cleaved using HF. Because these protecting groups are both acid-labile on each cycle of Boc deprotection some side chain deprotection and cleavage unavoidably occurs causing branching and reduced yield. Another downfall of this approach is the use of HF which is an extremely hazardous reagent requiring special equipment for safe handling. Some peptides are also not stable under these conditions and side reactions reduce yield. Figure 1-31 shows the Boc protecting group strategy contrasted to that of 9-fluorenylmethoxycarbonyl (Fmoc).

Problems of solvent accessibility to the resin also needed to be overcome, what was required was a more polar resin that would allow penetration of solvent and reagents. Sheppard's laboratory developed more polar resins that allowed for increased yields in the synthesis of peptides and nucleic acids.[120,121] As the resins used became better and sequences became longer the issue of cleavage from the resin by repeated acid treatment became the limiting factor for Boc chemistry. What was needed was an orthogonal approach in peptide synthesis.

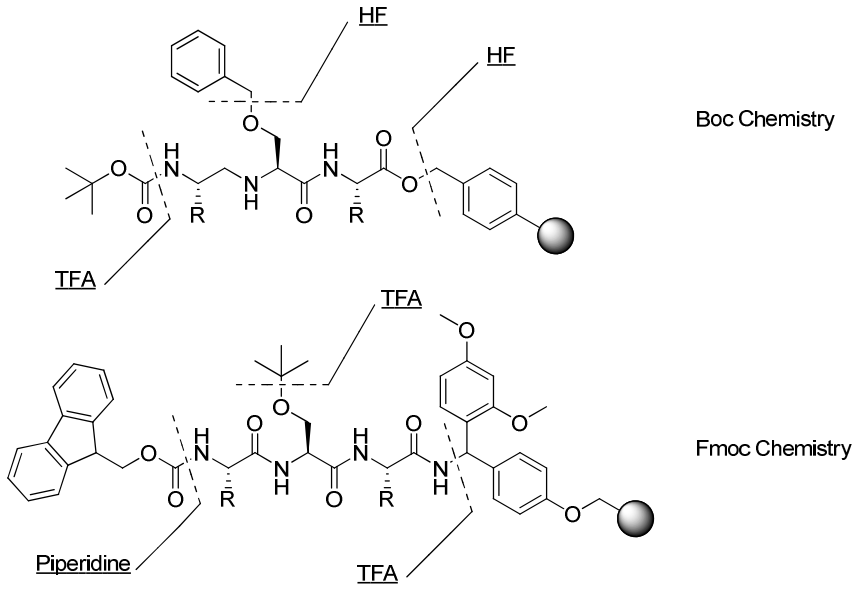


Figure 1-31: Contrast of Boc (graded) and Fmoc (orthogonal) protecting group strategies.

The introduction of the base labile fluorenylmethoxycarbonyl (Fmoc) nitrogen protecting group by Carpino[122] attracted little attention at first. Many existing amino protecting groups formed volatile or otherwise easily removed by-products and this was not the case with Fmoc. The dibenzofulvene formed can react further and polymerize making isolation of the desired product more difficult. A problem for solution phase chemistry but not so for solid-supported approaches as the by-products can be simply washed away (Fig. 1-32).

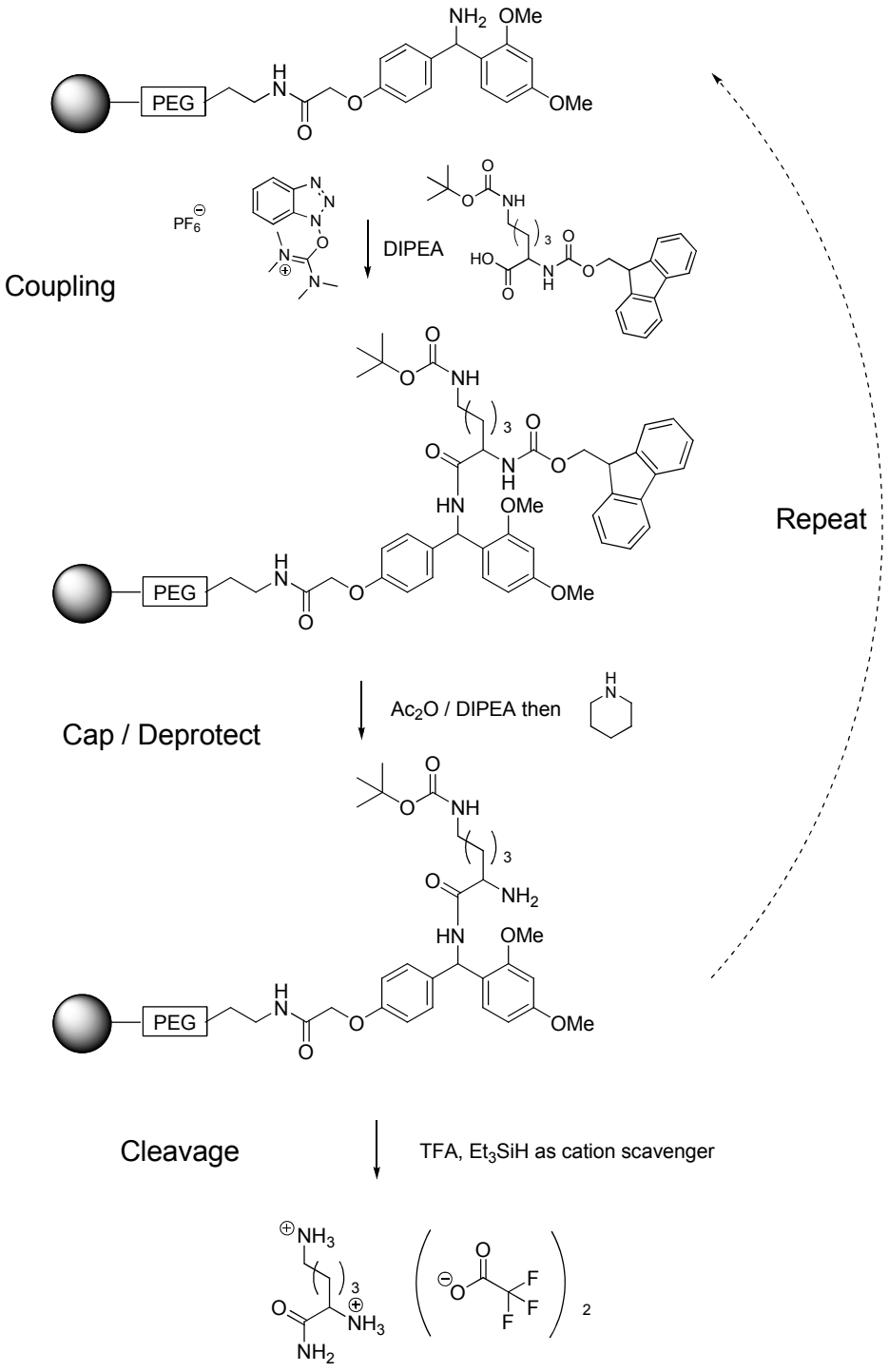


Figure 1-32: Modern Fmoc solid phase peptide synthesis.

A schematic representation of modern Fmoc-based peptide synthesis is shown above. In this case HBTU is used as a coupling reagent. The active ester is formed and coupled to the amine of the support (in this case NovaSyn® TGR) creating tetramethylurea and *N*-hydroxybenzotriazole as by-products that can readily be washed away. During this step a large excess of active ester (*ca.* 5 fold) is introduced in order to maximize stepwise coupling yields. A capping step allows for increased yields and easier purification by preventing the continued synthesis of truncated sequences. Acetylation renders any remaining amino groups unreactive. Removal of the Fmoc group is typically carried out with 20% piperidine in DMF. The use of a secondary amine for this purpose is important as it can further react with the liberated dibenzofulvene to form an adduct thus preventing its potential reaction with the growing oligomer or its polymerization. At this stage the progression of coupling can be monitored by the release of the dibenzofulvene either by conductivity readings or photospectroscopically. The process is then repeated until the desired length is reached and then cleaved from the resin using TFA. This process also removes permanent protecting groups; shown here is Boc. Purification of the trifluoroacetate salt is then carried out by HPLC or less often by electrophoresis and characterized by mass spectrometry. With the use of this resin a carboxamide C-terminated peptide is obtained; however, more traditional linkers result in the natural carboxylic acid.

The first report of solid-phase synthesis of nucleic acids[123] went largely unnoticed and it was not until 15 years or so later that its true potential was found. This was due to the problems with the chemistry at the time and the unsuitability of the resins in use. The most commonly used support today is the controlled pore glass[124] and is usually functionalized with a long-chain alkylamine. The alkylamine is conjugated to the base via a linker (often succinoyl) which can be cleaved by base. Original efforts concentrated on the phosphotriester approach but this and the H-phosphonate approach have now been overwhelmingly superceded by the phophoramidite approach. Shown in figure 1-33 is a modern solid-phase strategy using a universal support. Many approaches start with a pre-loaded support where the first sugar is attached directly to the succinoyl linker using a

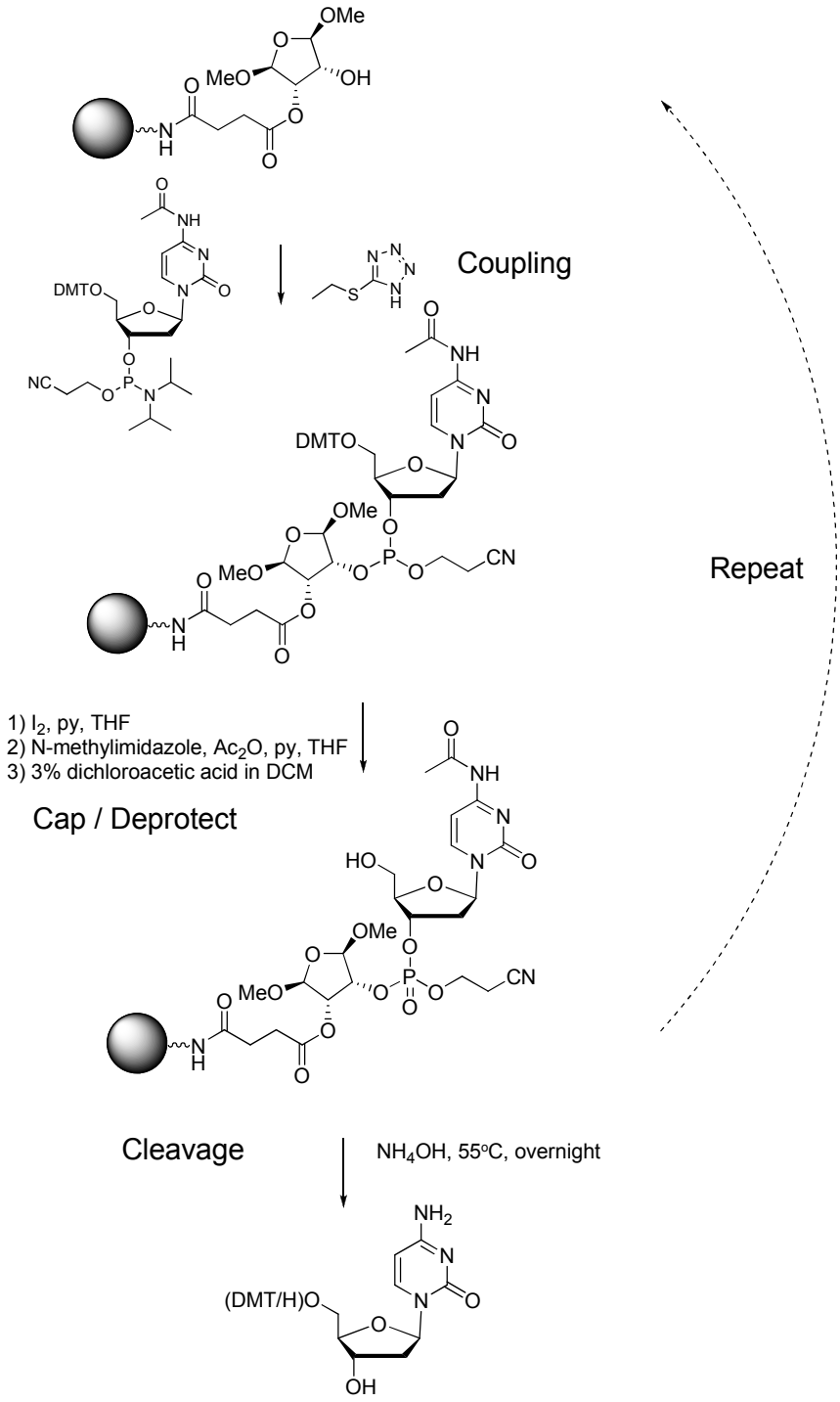


Figure 1-33: Modern solid supported oligonucleotide synthesis by the phosphoramidite method.

coupling reagent such as DCC yielding, upon cleavage, an oligonucleotide with a free 3'-OH. Use of a universal linker allows for the straight phosphoramidite coupling of the first nucleotide and upon cleavage yields an oligomer with a 3'-terminal OH (Fig 1-33). Tetrazole is still used by many as an activator; however, other activators are now also in common use such as dicyanoimidazole and ethylthiotetrazole. Ethylthiotetrazole offers advantages in that, due to a lowered pKa, it can be used in lower concentration. It also is more soluble in acetonitrile than tetrazole.[125] For these reasons Glen Research has termed it the "Rambo Activator". Research by the company found that it performs better in 0.25M concentration than 0.45 M tetrazole, particularly for RNA synthesis. After coupling, the phosphorus(III) must be oxidized to phosphorus(V) prior to capping and this is carried out with either I_2 or an organic peroxide. Capping then proceeds for the same reasons as demonstrated in peptide synthesis, followed by mild acidolysis of the DMT group. The cycle may then be repeated until the desired length oligomer is reached then cleaved from the resin with concomitant removal of base- and phosphate-protective groups using concentrated aqueous ammonia. HPLC purification is oftentimes carried out first with the 5'-terminal DMT group still attached, effectively separating the oligomer from truncated sequences. This is then removed in solution with 80% acetic acid, lyophilized and re-purified.

1.3 Oligonucleotides and their analogues as modifiers of gene expression

Altering gene expression using oligonucleotides, or their analogues, targeted to mRNA is an attractive proposal for several basic reasons. Up until the late 1980s nucleic acids were often overlooked as drug targets and still today most therapeutics act at the protein level, where there may be thousands of targets stemming from one corresponding transcript. Preventing translation of a single transcript has the potential to affect the production of a multitude of aberrant proteins or proteins associated with a particular disease state. Easily predictable design rules, based on Watson-Crick hybridization,[94] could simplify the procedure of identifying a lead compound and, moreover, could be wholly selective of a specific gene. With the sequence data now available and an increasing understanding of

molecular processes occurring in the cell; the antisense concept has only become more seductive.

The first demonstration of the aforementioned principle was carried out in a cell-free eukaryotic translation system.[126] The authors found that the translation of β globin mRNA was reversibly arrested when hybridized with a complementary DNA sequence. Shortly thereafter, Zamenick and Stevenson found a therapeutic value in this phenomenon in 1978[127] through the inhibition of replication of the Rous sarcoma virus. Chick embryo fibroblasts infected with the virus were cultured in the presence of a 13-mer DNA complementary to the 5' and 3' ends of the viral 35S RNA. The presence of the oligomer was found to decrease the infectivity and replication of the virus; the authors proposed a steric block mechanism as a possible therapeutic modality. Many broader targets have been addressed and will be discussed later in this chapter; but first, some more background.

Mechanisms of antisense dug action.

There are essentially three ways in which an antisense drug may exert its desired effect which is the inhibition of protein synthesis. The most conceptually simple of these is termed the steric block mechanism. In this case, the association of the antisense oligonucleotide (ASO) to the target mRNA prevents, by its occupancy, the association or progression of the ribosome across the transcript. This mechanism is most effective when the 5'-untranslated region (UTR) is targeted as this is where translation initiation occurs. Certain areas of the transcript are required to be structured for effective splicing or translation initiation, disturbance of these structured areas can also be of therapeutic value. The steric block mechanism can also be used for exon skipping and this therapeutic modality is seeing great progress as treatment for Duchenne and Becker muscular dystrophies. In this case a protein coding region (exon) within the dystrophin gene contains a mutation that results in a non-functional, truncated protein. By the binding of an ASO to the 5'-end of an exon an alternate splice site may be chosen resulting in expression of a shorter but still functional protein.[128,129] For this strategy to be effective the ASO must obviously not recruit degradative enzymes so only some ASOs

are available for use.

The second method works by the recruitment of RNaseH enzymes. RNaseH is expressed ubiquitously and is present both in the nucleus and the cytoplasm,[130] its natural function is to cleave the RNA strand of a DNA/RNA heteroduplex, this is required during the removal of RNA template strands during lagging strand DNA synthesis and for Okazaki fragment processing. Because of the importance of these functions knockout of the gene is embryonic lethal.[131] It recognizes A-form duplexes and the silencing is generally more robust than the previous mechanism as it is catalytic; one ASO can induce the degradation of multiple transcripts. The enzyme has limited sequence specificity and can therefore be manipulated easily with synthetic ASOs to cleave target RNA. Many of the most therapeutically important ASOs operate by this mechanism. This mechanism can also see the use of analogues that do not directly recruit RNaseH. So called gapmers consist of a central DNA sequence flanked at its 5' and 3' termini with modifications that impart enzymatic stability, improved binding, and pharmacokinetics.[132]

Perhaps the most remarkable of gene silencing mechanisms and currently the most studied is RNAi. Because of its dissimilarity to the previous two mechanisms described it is often placed in a category of its own. Using small RNAs as a guide, a complementary sequence is cleaved or its rate of translation is otherwise diminished by a series of endogenous enzymes that are thought to have evolved as a defense against viruses. The phenomenon was first discovered in plants[133] and then later in Caenorhabditis elegans.[134] Shortly thereafter, Fire and Mello discovered that double stranded RNAs were the cause of the observed silencing for which they would later receive the Nobel Prize.[135] Importantly, they noted that only a few molecules were required to give significant reduction of protein levels. This then ruled out a stoichiometric mechanism and implied that there was a catalytic or amplification process in operation. It was later found that larger dsRNAs were converted to shorter sequences and it was these that induced silencing.[136] The exact mechanism of gene silencing will not be described in great detail here, but, in short, the process begins as small double stranded RNA is loaded onto a RNA induced silencing complex (RISC). In mammals, this may consist of simply argonaute protein 2 (Ago2) loaded with a short interfering RNA (siRNA) or a very large

multiprotein complex. The associating factors enhance or otherwise modulate the silencing activity.[137-139] Using the antisense strand as a guide, RISC finds the cognate sequence and cleaves it giving a 5'-terminal phosphate and 2'- and 3'-terminal hydroxyl groups. Although, as we will see, the more traditional, single-stranded ASOs have seen extensive chemical modification, RNAi is yet to see substantial benefit from this chemistry.

Oligonucleotide analogues.

Enticing as oligonucleotide therapeutics may seem for the treatment of various diseases, we are yet to see comensurate translation into the clinic. At its advent, it was thought that antisense technology was the beginning of the end for many disorders. However, there are several inherent difficulties with the antisense approach that have brought the organic chemist to the table. DNA and RNA are readily degraded by enzymes, can invoke unwanted immune response, and are, for the most part, unable to penetrate cells.

In the following paragraphs, I will describe different oligonucleotide analogues and highlight the mechanism(s) of antisense action to which they pertain. We can divide the chemical modifications into distinct categories. The first two are phosphate and sugar modified respectively. The so-called first generation of modified oligonucleotide analogues are phosphate modified. Phosphorothioates (PS) were first synthesized in the late 1960s[140] with the aim of stabilizing polyribonucleotides that were used to induce interferon and are the most widely used backbone-modified ASO for several reasons (Fig. 1-34). Sulfurization of the phosphite triester can be easily carried out as opposed to the oxidation step in the phosphoramidite approach of DNA synthesis using Beaucage's reagent[141] or others such as 3-((dimethylaminomethylidene)amino)-3*H*-1,2,4-dithiazole-3-thione (DDTT).[142] This sulfurization results in the formation of a chiral centre which can complicate the HPLC purification as a large number of diastereomers are created. However, the benefits of this modification far outweigh this minor problem. They are far more resistant to nucleolytic degradation and therefore have long half lives in serum and cells and binding to serum proteins prevents rapid renal clearance.[143] This protein-binding ability adds the benefit of increasing cellular uptake but is also the cause of their

toxicity.[144] These ASOs operate by eliciting RNase H activity; however, their affinity to target RNA is reduced as compared to DNA. Several uniformly modified PS-ASOs are in clinical trials and one has FDA approval (Vitravene®, a 21-mer used to treat cytomegalovirus retinitis in immunocompromized patients) and many more are employed in gapmers (these will be explained soon). The PS backbone has also been introduced into siRNA and shown to be an effective and stable substitution although toxic at higher concentrations.[145]

The methylphosphonate modification (Fig.1-34) was first pursued to evaluate the effects of charge on hybridization efficiency of nucleic acids[146] and it was thought that the lack of charge, making the molecule more lipophilic, would allow for better cell penetration. Like PS ASOs the methyphosphonates are present as a mixture of diastereomers. These diastereomers vary greatly in their affinity towards target nucleic acids and do not recruit RNaseH.[147]

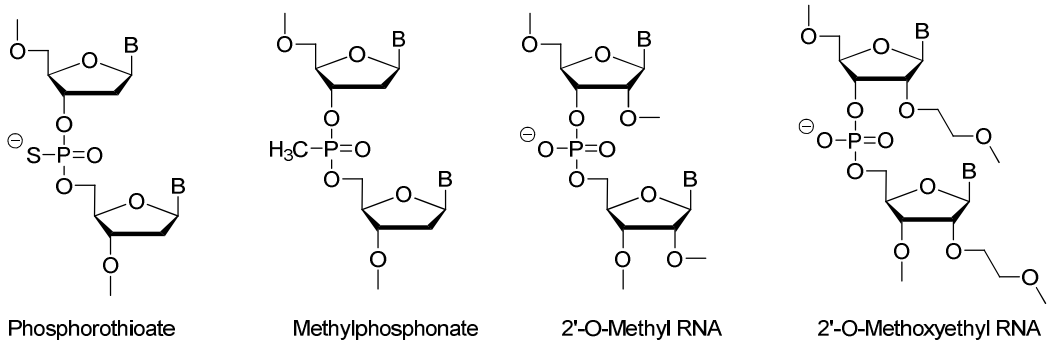


Figure 1-34: First (left) and second (right) generation antisense oligonucleotides (ASOs).

Two examples of second generation oligonucleotides are shown above (Fig. 1-34, right). 2'-O-Me and 2'-O-methoxyethyl RNAs[148,149] were sythesized to hopefully alleviate the shortcomings of the first generation ASOs. They maintain the 3'-endo conformation of RNA but the lack of an available proton of the 2'-O confers increased resistance to nucleases. Unfortunately they do not elecit RNaseH but may still be used in gapmers and are beginning to find success in RNAi experiments.

More dramatic departures from the natural structure of the sugar phosphate backbone

may be termed third generation ASOs, of which there are many. A select few of these structures are shown below in figure 1-35.

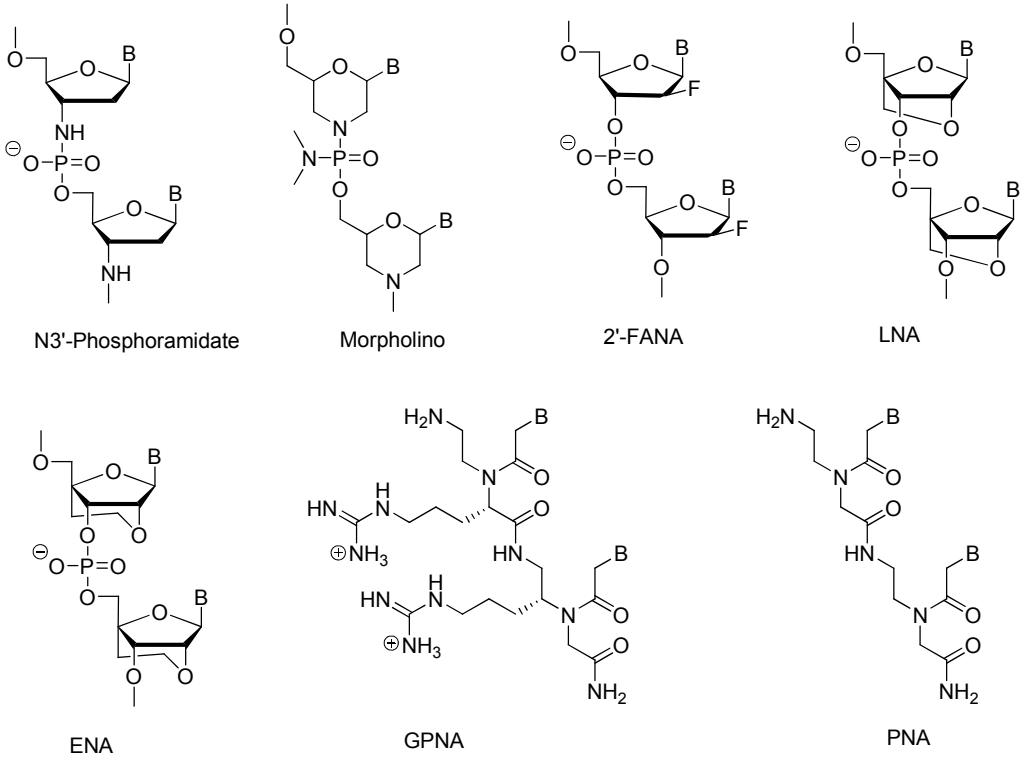


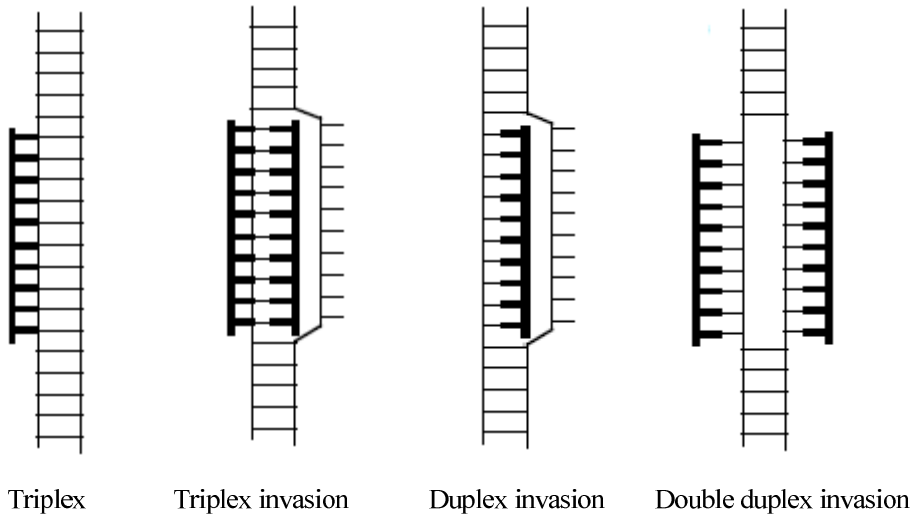
Figure 1-35: Third generation antisense oligonucleotides.

Wengel and Imanishi published simultaneously in 1998 on locked nucleic acid (LNA).[150,151] A methylene linker between 2'-O and 4'-C gives a conformationally stained nucleoside "locked" in the north or C-3' endo postion. This locking minimizes the entropic penalty associated with binding and results in high affinity towards DNA and RNA. It does not, however, elicit a response from RNaseH in and of itself but again can be used in gapmers or chimeras.

2'-Deoxy-2'-fluoroarabinonucleic acids (2'-FANAs) were a particularly rewarding modification for oligonucleotide therapeutics.[152] They were first oligomerized in 1993 by Watanabe's group.[153] Although arabinonucleic acids were found to hybridize poorly with DNA and RNA,[154] likely due to steric reasons, substitution of the C2'-hydroxyl with the

much smaller fluorine results with an analogue that binds strongly to both DNA and RNA.[155] The duplex formed with RNA is of A-type structure and is able to elicit RNA cleavage from RNaseH, this was a dramatic discovery at the time as up until that point only DNA/RNA heteroduplexes had been unequivocally proven as substates of this enzyme.[155] What is more is that the 2'-F modification imparts stability towards nucleases[156] and the electronegative fluorine also make the nucleosides highly resistant to glycosylases. This modification finds much use in gapmers, altimers and chimeras for RNaseH supported cleavage and in siRNA.[157-159]

Peptide nucleic acid (PNA) is a neutral polyamide first synthesized in 1991 by Nielsen and coworkers.[160] Its backbone departs dramatically from natural oligonucleotides; it is composed of *N*-(2-aminoethyl)glycine units to which the bases are attached via a carboxymethyl linkers (Fig.1-35). The molecule was originally designed for triplex formation but it was found that it can displace the opposing strand of a DNA duplex and follows Watson-Crick hydrogen bonding rules.[160,161] Partly because of its non-ionic nature, PNA binds to DNA and RNA with greater affinity than either do with themselves and has therefore been studied in great detail as an ASO operating by a steric block mechanism. PNA may bind to its targets in different ways depending on concentration and sequence identity (Fig. 1-36).



1-36: PNA binding modes withe PNA shown in bold.

PNA binds to DNA and RNA in an antiparallel fashion if one considers the N-terminus of the peptide to be pseudo-5'. This is not to say that parallel duplex formation is impossible, only that is a markedly less stable configuration.[161] Homopyrimidine PNA oligomers form highly stable triplex invasion complexes with target DNA while other sequences may form duplex invasion complexes as observed in Nielsen's initial publication. With the aid of the modified bases 2,6-diaminopurine and 2-thiouracil it is possible to form double duplex invasion species that prevent transcription and can also be used effectively as artificial restriction endonucleases.[162,163] These attributes make PNAs suitable for targeting DNA in anti-gene therapies as they have been shown to suppress transcription and,[164] curiously enough, activate it also.[165] When targeted to mRNA PNA can prevent translation if targeted to the 5'-UTR of the transcript and occasionally when targeted to the coding region. Due to its structural dissimilarity to any natural oligonucleotide or peptide it operates only by a steric block method as it is not recognized by RNaseH.[166] Although, as with other analogues, it can still be used in chimeras with analogues that do recruit RNaseH.[167] PNA and its applications will be discussed further in chapters 3, 4 and 7.

In short, the objectives of the work contained in the following chapters are twofold. We wish to increase the scope of cytosine analogues to be used as base-discriminating fluorophores. This will be approached in two ways. Firstly, using the Huisgen cycloaddition, we will generate a range of analogues in a modular fashion. In addition to this, we will modify the pyrrolocytosine scaffold to give a range of useful fluorophores for use in mismatch detection. The second objective is to design oligomers with high-affinity towards DNA and RNA, either by the attachement of a platinum centre or through the incorporation of a modified nucleobase to PNA, for use as tools for the modulation of gene expression.

References

1. Tyagi, S.; Kramer, S.R. *Nat. Biotechnol.* **1996**, *14*, 303-308.

2. Marti, A.A.; Jockusch, S.; Stevens, N.; Ju, J.Y.; Turro, N.J. *Acc. Chem. Res*. **2007**, *6*, 402-418.

3. Orum, H.; Nielsen, P.E.; Egholm, M.; Berg, R.H.; Buchardt, O.; Stanley, C. *Nucleic Acids Res*. **1993**, *21*, 5332-5336.

4. Chen, C.Y.; Shiesh, S.C.; Wu, S. *Clin. Chem.* **2004**, *50*, 481-489.

5. Dabritz, J.; Hanfler, J.; Preston, R.; Stieler, J.; Oettle, H. *Brit. J. Cancer* **2005**, *92*, 405-412.

6. Urata, M.; Wada, Y.; Kim, S.H.; Chumpia, W.; Kayamori, Y.; Hamasaki, N.; Kang, D. *Clin. Chem.* **2004**, *11*, 2045-2051.

7. Kumamoto, S.; Watanabe, M.; Kawakami, N.; Nakamura, M.; Yamana, K. *Bioconjugate Chem*. **2008**, *19*, 65-69.

8. Drummond, T.G.; Hill, M.G.; Barton, J.K. *Nat. Biotechnol.* **2003**, *21*, 1192-1199.

9. Okamoto, A.; Saito, Y.; Saito, I. *J. Photochem. Photobiol.* **2005**, *6*, 108-122.

10. Park, S.H.; Krull, U. *Anal. Chim. Acta.* **2006**, *564*, 133-140.

11. Callis, P.R. *Annu. Rev. Phys. Chem.* **1983**, *34*, 329-357.

12. Greco, N.J.; Tor, Y. *Tetrahedron* **2007**, *63*, 3515-3527.

13. Hudson, R.H.E.; Ghorbani-Choghamarani, A. *Synlett* **2007**, 870-873.

14. Okamoto, A.; Tainaka, K.; Saito, I. Monitoring DNA *J. Am. Chem. Soc.* **2005**, *127*, 13128-13129.

15. Ward, D.C.; Reich, E.; Stryer, L. *J. Biol. Chem.* **1969**, *244*, 1228-1237.

16. Rist, M.J.; Marino, J.P. *Current Org. Chem.*, **2002** *9*, 775-793.

17. Ballin, J.D.; Bharill, S.; Fialcowitz-White, E.J.; Gryczynsky, I.; Gryczynsky, Z.; Wilson, G.M. *Biochemistry* **2007**, *46*, 13948-13960.

18. Da Costa, C.P.; Fedor, M.J.; Scott, L.G. *J. Am. Chem. Soc.* **2007**, *129*, 3426-3432.

19. Miyaura, N.; Suzuki, A. *Chem. Commun.* **1979**, 866-867.

20. Sun, K.M.; McLaughlin, C.K.; Lantero, D.R.; Manderville, R.A. *J. Am. Chem. Soc*. **2007**, *129*, 1894-1895.

21. McLaughlin, C.K.; Lantero, D.; Manderville, R.A. *J. Phys. Chem.* **2006**, *110*, 6224-6230.

22. Vrábel, M.; Pohl, R.; Klepetářová, B.; Votruba, I.; Hocek, M. *Org. Biomol. Chem*. **2007**, *5*, 2849-2857.

23. Dyrager, C.; Börjesson, K.; Dinér, P.; Elf, A.;Albinsson, B.; Wilhelmsson, L.M.; Grøtli, M. *Eur. J. Org. Chem*. **2009**, 1515-1521.

24. Dierckx, A.; Dinér, P.; El-Sagheer, A.H.; Kumar, J.D.; Brown, T., Grøtli, M.; Wilhelmsson, L.M. *Nucleic Acids Res*. 2011 doi: 10.1093/nar/gkr010.

25. Fujiwara, T.; Kimoto, M.; Sugiyama, H.; Hirao, I.; Yokoyama, S. *Bioorg. Med. Chem. Lett*. **2001**, *11*, 2221-2223.

26. Hirao, I.; Ohtsuki, T.; Fujiwara, T.; Mitsui, T.; Yokogawa, T.; Okuni, T.; Nakayama, H.; Takio, K.; Yabuki, T.; Kigawa, T.; Kodama, K.; Yokogawa, T.; Nishikawa, K.; Yokoyama, S. *Nat. Biotechnol*. **2002**, *20*, 177183.

27. Mitsui, T.; Kimoto, M.; Kawai, R.; Yokoyama, S.; Hirao, I. *Tetrahedron* **2007**, *63*, 3528-3537.

28. Hopkins, F.G. Phil. *Trans. Rot. Soc., London, Ser. B.* **1895**, *186*, 661-663.

29. Hawkins, M.E., Pfleiderer, W., Jungmann, O., Balis, F.M. *Anal. Biochem.* **2001**, *298*,

231-240.

30. Hawkins, M.E. *Cell Biochem. Biophys.* **2001**, *34*, 257-281.

31. Hudson, R.H.E.; Ghorbani-Choghamarani, A. *Org. Biomol. Chem.* **2007**, *5*, 1845-1848.

32. Greco, N.J.; Tor, Y. *J. Am. Chem. Soc.* **2005**, *127*, 10784-10785.

33. Xiao, Q.; Ranasinghe, R.T.; Tang, A.M.P.; Brown, T. *Tetrahedron* **2007**, *63*, 3483-3490.

34. Sessler, J.L; Sathiosatham, M.; Brown, C.T.; Rhodes, T.A.; Wiederrecht, G. *J. Am. Chem. Soc*. **2001**, *123*, 3655-3660.

35. Ryu, J.H.; Seo, Y.J.; Hwang, G.T.; Lee, J.Y.; Kim, B.H. *Tetrahedron* **2007**, *63*, 3538-3547.

36. Wojciechowski, F.; Hudson, R.H.E. *Nucleos. Nucleot. Nucl.* **2007**, *26*, 1199-1202.

37. Weber, G.; Farris, F.J. *Biochemistry* **1979**, *18*, 3075-3078.

38. Kimura, T.; Kawai, K.; Majima, T. *Org. Lett*. **2005**, *7*, 5829-5832.

39. Kimura, T.; Kawai, K.; Majima, T. *Chem. Commun*. **2006**, 1542-1544.

40. Okamoto, A.; Tainaka, K.; Unzai, T.; Saito, I. Synthesis *Tetrahedron* **2007**, *63*, 3465-3470.

41. Wigerinck, P.; Pannecouque, C.; Snoeck, R.; Claes, P.; De Clercq, E.; Herdewijin, P. *J. Med. Chem.* **1991**, *34*, 2383-2389.

42. Srivatsan, S.G.; Tor, Y. *J. Am. Chem. Soc.* **2007**, *129*, 2044-2053.

43. Reichardt, C. *Chem. Rev*. **1994**, *94*, 2319-2358.

44. Sinkeldam, R.W.; Tor, Y. *Org. Biomol. Chem*. **2007**, *5*, 2523-2528.

45. Greco, N.J.; Sinkeldam, R.W.; Tor. Y. *Org. Lett.* **2009**, *11*, 1115-1118.

46. Sinkeldam, R.W.; Greco, N.J.; Tor, Y. *ChemBioChem* **2008**, *9*, 706-709.

47. Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem. Int. Ed.* **2001**, *40*, 2004-2021.

48. Kocalka, P.; Andersen, N.K.; Jensen, F.; Nielsen, P. *ChemBioChem* **2007**, *17*, 2106-2116.

49. Seela, F.; Sirivolu, V.R.; Chittepu, P. *Bioconjugate Chem*. **2008**, *19*, 211-221.

50. Gramlich, P.M.E.; Wirges, C.T.; Manetto, A.; Carell, T. *Angew. Chem. Int. Ed.* **2008**, *47*, 8350-8358.

51. Park S.M.; Lee Y.S.; Kim B.H. *Chem. Commun.* **2003**, 2912-2913.

52. Park S.M.; Shen Y.; Kim B.H. *Org. Biomol. Chem.* **2007**, *5*, 610-612.

53. Dodd, D.W.; Swanick, K.N.; Price, J.T.; Brazeau, A.L.; Ferguson, M.J.; Jones, N.D.; Hudson, R.H.E. *Org. Biomol. Chem.* **2010**, *8*, 663-666.

54. Gutierrez, A.J.; Terhorst, T.J.; Matteucci, M.D.; Froehler, B.C. *J. Am. Chem. Soc.* **1994**, *116*, 5540-5544.

55. Capobianco, M.L.; Cazzato, A.; Alesi, S.; Barbarella, G. *Bioconjugate Chem*. **2008**, *19*, 171-177.

56. Leonard, N.J. *Acc. Chem. Res.* **1982**, *15*, 128-135.

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

58. Liu, H.; Gau, J.; Maynard, L.; Saito, Y.D.; Kool, E.T. *J. Am. Chem. Soc.* **2004**, *126*, 1102-1109.

59. Lee, A.H.F.; Kool, E.T. *J. Am. Chem. Soc.* **2006**, *128*, 9219-9230.

60. Tor, Y.; Del Valle, S.; Jaramillo, D.; Srivatsan, S.G.; Rios, A.; Weizman, H. *Tetrahedron* **2007**, *63*, 3608-3514.

61. Srivatsan, S.G.; Weizman, H; Tor, Y. *Org. Biomol. Chem.* **2008**, *6*, 1334-1338.

62. Srivatsan, S.G.; Greco, N.J.; Tor, Y. *Angew. Chem. Int. Ed.* **2008**, *47*, 6661-6665.

63. Curd, F.H.S.; Landquist, J.K.; Rose, F.L. *J. Chem. Soc*. **1948**, 1759-1766.

64. Godde, F.; Toulme, J. ; Moreau, S. *Biochemistry* **1998**, *37*, 13765-13775.

65. Moser, H.E.; Dervan, P.B. *Science* **1987**, *238*, 645-650.

66. Le Doan, T.; Perrouault, L.; Praseuth, D.; Habhoub, N.; Decout, J.L.; Thuong, N.T.; Lhomme, J.; Hélène, C. *Nucleic Acids Res.* **1987**, *15*, 7749-7760.

67. Major, D.T.; Fischer, B. *J. Phys. Chem. A* **2003**, *107*, 8923-8931.

68. Fischer, B.; Kabha, B.; Gendron, F.P.; Beaudoin, A.R. *Nucleos. Nucleot. Nucl.* **2000**, *19*, 1033-1054.

69. Barrio, J.R.; Secrist, J.A.; Leonard, N.J. *Proc. Natl. Acad. Sci. USA* **1972**, *69*, 2039-2042.

70. Secrist, J.A.; Barrio, J.R.; Leonard, N.J. *Science* **1972**, *175*, 646-647.

71. Seela, F.; Schweinberger, E.; Xu, K.; Sirivolu, V.R.; Rosemeyer, H.; Becker, E.M. *Tetrahedron* **2007**, *63*, 3471-3482.

72. Leonard, G.A.; McAuley-Hecht, K.E.; Gibson, N.J.; Brown, T.; Watson, W.P.; Hunter, W.N. *Biochemistry* **1994**, *33*, 4755-4761.

73. Lin, K.; Jones, R.J.; Matteucci, M. *J. Am. Chem. Soc.* **1995**, *117*, 3873-3874.

74. Roth, B.; Schloemer, L.A. *J. Org. Chem.* **1963**, *28*, 2659-2672.

75. Wilhelmsson, L.M.; Holmén, A.; Lincoln, P.; Nielsen, P.E.; Nordén, B. *J. Am. Chem. Soc.* **2001**, *123*, 2434-2435.

76. Sandin, P.; Wilhelmsson, L.M.; Lincoln, P.; Powers, V.E.C.; Brown, T.; Albinsson, B.

Nucleic Acids Res. **2005**, *33*, 5019-5025.

77. Sandin, P.; Börjesson, K.; Li, H.; Mårtensson, J.; Brown, T.; Wilmelmsson, L.M.; Albinsson, B. *Nucleic Acids Res.* **2008**, *36*, 157-167.

78. Stengel, G.; Gill, J.P.; Sandin, P.; Wilhelmsson, L.M.; Albinsson, B.; Nordén, B.; Millar, D. *Biochemistry* **2007**, *46*, 12289-12297.

79. Barhate, N.; Cekan, P.; Massey, A.P.; Sigurdsson, S.N. *Angew. Chem. Int. Ed.* **2007**, *46*, 2655-2658.

80. Cekan, P.; Sigurdsson, S.Th. *Chem. Commun*. **2008**, 3393-3395.

81. Lin, K.-Y.; Matteucci, M. *J. Am. Chem. Soc.* **1998**, *120*, 8531-8532.

82. Holmes, S.C.; Arzumanov, A.A.; Gait, M.J. *Nucleic Acids Res*. **2003**, *31*, 2759-2768.

83. Ortega, J.A.; Blas, J.R.; Orozco, M.; Grandas, A.; Pedroso, E.; Robles, J. *Org. Lett.* **2007**, *9*, 4503-4506.

84. Sazani, P.; Astriab-Fischer, A.; Kole, R. *Antisense Nucleic A*. **2003**, *13*, 119-128.

85. Wilds, C.J.; Maier, M.A.; Tereshko, V.; Manoharan, M.; Egli, M. *Angew. Chem. Int. Ed.* **2002**, *41*, 115-117.

86. Rajeev, K.G.; Maier, M.A.; Lesnik, E.A.; Manoharan, M. *Org. Lett.* **2002**, *4*, 4395-4398.

87. Nakagawa, O.; Ono, S.; Tsujimoto, A.; Li, Z.; Sasaki, S. *Nucleos. Nucleot. Nucl*. **2007**, *26*, 645-649.

88. Wilson, J.N.; Kool, E.T. *Org. Biomol. Chem.* **2006**, *4*, 4265-4274.

89. Millar, D.P. *Curr. Opin. Struct. Biol.* **1996**, *6*, 322-326.

90. Silverman, A.P.; Kool, E.T. *Chem. Rev.* **2006**, *106*, 3775-3789.

91. Dodd, D.W.; Hudson, R.H.E. *Mini-Rev. Org. Chem.* **2009**, *6*, 378-391.

92. Sinkeldam, R.W.; Greco, N.; Tor, Y. *Chem. Rev.* **2010**, *110*, 2579-2619.

93. Wilhelmsson, L.M. *Q. Rev. Biophys.* **2010**, *43*, 159-183.

94. Watson, J.D.; Crick, F.H.C. *Nature* **1953**, *171*, 737-738.

95. Miescher, F. *Medicinisch-chemische Untersuchungen* **1871**, *4*, 441–460.

96. Kossel, A. *Untersuchungen über die Nukleine und ihre Spaltungsprodukte,* **1881**.

97. Levene, P.A.; Mikeska, L.A.; Mori, T.J. *J. Biol. Chem.* **1930**, *85*, 785-787.

98. Levene, P.A.; Tipson, R.S. *Science* **1935**, *25*, 98-99.

99. Griffith, F. *J. Hygiene* **1928**, *27*, 113.

100. Avery, O.T.; MacLeod, C.M.; McCarty, M. *J. Exp. Med.* **1944**, *79*, 137–158.

101. Hershey, A.D.; Chase, M. *J. Gen. Physiol.* **1952**, *36*, 39–56.

102. Chargaff, E. *Experientia* **1950**, *6*, 201-209.

103. Chargaff, E. *Fed. Proc.* **1951**, *10*, 654-659.

104. Clark, V.M.; Todd, A.R.; Zussman, J. *J. Chem. Soc.* **1951**, 2952-2958.

105. Brown, D.M.; Todd, A.R. *J. Chem. Soc.* **1952**, 52-58.

106. Michelson, A.M.; Todd, A.R. *J. Chem. Soc.* **1955**, 2632-2638.

107. Pauling, L.; Corey, R.B. *Proc. Natl. Acad. Sci USA*. **1953**, *38*, 84-97.

108. Reese, C.B. *Org. Biomol. Chem*. **2005**, *3*, 3851, 3868.

109. Khorana, H.G.; Tener, G.M; Moffat, J.G.; Pol, E.H. *Chem. Ind.* **1956**, 1523.

110. Khorana, H.G.; Razzell, W.E.; Gilham, P.T.; Tener, G.H.; Pol, E.H. *J. Am. Chem. Soc.* **1957**, *79*, 1002-1003.

111. Schaller, H.; Weimann, G.; Lerch, B.; Khorana, H.G. *J. Am. Chem. Soc*. **1963**, *85*, 3821-3827.

112. Letsinger, R.L.; Ogilvie, K.K. *J. Am. Chem. Soc*. **1967**, *89*, 4801, 4803.

113. Letsinger, R.L.; Ogilvie, K.K.; Miller, P.S. *J. Am. Chem. Soc.* **1969**, *91*, 3360-3365.

114. Letsinger, R.L.; Lunsford, W.B. *J. Am. Chem. Soc*. **1976**, *98*, 3655-3661.

115. Beaucage, S.L.; Caruthers, M.H. *Tetrahedron Lett*. **1981**, *22*, 1859-1862.

116. McBride, L.J.; Caruthers, M.H. *Tetrahedron Lett*. **1983**, *24*, 245-248.

117. Sinha, N.D.; Biernat, J.; Köster, H. *Tetrahedron Lett*. **1983**, *24*, 5843-5846.

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

119. Merrifield, R.B. *J. Am. Chem. Soc*. **1964**, *86*, 304-305.

120. Atherton, E.; Clive, D.L.; Sheppard, R.C.; *J. Am. Chem. Soc.* **1975**, *97*, 6584-6587.

121. Gait, M.J.; Sheppard, R.C. *J. Am. Chem. Soc*. **1976**, *98*, 8154-8516.

122. Carpino, L.A.; Han, G.Y. *J. Org. Chem*. **1972**, *37*, 3404-3409.

123. Letsinger, R.L.; Mahadevan, V. *J. Am. Chem. Soc.* **1965**, *87*, 3256-3527.

124. Gough, G.R., Brunden, M.J. Gilham, P.T. *Tetrahedron Lett*. **1981**, *22*, 4177-4180.

125. Tsou, D.; Hampel, A.; Andrus, A.; Vinayak, R. *Nucleos. Nucleot*., **1995**, *14*, 1481-1492.

126. Paterson, B.M.; Roberts, B.E.; Kuff, E.L. *Proc. Natl. Acad. Sci. USA* **1977**, *74*, 4370-4374.

127. Zamenick, P.C.; Stevenson, M.L. *Proc. Natl. Acad. Sci. USA*, **1978**, *75*, 280-284.

128. Sazani, P.; Kole, R. *J. Clin. Invest*. **2003**, *112*, 481-486.

129. Wang, Q.S.; Yin, H.F.; Camelliti, P.; Betts, C.; Moulton, H.; Lee, H.; Saleh, A.F.; Gait, M.J.; Wood, M.J.A. *J. Gene Med*. **2010**, *12*, 354-364.

130. Wu, H.; Lima, W.F.; Zhang, H.; Fan, A.; Sun, H.; Crooke, S.T. *J. Biol. Chem*. **2004**, *279*, 17181-17189.

131. Cerritelli, S.M.; Frovola, E.G.; Feng, C.; Grinberg, A.; Love, P.E.; Crouch, R.J. *Mol. Cell* **2003**, *11*, 807-815.

132. McKay,R.A.; Miraglia, L.J.; Cummins, L.L.; Owens, S.R.; Sasmor, H.; Dean N.M. *J. Biol. Chem*. **1999**, *274*, 1715-1722.)

133. Napoli, C.; Lemiex, C.; Jorgensen, R. *Plant Cell* **1990**, *2*, 279-290.

134. Guo, S.; Kemphues, K.J. *Cell* **1995**, *81*, 611-620.

135. Fire, A.; Xu, S.; Montgomery, M.K.; Kostas, S.A.; Driver, S.E.; Mello, C.C. *Nature* **1998**, *391*, 806-811.

136. Hamilton, A.J.; Baulcombe, D.C. *Science* **1999**, *286*, 950-952.

137. Martinez, J.; Patkaniowska, A.; Urlaub, H.; Lührmann, R.; Tuschl, T. *Cell* **2002**, *110*, 563-574.

138. Rivas, F.V.; Tolia, N.H.; Song, J.J.; Aragon, J.P.; Liu, J.; Hannon, G.J.; Joshua-Tor, L. *Nat. Struct. Biol*. **2005**, *12*, 340-349

139. Pham, J.W.; Pellino, J.L.; Lee, Y.S.; Carthew, R.W.; Sontheimer, E.J. *Cell* **2004**, *117*, 83-94.

140. DeClercq, E.; Eckstein, F.; Merigan, T.C. *Science* **1969**, *165*, 1137-1140.

141. Iyer, R.P.; Egan, W.; Regan, J.B.; Beaucage, S.L. *J. Am. Chem. Soc*. **1990**, *112*, 1253-1254.

142. Guzaev, A.P. *Tetrahedron Lett.* **2011**, *52*, 434–437.

143. Stein, C.A.; Subasinghe, C.; Shinozuka, K.; Cohen, J. S. *Nucleic Acids Res*. **1988**, *16*, 3209-3221.

144. Geary, R.S.; Yu, R.Z.; Levin, A.A. *Curr. Opin. Investig. Drugs* **2001**, *2*, 562-573.

145. Braasch, D.A.; Jensen, S.; Liu, Y.H.; Kaur, K.; Arar, K.; White, M.A.; Corey, D.R. *Biochemistry* **2003**, *42*, 7967-7975.

146. Barrett, J.C.; Miller, P.S.; Ts'o, P.O. *Biochemistry* **1974**, *13*, 4897-4906.

147. Lesnikowski, Z.J.; Jaworska, M.; Steck, W. *Nucleic Acids Res*. **1990**, *18*, 2109-2115.

148. Baker, B.F.; Lot, S.S.; Condon, T.P.; Cheng-Fluornoy, S.; Lesnik, E.A.; Sasmor, H.M.; Bennet, C.F.; *J. Biol. Chem*. **1997**, *272*, 11994-11998.

149. Crooke, S.T.; Lemonidis, K.M.; Nielson, L.; Griffey, R.; Lesnick, E.A.; Monia, B.P. *Biochem. J.* **1995**, *312*, 599-602.

150. Koshkin, A.A.; Rajwanski, V.K.; Wengel, J. *Tetrahedron Lett*. **1998**, *39*, 4381-4382.

151. Obika, S.; Nanbu, D.; Hari, Y.; Andoh, J.-I.; Morio, K.-I.; Doi, T.; Imanishi, T. *Tetrahedron Lett*. **1998**, *39*, 5401-5404.

152. Watts, J.K.; Damha, M.J. *Can. J. Chem*. **2008**, *86*, 641-656.

153. Rosenberg, I.; Soler, J.F.; Tocik, Z.; Ren, W.Y.; Ciszewski, L.A.; Kois, P.; Pankiewicz, K.W.; Spassova, M.; Watanabe, K.A. *Nucleos. Nucleot*. **1993**, *12*, 381-401.

154. Noronha, A.M.; Wilds, C.J.; Lok, C.-N.; Viazkovina, K.; Arion, D.; Parniak, M.A.; Damha, M.J. *Biochemistry* **2000**, *39*, 7050-7062.

155. Damha, M.J.; Wilds, C.J.; Noronha, A.; Bruckner, I.; Borkow, G.; Arion, D.; Parniak, M.A. *J. Am. Chem. Soc*. **1998**, *120*, 12976-12977.

156. Kois, P.; Tocik, Z.; Spassova, M.; Ren, W.Y.; Rosenberg, I.; Soler, J.F.; Watanabe, K.A. *Nucleos. Nucleot*. **1993**, *12*, 1093-1109.

157. Dowler, T.; Bergeron, D.; Tedeschi, A.L.; Paquet, L.; Ferrari, N.; Damha, M.J. Nucleic Acids Res. 2006, 34, 1169-1175.

158. Watts, J.K.; Choubdar, N.; Sadalpure, K.; Robert, F; Wahba, A.S.; Pelletier, J.; Pinto, B.M.; Damha, M.J. *Nucleic Acids Res*. **2007**, *35*, 1441-1447.

159. Watts, J.K.; Deleavey, G.F.; Damha M.J. *Drug Discov. Today* **2008**, *13*, 842-855.

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

161. Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S.M; Driver, D.A.; Berg, R.H.; Kim, S.K.; Norden, B.; Nielsen, P.E. *Nature* **1993**, *365*, 566-568.

162. Nielsen, P.; Egholm, M.; Buchardt, O. J. Mol. Recognit. 1994, 7, 165-170.

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

164. Janowski, B.A.; Kaihatsu, K.; Huffman, K.E.; Schwartz, J.C.; Ram, R.; Hardy, D.; Mendelson, C.R.; Corey D.R. *Nat. Chem. Biol*. **2005**, *1*, 210-215.

165. Møllengaard, N.E.; Buchardt, O.; Egholm, M.; Nielsen, P.E. *Proc. Natl. Acad. Sci. USA*, **1994**, *91*, 3892-3896.

166. Shiraishi, T.; Nielsen, P.E. *Nucleic Acids Res*. **2004**, *32*, 1921-1927.

167. Pâtureau, B.; Hudson, R.H.E.; Damha, M.J. *Bioconjugate Chem*. **2007**, *18*, 421-430.

Chapter 2

2 Synthesis and evaluation of the *in vitro* DNA binding properties of chiral *cis*-dichloro(pyridyloxazoline)platinum(II) complexes

A series of chiral *cis*-dichloro(pyridyloxazoline)platinum(II) and palladium(II) complexes were synthesized and their reactivity towards a defined sequence of single-stranded- and double-stranded DNA was investigated in comparison to cisplatin. The compounds differed in the nature and absolute configuration of the substituent at the *C*4 position of the oxazoline ring. The DNA binding ability of these compounds was evaluated by HPLC analysis, post metal exposure, of enzymatic digests of an undecamer duplex containing one putative metallation site. Polyacrylamide gel electrophoresis (PAGE) and thermal denaturation confirmed the results of the HPLC analysis which showed that the stereochemistry and character of the substituent at the *C*4 position of the oxazoline ring had little effect on DNA binding, possibly due to the formation of monofunctional adducts. Excerpts of this chapter have been taken from Dodd, D.W.; Toews, H.E.; Trevail, M.J.; Jennings, M.C.; Hudson, R.H.E.; Jones, N.D.. *Can. J. Chem.* **2009**, *87*, 321-327.

2.1 Introduction

The serendipitous discovery of the antiproliferative properties of *cis*-dichlorodiammine-platinum(II) (cisplatin) over 30 years ago led to its eventual use as a chemotherapeutic agent.[1] Today, cisplatin remains one of the most widely used drugs for the treatment of solid cell tumours.[2] Motivated by cisplatin's effectiveness, similar treatments have been pursued; however, despite vigorous research towards this end, only a handful of successful analogues have been developed.[3] While complexes of other metals such as titanium, rhenium, ruthenium and gallium have shown promise as potential antineoplastic agents, the literature of metal anticancer theraputics is dominated by platinum complexes.[4] The efficacy of cisplatin in the treatment of carcinomata is almost universally thought to be due to its ability to form intrastrand adducts between adjacent

purines in DNA.[2-5] The design principles of cisplatin analogues were described two years after the first clinical trials: the complexes should be neutral and four coordinate with two coordination sites occupied by labile leaving ligands and two occupied by inert am(m)ine ligand(s) in a *cis* geometry.[6]

Once cisplatin enters the cell, the lower intracellular chloride concentration permits dissociation of the halides resulting in the formation of an aquated platinum centre. This cationic species reacts predominantly with the *N*7 of guanine or adenine nucleobases.[5] The guanine adduct is thought to be stabilized through hydrogen bonding of an ammine proton with the carbonyl oxygen of *C*6 making deoxyguanosine dinucleotide (-dGpdG-) the preferred target sequence.[7] Binding to DNA in this fashion causes a kinking of the double helix and a widening of the minor groove; this can inhibit transcription and lead to apoptosis by inducing the binding of high mobility group (HMG) box proteins (5). Although cisplatin is indeed an effective drug, there are problems associated with prolonged dosages such as toxicity and acquired resistance. Therefore the search continues for improved platinum based therapies.

While several thousand cisplatin analogues have been synthesized and tested for biological activity,[8] only a select few have been entered into clinical trials. This has led many in the scientific community to believe that the search for Pt(II) chemotherapeutics is futile. Discovery of a drug suitable for clinical trials is a rare occurrence in every field, yet failure in cisplatin analogue discovery draws inordinate attention. Perhaps the reason for the reluctance to continue in this area is due to the remarkable simplicity of cisplatin itself: even minor substitutions cause drastic changes in efficacy. Although the structure of cisplatin is simple, the mode of action certainly is not, and many questions remain unanswered.[5]

Although it is unlikely that a simple substitution of the ammine ligands will yield a complex with activity surpassing that of cisplatin, more elaborate changes, such as variation in the σ-donation character of the ligand and the introduction of chirality could yield complexes with novel reactivities *in vitro* and corresponding cytostatic effects *in vivo*.

As DNA is a chiral molecule, binding of chiral platinum complexes results in diastereomeric adducts. Different binding affinities for enantiomeric complexes have been reported and (of no lesser importance) it has also been found that the diastereomeric adducts may be dealt with by cellular machinery in vastly different ways.[9] For example, the *S,S* enantiomer of (1,2-diaminocyclohexane)oxalatoplatinum(II) (*S,S*-**I**, Fig. 2-1) is significantly more mutagenic than the *R,R* enantiomer and consequently only the latter was pursued in clinical trials.[9,10] Other examples of chiral DNA-binding Pt complexes that have been studied in detail include dichloro(*N*-methyl-2-aminomethylpyrrolidine)platinum(II) (**II**, Fig. 2-1)[10] and complexes involving 1,2-bis(aminomethyl)carbobicyclic ligands (**III**, Fig. 2-1).[11] Differences in reactivities of enantiomeric complexes towards DNA are rare.[12]

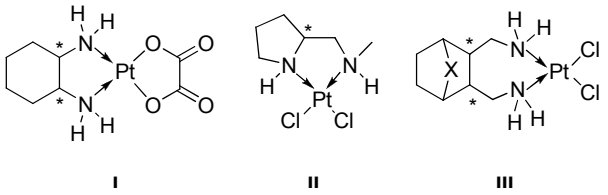
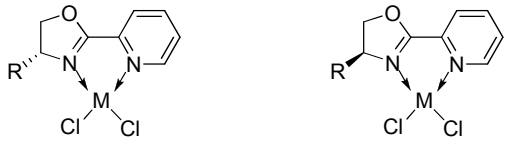


Figure 2-1: Chiral platinum complexes I-III studied as cisplatin analogues.[9-11] Chiral centres are denoted by asterisks.

Although hydrogen bonding involving coordinated NH groups have been considered necessary for cytostatic activity,[13] there are examples of compounds lacking an NH proton that display significant antitumour effects.[14,15]

Pyridyloxazolines, derived from the condensation of a chiral β-amino alcohol with 2-cyanopyridine, are planar, bidentate ligands for platinum(II) and are necessarily bound *cis* (Fig. 2-2). The substituent at the α-position of the amino alcohol, *C*4 in the resultant



R-**1** R = iPr, M = Pt
R-**2** R = Bn, M = Pt
R-**3** R = Ph, M = Pt
R-**4** R = iPr, M = Pd

S-**1** R = iPr, M = Pt
S-**2** R = Bn, M = Pt
S-**3** R = Ph, M = Pt
S-**4** R = iPr, M = Pd
S-**5** R = sBu, M = Pt
6 R = H, M = Pt

Figure 2-2: Pyridyloxazoline platinum and palladium dichlorides employed in this study. The chiral centre is derived from the enantiomerically pure amino acid.

oxazoline ring, can be varied greatly and inexpensively as the alcohols are obtained by the reduction of the corresponding enantiopure amino acids.[16,17]

Pyridyloxazoline Pt(II) complexes have found use in enantioselective enyne cyclizations[18] and the unsubstituted complex **6** has shown some promise as an anticancer therapeutic, displaying significant toxicity towards rat brain tumour cell lines *in vitro*.[19] The palladium analogues have not been used in DNA binding studies, but have been investigated as asymmetric catalysts for the intermolecular asymmetric Heck reaction.[16]

In this work, binding of these complexes to a defined sequence of oligomeric double stranded (ds) DNA containing one putative metallation site was investigated by HPLC analysis of enzymatic digests. The duplexes that were exposed to the metal complexes were also analyzed by thermal denaturation and polyacrylamide gel electrophoresis (PAGE).

Difference in binding abilities between enantiomers and sidechain identity was investigated, the objective being to gain a better understanding of the SAR of chiral platinum antineoplastic agents.

2.2 Results and discussion

The modular nature of the pyridyloxazoline ligand scaffold, combined with its brief synthesis from readily-available enantiopure starting materials and facile metal coordination, permitted the rapid development of small libraries of enantiomeric Pt(II) and Pd(II) complexes. The known pyridyloxazoline ligands required for the syntheses of complexes **1**(**4**)–**6** were made by the $ZnCl_2$-catalyzed coupling of 2-cyanopyridine with the β-amino alcohols, valinol, phenylalaninol, phenylglycinol, isoleucinol, and

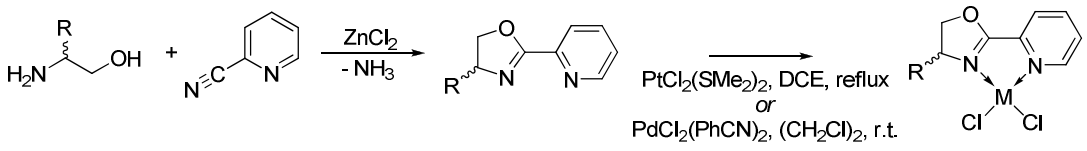


Figure 2-3: Synthesis of dichloro(pyridyloxazoline)metal(II) complexes.

ethanolamine, respectively, using the procedure developed by Bolm *et al*. (Fig. 2-3).[17] The amino alcohols, except for ethanolamine, were themselves readily made by $NaBH_4/I_2$-reduction of both enantiomers of the corresponding amino acids, during which time no racemization occurs.[26] The metal(II) complexes were made in good-excellent yields by the reaction of slight molar excesses of the ligands either with cis/trans-$PtCl_2(SMe_2)_2$ or trans-$PdCl_2(PhCN)_2$ in refluxing 1,2-dichloroethane or CH_2Cl_2 at r.t. The complexes were yellow-orange, microcrystalline solids that were freely soluble in polar organic solvents, such as $CHCl_3$ and CH_2Cl_2, and insoluble in Et_2O and hexanes. All complexes were also sufficiently soluble in H_2O to allow DNA-binding studies.

The molecular structure of *S*-**1** is shown in Figure 2-4; it is isostructural with the known precisely analogous palladium derivative.[16] As expected, the metal adopts an approximately square planar coordination geometry, and the pyridyloxazoline ligand is roughly planar except for the *iso*-propyl substituent at the 4-position of the oxazoline ring which projects prominently from one face of the complex and makes a torsion angle with it of ca. 62 °. The ligand has a "bite angle" of 79.711(4) °, which is nearly identical to that found for the palladium analogue, but unlike in that compound, whose M–oxazoline bond (2.017(5) Å) is significantly shorter than its M–pyridine bond (2.048(6) Å), there is no difference in the two M–N distances, which are both 2.015(11, 12) Å in *S*-**1**. However, there is, as in the palladium analogue, a significant difference in the M–Cl distances: in complexes of both metals, the M–Cl bond trans to the oxazoline ring is longer than that trans to the pyridine ring, which consistently suggests that oxazoline is a stronger field ligand than pyridine, and implies strong contribution from the resonance form that places a formal negative charge on the N-atom of the free oxazoline.

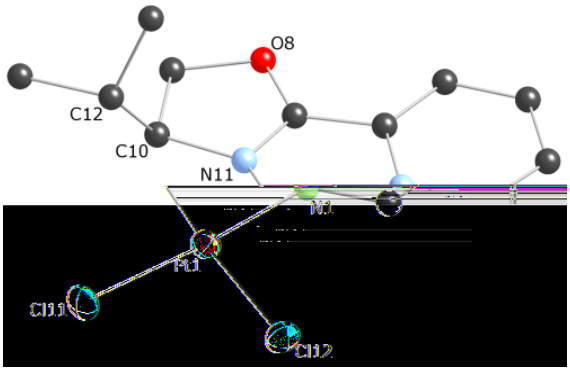


Figure 2-4: Partial ORTEP representation of the molecular structure of S-1, with H-atoms omitted for clarity. Thermal ellipsoids for Pt- and Cl-atoms are shown at 50 % probablility; all other atoms are refined isotropically (see Experimental section). Selected bond distances (Å) and angles (°), with estimated standard deviations in parentheses: Pt1–N1 2.015(11), Pt1–N11 2.015(12), Pt1–Cl11 2.2869(3), Pt1–Cl12 2.298(4), N1–Pt1–N11 79.711(4), Pt1–N11–C10–C12 61.629(19).

The particular undecamer sequence chosen for the present studies (shown below) has been used by Lippard and co-workers to obtain a solution NMR structure of a platinated duplex and is known to adopt the typical B-form of DNA in the absence of platination.[27] The putative metallation site is underscored.

Top strand - 5'-CTCTCGGTCTC-3'

Bottom strand – 3'-GAGAGCCAGAG-5'

After incubation of the ds DNA test sequence with complexes **1**-**6** and complete digestion of the oligonucleotides, HPLC analyses were carried out. Nucleoside peaks were identified by the coinjection of standards and integrated peak area ratios were calculated. The ratios of nucleosides were calculated and analyzed as it independent of volume and concentration of reaction, and is not influenced by volume of sample used for HPLC analysis and eliminates errors associated with pipetting technique. Binding of cisplatin analogues to a dsDNA sequence is consistent with a marked decrease in dG/dC and a slight decrease in dA/dC because dG and dA are known to coordinate metal complexes. This drastically shifts their retention times in RP-HPLC.[28] The nucleoside peak area

ratios for the control digests conformed to the stoichiometry of the nucleosides within the DNA strands given their relative absorptivities.[29] The dG/dA ratio can be used as a measure of binding selectivity for dG over dA. Results are shown in Figure 4 for 1 eq. of complex per duplex DNA. From the ratios of dG/dC it can be seen that the pyridyloxazoline complexes tested are less reactive towards dG than cisplatin. From comparison with the control experiment, the composition of nucleosides is reduced in dG by *ca.* 40% upon exposure to cisplatin. Higher degrees of platination may be obtained by reaction at elevated temperatures with an excess of platinum salt or through the use of *bis*-DMF/NO_3^- compounds.[27,28] The low thermal stability of our test sequence precluded us from the use of the former and we avoided the latter as we wished not for complete saturation, but to determine the differences in binding between enantiomeric complexes. In all digests, dC/dT remains approximately constant implying that these bases are not reactive towards the platinum complexes, as expected. Enantiomeric complexes showed no discernable patterns in their DNA binding ability. In the case where R = iPr (*R*-**1**, *S*-**1**) the *S* enantiomer appears less reactive towards dG, but in the case where R = Ph (*R*-**2**, *S*-**2**) it is the *R* enantiomer which is less reactive. In the case of R = Bn (*R*-**3**, *S*-**3**), there is little difference due to stereochemistry (Fig. 2-5). The unsubstituted ligand complex **6** (R = H) appeared to be the least reactive of the platinum complexes; this was somewhat surprising as it was thought the decreased steric constraints would allow for greater DNA binding. It has been shown formerly that this compound has antineoplastic effects on cancer cells *in vitro*.[19] The lack of binding in our study implies that this compound may have a different biological target and/or mode of action. Both palladium complexes tested showed no significant deviations from the negative control; palladium complexes in a *cis* geometry are known to have lower binding affinities towards DNA than *cis*-dichloroplatinum compounds; however, the comparable *trans*-complexes often offer cytostatic effects surpassing that of cisplatin.[30]

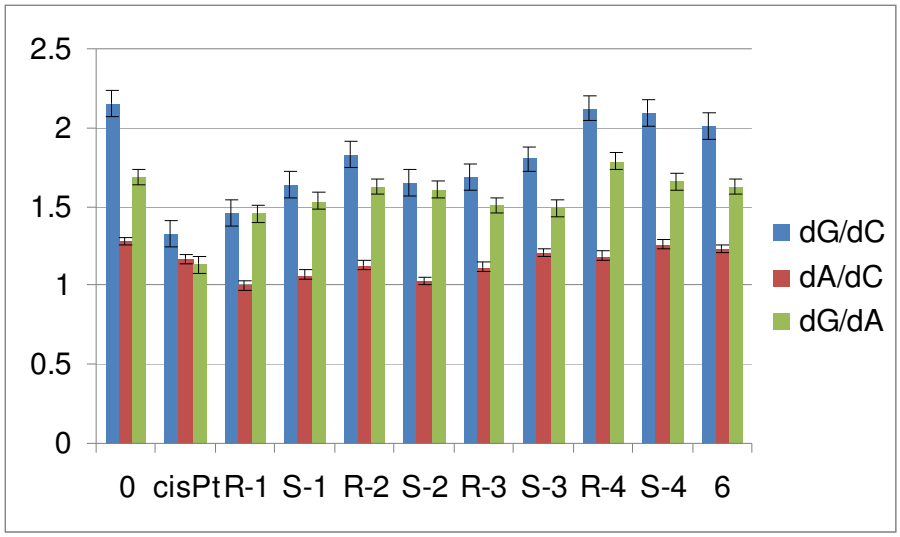


Figure 2-5: Ratios of peak area for enzymatic digests of the undecamer duplex after exposure to no Pt containing compound (0), cisplatin (cisPt), and compounds 1-6. Error bars are +/- standard deviation.

In the single stranded digests, lack of secondary structure should make the stereochemistry of the complexes less relevant. The complexes exhibited reactivity different from what was observed in the case of the duplex. In the reaction with the bottom strand (5'-GAGACCGAGAG-3'), where the only adjacent purines are dApdG, *R*-**1**, *S*-**1** and *S*-**2** showed reactivity towards dG comparable to that of cisplatin (Fig. 2-6).

All complexes showed greater reactivity towards dA than cisplatin. Decreased selectivity for dG is possibly due to the lack of an available NH proton. Digests of the top strand (5'-CTCTCGGTCTC-3') yielded very similar results to the duplex digestions where all compounds bound to dG less efficiently than cisplatin.

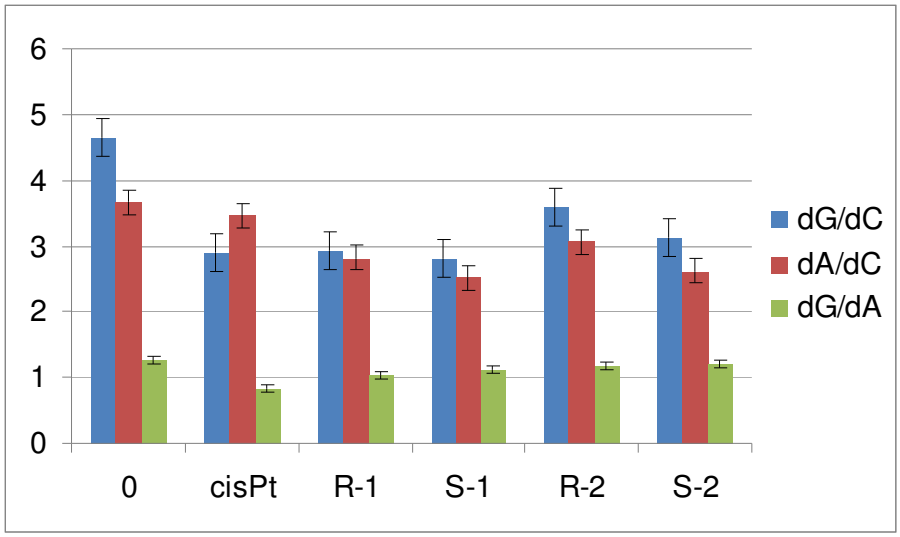


Figure 2-6: Ratio of peak areas for the enzymatic digests of the bottom strand undecamer after exposure to no Pt containing compound (0), cisplatin (cisPt), and compounds 1 and 2. Error bars are +/- standard deviation.

Thermal denaturation of platinated duplexes showed that cisplatin lowered the melting temperature (T_m) of the duplex, which is consistent with the formation of 1,2-intrastrand adducts (the stability of the duplex is compromised to the extent that a ΔT_m of -20 °C can be observed),[2,31] while the pyridyloxazoline complexes decreased the T_m only very slightly (ΔT_m = 1-3°C). This indicates that 1,2 adducts are likely not formed.

Denaturing polyacrylamide gel electrophoresis (PAGE) of the platinated ds and ss DNA again confirmed the generally low reactivity of the compounds. However, DNA platination is revealed by electrophoretic analysis of the reaction mixtures employing the most reactive compound used in this study (*R*-**1**), (Fig. 2-7). The accumulation of positive charge which occurs after platination of the sequence results in lower mobility of the oligonucleotide within the gel matrix.

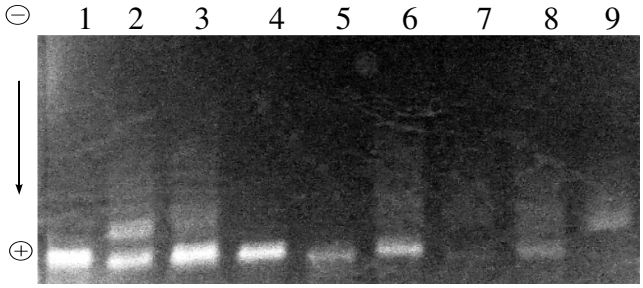


Figure 2-7: Denaturing polyacrylamide gel electrophoresis (PAGE) (negative exposure) of top strand (TS), bottom strand (BS) and double stranded (ds) DNA subjected to 1 eq. of either *R*-1, cisplatin or no Pt. Lanes: 1: dsDNA, negative control; 2: dsDNA + cisplatin; 3: dsDNA + *R*-1; 4: BS, negative control; 5: TS, negative control; 6: BS + *R*-1; 7: TS + *R*-1; 8: BS + cisplatin; 9: TS + cisplatin.

A significant shifting of the bands in the lanes containing cisplatin was observed. The lanes containing the pyridyloxazoline complex *R*-**1** either showed one band identical to the control (no platinum) or two bands, the top band (corresponding to a platinated ss sequence) appearing very faintly.

It is possible that if left for longer periods of time the compounds tested may show differences in binding between enantiomers as the diastereomeric adducts may be resolved towards the thermodynamic product. It was however found that the ratios determined at 36 and 48 h were not substantially different.

2.3 Conclusions

In conclusion, the pyridyloxazoline complexes of Pt and Pd tested in this study likely did not form 1,2-adducts with DNA, but had a higher propensity towards the formation of complexes with dA than cisplatin. It was also found that the nature and absolute configuration of the substituent at *C*4 of the oxazoline ring had no systematic or simple and predictable effect on reactivity. The previous work demonstrating the antineoplastic effects of **6** [19] may imply that monofunctional adducts formed are pharmacologically relevant. *In vivo* or *in vitro* biological assays may, in turn, highlight whether the chirality of these complexes is pertinent in regards to their efficacy as anticancer agents.

2.4 Experimental

General

Pyridyloxazoline ligands and the metal precursors $PdCl_2(PhCN)_2$ and cis/trans-$PtCl_2(SMe_2)_2$ were synthesised according to literature procedures and characterization data matched those of the known compounds.[20,21] Platinum compounds **1**-**3**, **5**, **6** were synthesized by a general procedure as outlined for *R*-**1**; spectral data for enantiomeric complexes were identical, as expected. Solution ^1H NMR spectra were collected using a Varian 400 spectrometer (400.09 MHz for ^1H and 100.61 MHz for ^{13}C) at r.t. in $CDCl_3$ (unless otherwise noted) and referenced to residual ^1H or ^{13}C in the deuterated solvent. Coupling constants are given in Hz. Elemental analyses were conducted by Chemisar Laboratories Inc., Guelph ON. Denaturing polyacrylamide gels were cast and run according to literature procedures.[22] The resultant gels were documented by colour digital photography over fluorescent TLC plates illuminated at 258 nm. For clarity, the image presented has been rendered in grayscale, inverted and contrast/brightness adjusted in Adobe Photoshop.

Optical rotations were recorded on an Atago Polax-2L polarimeter at a concentration of 1g/100 mL at 589 nm in CH_2Cl_2 solution.

Stock solutions of the complexes (40 μM) were prepared by dissolution in a minimum amount of DMF (< 0.5 mL) followed by dilution with Milli-Q deionized water to 100 mL and stored at r.t. in the dark for no longer than 2 weeks. DNA was purchased from the University of Calgary DNA Services and used without further purification.

Syntheses

Dichloro-2-[(4R)-4-iso-propyl-4,5-dihydro-1,3-oxazol-2-yl]pyridineplatinum(II) (R-1)

The title complex was synthesized by the addition of cis/trans-$PtCl_2(SMe_2)_2$ (0.23 g, 0.60 mmol) to 5 mL of 1,2-dichloroethane containing a slight excess of pyridyloxazoline ligand (0.14 g, 0.72 mmol). The resultant yellow solution was heated to reflux for 4 h over which time it darkened in colour. The solvent was removed in vacuo and the

remaining oil triturated with Et_2O (10 mL). Vacuum filtration gave 0.27 g (94%) of *R*-**1** as an orange/yellow powder. Anal. calcd. for $C_{11}H_{14}N_2Cl_2OPt$: C, 28.96; H, 3.09; N, 6.14%. Found: C, 28.90; H, 3.10; N, 6.21%. ^{1}H NMR δ: 9.51 (m, 1H, py), 8.17 (m, 1H, py), 7.76 (m, 1H, py), 7.63 (m, 1H, py), 5.00 (m, 1H, CH_2), 4.86 (m, 1H, CH), 4.71 (m, 1H, CH_2), 2.99 (m, 1H, CH), 0.98 (d, 3H, CH_3, ^{3}J=7.1), 0.83 (d, 3H, CH_3, ^{3}J=6.9). ^{13}C{^{1}H} NMR δ: 173.2, 149.7, 145.0, 139.4, 129.4, 126.6, 73.3, 68.2, 28.9, 18.9, 14.1. HRMS (ESI) calcd. for (M + Na) $C_{11}H_{14}N_2Cl_2ONaPt$ 477.0008; found: 477.0015. [α] = -65.0°.

Dichloro-2-[(4S)-4-iso-propyl-4,5-dihydro-1,3-oxazol-2-yl]pyridineplatinum(II) (S-1)

Yield: 94%. (ESI) calcd. for (M + Na) $C_{11}H_{14}N_2Cl_2ONaPt$ 477.0008; found: 476.9996. [α] = 70.0°. Orange plates suitable for X-ray diffraction analysis were provided by slow diffusion of Et_2O into a CH_2Cl_2 solution of the title complex.

Dichloro-2-[(4R)-4-benzyl-4,5-dihydro-1,3-oxazol-2-yl]pyridineplatinum(II) (R-2)

Yield: 72%. Anal. calcd. for $C_{15}H_{14}N_2Cl_2OPt$: C, 35.73; H, 2.80; N, 5.56%. Found: C, 36.02; H, 2.76; N, 5.54%.^{1}H NMR δ: 9.63 (m, 1H, py), 8.15 (m, 1H, py), 7.67 (m, 1H, py), 7.39 (m, 1H, py), 7.27 (m, 5H, Ar), 5.01 (m, 1H, CH_2), 4.93 (m, 1H, CH), 4.85 (m, 1H, CH_2), 3.66 (dd, 1H, CH_2, ^{2}J=13.9, ^{3}J=10.7), 3.11 (dd, 1H, CH_2, ^{2}J=13.9, ^{3}J = 5.8). ^{13}C{^{1}H} NMR δ: 173.9, 150.1, 145.3, 139.7, 135.9, 130.2, 129.5, 129.2, 127.6, 126.8, 64.5, 39.8. HRMS (ESI) calcd. for (M + Na) $C_{15}H_{14}N_2Cl_2ONaPt$ 525.0008; found: 525.0033.

Dichloro-2-[(4S)-4-benzyl-4,5-dihydro-1,3-oxazol-2-yl]pyridineplatinum(II) (S-2)

Yield: 68%. Anal. calcd. for $C_{15}H_{14}N_2Cl_2OPt$: C, 35.73; H, 2.80; N, 5.56%. Found: C, 36.03; H, 3.06; N, 5.52%. HRMS (ESI) calcd. for (M + Na) $C_{15}H_{14}N_2Cl_2ONaPt$ 525.0008; found: 525.0033.

Dichloro-2-[(4R)-4-phenyl-4,5-dihydro-1,3-oxazol-2-yl]pyridineplatinum(II) (R-3)

Yield: 94%. Anal. calcd. for $C_{14}H_{12}N_2Cl_2OPt$: C, 34.30; H, 2.47; N, 5.71%. Found: C,

34.53; H, 2.72; N, 5.52%. ^{1}H NMR δ: 9.45 (m, 1H, py), 8.20 (m, 1H, py), 7.86 (m, 1H, py), 7.62 (m, 1H, py), 7.38 (m, 5H, Ph), 5.88 (dd, 1H, CH_2, ^{2}J=10.2, ^{3}J=4.9), 5.46 (dd, 1H, CH, ^{3}J=4.9, ^{3}J=3.8), 4.92 (dd, 1H, CH_2, ^{2}J=10.2, ^{3}J=3.8). ^{13}C{^{1}H} NMR (CD_2Cl_2): δ: 174.3, 150.0, 149.6, 144.8, 139.5, 139.1, 130.3, 129.9, 129.1, 127.1, 126.8, 126.3, 81.1, 66.4. HRMS (ESI) calcd. for (M + Na) $C_{14}H_{12}N_2Cl_2ONaPt$ 510.9851; found: 510.9845.

Dichloro-2-[(4S)-4-phenyl-4,5-dihydro-1,3-oxazol-2-yl]pyridineplatinum(II) (S-3)

Yield: 92%. Anal. calcd. for $C_{14}H_{12}N_2Cl_2OPt$: C, 34.30; H, 2.47; N, 5.71%. Found: C, 34.56; H, 2.75; N, 5.73%. HRMS (ESI) calcd. for (M + Na) $C_{14}H_{12}N_2ONaCl_2Pt$ 510.9851; found: 510.9868.

Dichloro-2-[(4R)-4-iso-propyl-4,5-dihydro-1,3-oxazol-2-yl]pyridinepalladium(II) (R-4)

The title complex was synthesised as described in the literature. Characterization data matched those of the known compound.[17]

Dichloro-2-[(4S)-4-iso-propyl-4,5-dihydro-1,3-oxazol-2-yl]pyridinepalladium(II) (S-4)

The title complex was made in the same way as *R*-**4**. Spectroscopic data matched those of the enantiomeric compound.

Dichloro-2-[(4S)-4-sec-butyl-4,5-dihydro-1,3-oxazol-2-yl]pyridineplatinum(II) (5)

Yield: 84%. ^{1}HNMR δ: 9.26 (d, 1H, py, ^{3}J=5.5), 7.8 (dd, 1H, py, ^{3}J=5.5, ^{3}J=7.8), 7.76 (d, 1H, py, ^{3}J=7.6), 7.56 (dd, 1H, py, ^{3}J=7.8, ^{3}J=7.6), 5.13 (m, 1H, CH_2), 4.87 (m, 1H, CH), 4.78 (m, 1H, CH_2), 2.75 (m, 1H, CH), 1.29 (m, 2H, CH_2), 1.00 (t, 3H, CH_3, ^{3}J=7.4), 0.78 (d, 3H, CH_3, ^{3}J=6.8). ^{13}C{^{1}H} NMR δ: 173.2, 149.8, 145.0, 139.3, 129.3, 126.5, 73.3, 66.9, 35.5, 26.5, 18.0, 11.9. HRMS (ESI) calcd. for (M + Na) $C_{12}H_{16}N_2ONaCl_2Pt$ 491.0164; found: 491.0177.

Dichloro-2-[4,5-dihydro-1,3-oxazol-2-yl]pyridineplatinum(II) (6)

The title compound has been previously made by a different route.[19] Following the procedure for *R*-**1**, 6 was obtained in 72% yield. The limited solubility of this complex

did not allow for the measurement of NMR spectral data. Anal. calcd. for $C_8H_8N_2Cl_2OPt$: C, 23.20; H, 1.95; N, 6.76%. Found: C, 23.50; H, 1.89; N, 6.66%.

Crystallography

An orange plate was cut loose and then mounted on a glass fibre. Data were collected at low temperature (-123 °C) on a Nonius Kappa-CCD area detector diffractometer with COLLECT (Nonius B.V., 1997-2002). The unit cell parameters were calculated and refined from the full data set. Crystal cell refinement and data reduction were carried out using HKL2000 DENZO-SMN.[23] The absorption correction was applied using HKL2000 DENZO-SMN (SCALEPACK). The crystal data and refinement parameters of S-1 are listed in table 2-1.

Table 2-1-1: X-ray crystallographic structural parameters for compound S-1.

Crystal	*S*-**1**
Empirical formula	$C_{11}H_{14}Cl_2N_2Opt$
Formula mass	456.23
Colour, habit	Orange, plates
Crystal dimensions (mm)	0.17 x 0.15 x 0.10^3
Crystal system	Monoclinic
Space group	$P\ 2_1$
Z	4
a (Å)	6.7455(3)
b (Å)	20.2782(9)
c (Å)	9.5172(4)
α (°)	90
β (°)	93.10(2)
γ (°)	90
Collection ranges	-4<=*h*<=4; -24<=*k*<=24; -11<=*l*<=11
Temperature (K)	150(2)
Volume ($Å^3$)	1299.93(10)
D_{calcd} (Mg m^{-3})	2.331
Radiation	Mo Kα (λ = 0.71073 Å)
Absorption coeff. (μ) (mm^{-1})	11.189
Absorption correction	Semi-empirical from equivalents
F(000)	856
θ range for data collection (°)	2.94 to 25.03
Observed reflections	11811
Independent reflections	2550 (R_{int} = 0.0760)
Data/restraints/parameters	2550/37/171
Goodness-of-fit on F^2	1.023

Final R indices ($I > 2\sigma(I)$)	$R_1 = 0.0321$, $wR_2 = 0.0664$
R indices (all data)	$R_1 = 0.0395$, $wR_2 = 0.0693$
Absolute structure parameter	-0.009(18)
Largest diff. peak and hole (e Å^3)	0.889 and -0.702

The reflection data and systematic absences were consistent with a monoclinic space group: P21). The SHELXTL/PC V6.14 for Windows NT (Sheldrick, G.M., 2001) suite of programs was used to solve the structure by direct methods. Subsequent difference Fourier syntheses allowed the remaining atoms to be located. There were two independent molecules in the asymmetric unit. For this crystal, twinning was also an issue; it is discussed below. Only the heavy atoms (Pt, Cl) were refined anisotropically as some of the lighter atoms became non-positive definite when they were allowed to refine freely. The hydrogen atom positions were calculated geometrically and were included as riding on their respective carbon atoms. The largest residue electron density peak (0.877 e/Å3) was associated with one of platinum atoms. Full-matrix least squares refinement on F^2 gave $R_1 = 3.21$ for 2σ data and $wR_2 = 6.66$ for all data (GooF = 1.021).

The structure was solved readily, but the initial refinement stalled at R = 12.9 %. Further, the light atoms did not refine successfully with an anisotropic model. There were signs of non-merohedral twinning in the E-statistics and the Fobs values were consistently higher than the Fcalcs. WinGX [24] was used to "detwin" the data. ROTAX [25] found the Twin Law, which turned out to be a 180° rotation about the direct lattice direction (1 0 0). "Make HKLF5" was used to generate the detwinned file used in further refinement.

The structure refinement improved considerably with better values for R_1, wR_2, GooF, standard uncertainties and residual electron density. The final R_1 value came down to 3.21%. The final solution, however, could still not be refined completely anisotropically, and the light atoms were left isotropic in the refinement. The constant for the Twin Law refined to a value of 0.47309. All these points served to confirm the correct Twin Law was chosen.

Nuclease Digest Studies

Duplex DNA was prepared by heating a 1:1 mixture of the complementary sequences in

10 mM $NaClO_4$ to 90 ºC and allowing the solution to come slowly to rt. The solution was then cooled to 4ºC for 2h. Both single and double stranded DNA (2 μM) were reacted with the complexes in a 1:1 molar ratio in 200 μL of 10 mM $NaClO_4$ for 48 h in the dark. The buffer was then adjusted to 50 mM Tris, 20 mM $MgCl_2$, pH 7.5. Bovine serum albumin (BSA) was added to a final concentration of 0.02 μg/mL and the DNA was digested sequentially with bovine pancreas DNase I (Sigma, 36 U for 4 h), nuclease P1 (Sigma, 10 U for 18 h) and alkaline phosphatase (Promega, 20 U for 4 h). The samples were then heated to 90 ˚C for 2 min to denature the enzymes and centrifuged for 10 min at 10,000 g. The supernatant was analyzed at 254 nm by RP-HPLC on a Varian Microsorb 100 Å C18 column using a flow rate of 1 mL min^{-1} and a solvent system of 100:0 0.1 M aqueous $[Et_3NH][OAc]$ (pH 6.5):MeCN from 0-5 min followed by a linear ramp to 90:10 over a period of 20 min. Nucleoside peaks were identified by co-injection of standards. The areas under the peaks were found by integration and analyzed in relation to one another. These digests and the HPLC analyses thereof were carried out in triplicate.

2.5 References

1. Rosenberg, B. *Interdiscip. Sci. Rev.* **1978**, *3*, 134.

2. Wong, E.; Giandomenico, C.M. *Chem. Rev.* **1999**, *99*, 2451.

3. Poklar, N.; Pilch, D.S.; Lippard, S.J.; Redding, E.A.; Dunham, S.U.; Breslauer, K.J. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 7606.

4. Jakupec, M.A.; Galanski, M.; Arion, V.B.; Hartinger, C.G.; Keppler, B.K. *Dalton Trans.* **2008**, 183.

5. Jung, Y.; Lippard, S.J. *Chem. Rev.* **2007**, *107*, 1387.

6. Cleare, M.J.; Hoeschele, J.D. *Bioinorg. Chem.* **1973**, *2*, 7792.

7. Benedetti, M.; Marzilli, L.G.; Natile, G. *Chem. Eur. J.* **2005**, *11*, 5302.

8. Weiss, R.B.; Christian, M.C. *Drugs* **1993**, *46*, 360.

9. Delalande, O.; Malina, J.; Brabec, V.; Kozelka, J. *Biophys. J.* **2005**, *88*, 4159.

10. Hambley, T.W.; Ling, E.C.H.; Munk, V.P.; Davies, M.S. *J. Biol. Inorg. Chem*. **2001**, *6*, 534.

11. Mier-Vinué, J.d.; Gay, M.; Montaña, A.M.; Sáez, R.; Moreno, V.; Kasparkova, J.; Vrana, O.; Heringova, P.; Brabek, V.; Boccarelli, A.; Coluccia, M.; and Natile, G. *J. Med. Chem*. **2008**, *51*, 424.

12. Natile, G.; Marzilli, L.G. *Coord. Chem. Rev*. **2006**, *250*, 1315.

13. Reedijk, J. *Chem. Commun*. **1996**, 801.

14. Bloemink, M.J.; Engelking, H.; Gütschow, N.; Karentzopoulos, S.; Krebs, B.; Reedijk, J. *J. Inorg. Biochem*. **1995**, *59*, 222.

15. Bloemink, M.J.; Engelking, H.; Karentzopoulos, S.; Krebs, B.; Reedijk, J. *Inorg. Chem.* **1996**, *35*, 619.

16. Dodd, D.W.; Toews, H.E.; Carneiro, F.d.S.; Jennings, M.C.; Jones, N.D. *Inorg. Chim. Acta*. **2006**, *359*, 2850.

17. Bolm, C.; Weickhardt, K.; Zehnder, M.; Ranff, T. *Chem. Ber*. **1991**, *124*, 1173.

18. Muñoz, M.P.; Adrio, J.; Carretero, J.C.; Echavarren, A.M. *Organometallics*, **2005**, *24*, 1293.

19. Paschke, N.; Röndigs, A.; Poppenborg, H.; Wolff, J.E.A.; Krebs, B. *Inorg. Chim. Acta*. **1997**, *264*, 239.

20. Hartley, F.R. *Organomet. Rev. A*. **1976**, *6*, 119.

21. Hill, G.S.; Irwin, M.J.; Levy, C.J.; Puddephatt, R.J. *Inorg. Synth*. **1998**, *32*, 149.

22. Ellington, A.; Pollard, J.D. *Curr. Prot. Mol. Bio.* **1998**, 2.12.1-2.12.7.

23. Otwinowski, Z.; Minor, W. *Methods in Enzymology: Macromolecular Crystallography*

Part A. **1997**, *276*, 307.

24. Farrugia, L.J. *J. Appl. Cryst*. **1999**, *32*, 837.

25. Cooper, R.I.; Gould, R.O.; Parsons, S.; Watkin, D.J. *J. Appl. Cryst*. **2002**, *35*, 168.

26. McKennon, M.J.; Meyers, A.I.; Drauz, K.; Schwarm, M. *J. Org. Chem*. **1993**, *58*, 3568.

27. Dunham, S.U.; Turner, C.J.; Lippard, S.J. *J. Am. Chem. Soc*. **1998**, *120*, 5395.

28. Nikolcheva, L.G.; Vogels, C.M.; Stefan, R.A.; Darwish, H.A.; Duffy, S.J.; Ireland, R.J.; Decken, A.; Hudson, R.H.E.; Westcott, S.A. *Can. J. Chem*. **2003**, *81*, 269.

29. Sober, H.A.; Harte, R.A.; Sober, E.K. CRC Handbook of Biochemistry and Selected Data for Microbiology 2nd Ed. Chemical Rubber Co. Cleveland OH **1970**. Section G.

30. . Ray, S.; Mohan, R.; Singh, J.K.; Samantaray, M.K.; Shaikh, M.M.; Panda, D.; Gosh, P. *J. Am. Chem. Soc*. **2007**, *129*, 15042.

31. Van Hemelryck, B.; Guiettet, E.; Chottard, G.; Girault, J.P.; Huynh-Dinh, T; Lallemand, Y.J.; Igolen, J.; Chottard, J.C. *J. Am. Chem. Soc*. **1984**, *106*, 3037.

Chapter 3

3 Peptide Nucleic Acid Pt(II) Conjugates: A Preliminary Study of Antisense Effects in *Xenopus laevis*

The avid hybridization of peptide nucleic acid (PNA) to DNA and RNA, coupled with the analogue's stability towards enzymatic degradation, has led to its investigation as an antigene/antisense agent. PNA targeted toward the 5'-UTR of an mRNA transcript can effect efficient silencing; however, if targeted to an area within the coding region, the PNA can be displaced by the moving ribosome and be an ineffective antisense agent. Platinum-appended and standard PNAs antisense to an area within the open reading frame of the gene noggin, were injected into *Xenopus laevis* embryos. Phenotypic responses were observed and the preliminary results are reported herein.

3.1 Introduction

Antisense PNA (asPNA) exerts its effects differently from the most well established and widely used oligodeoxynucleotide analogues, such as phosphorothioates, in that it acts independently of RNaseH (please see chapter 1.3 for a more detailed discussion). Mechanistically, with respect to translation, asPNA relies on steric blockage of the mRNA transcript. This steric interference is most robust within the 5'-UTR, where asPNA prevents ribosomal binding, effective knockdown of transcription can be readily achieved. However, oftentimes when asPNAs targeting areas within the coding region are used, little or no antisense effects are observed.[1-3] When down-regulating a gene product by asPNA, the target chosen must share little homology to other sequences within the transcriptome. If selection of this target is limited to the 5'-UTR it can be difficult to find a unique sequence. If the whole of the transcriptome was available for this purpose then more sequences could be assessed. Since the discovery of RNAi, post-transcriptional down-regulation of a gene by targeting the coding region is routinely possible although it is often difficult to predict the efficiency/specificity. The use of PNA also offers clinical advantages relating to serum stability, lower molecular weight and fewer off-target and immunogenic effects.[4]

The drug *cis*-diamminedichloroplatinum (II) (cisplatin) is a successful anti-cancer agent, particularly in the treatment of testicular cancer where the disease can, in the vast majority of cases, be cured.[5] The mechanism of action is almost universally thought to be brought about by the coordination of adjacent purines (preferentially G) to the metal centre in an *N*7-Pt-*N*7 bonding arrangement.[6] Platinum cross-links have been investigated using transplatin modified PNA,[7] and 2'-*O*-methylribonucleotides whereby the metal centres are linked to the oligonucleotide analogue in a 1,3-intrastrand *N*7-Pt-*N*7 bonding arrangement.[8,9] A Pt chelating aminoethylglycine moiety has previously been incorporated into a solid phase synthesis strategy with the aim of producing a library of peptide based cisplatin analogues.[10,11] Our aim is to incorporate this same platinum chelating moiety into a PNA synthesis thereby combining the sequence specificity of PNA and the irreversible, covalent platination of DNA and mRNA.

Xenopus laevis is an African frog which finds much use in developmental biology and cell signaling studies. Eggs of this species are easily obtained, fertilized and cultivated, and embryos are large enough to be injected with up to 20 nL of an aqueous solution.[12] In choosing a gene to knockdown, several criteria must be met: the gene must be expressed early in development, show a clear phenotypic response upon knockdown (with literature precedent), and it would also be desirable that an antibody is commercially available to allow for Western analysis of protein levels. Noggin, a small signaling protein involved in the cephalization cascade,[13,14] fulfilled all of these criteria. The protein is thus named due to the loss-of-function mutants displaying a headless phenotype. We therefore aimed to inject Pt-containing and standard PNA complementary to an area within the open reading frame of *Xenopus* noggin with the hypothesis that an increased antisense effect would be brought about by the formation of Pt mediated cross-links.

3.2 Results and Discussion

Fmoc protection of β-alanine and aminoethylglycine proceeded smoothly and monomers **1**[15] and **2**[10] were incorporated into the solid phase supported synthesis of a PNA 15-mer complementary to *X. laevis* noggin mRNA. Following removal of the Fmoc groups, the oligomers were subjected to platinating conditions[10] while on the solid support with base-protective groups still present. As a control, a portion of the PNA was not platinated.

After TFA cleavage from the resin and purification by RP-HPLC, the oligomers were characterized by ESI-HRMS and injected into *X. laevis* embryos as unbuffered aqueous solutions, figure 3-1. Once hybridized, the nearby GpG of the target may coordinate the pendant metal centre. The target region is within the open reading frame of noggin, 376 bases downstream of the AUG start codon in an area predicted to have limited secondary structure.

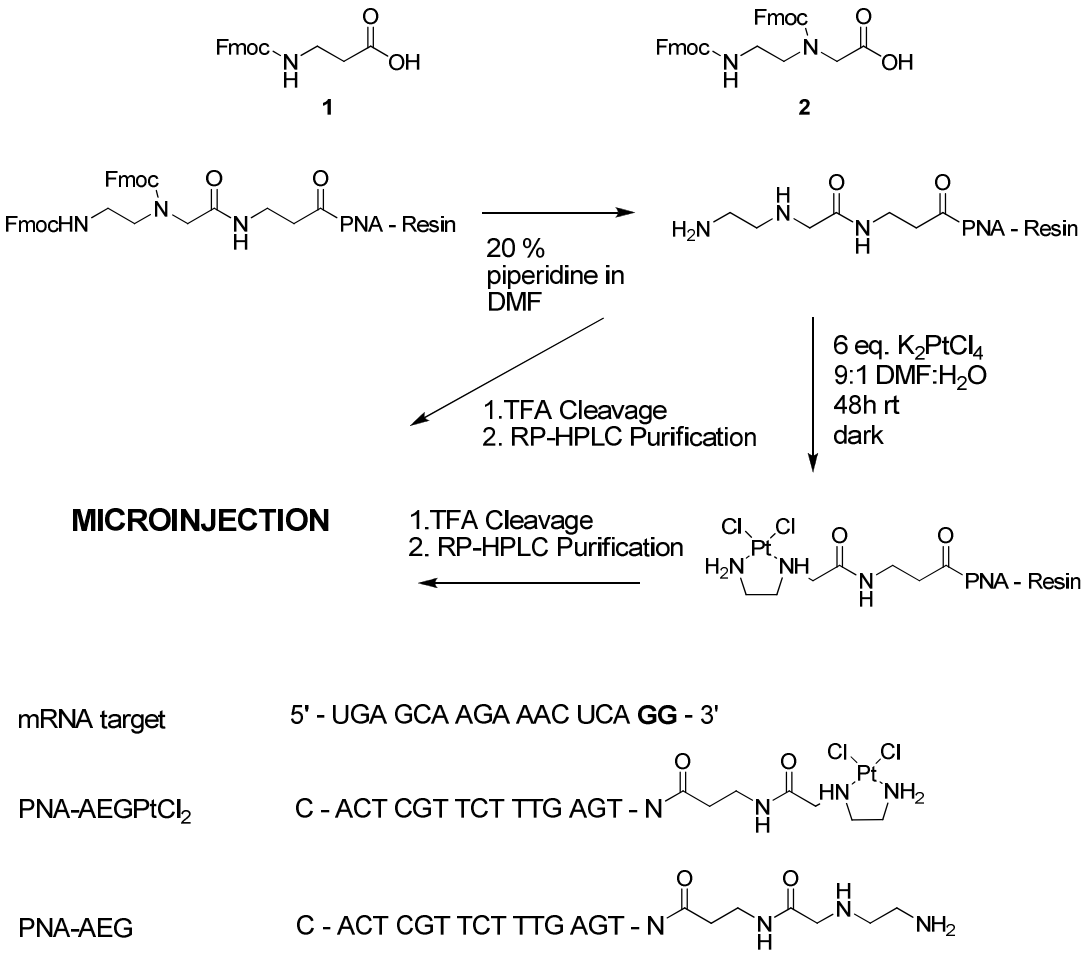


Figure 3-1: Solid-supported synthesis of a platinum-bearing PNA. Sequence identity of the target with the putative platination site shown in bold.

In vitro characterization of binding was carried out with a DNA target for simplicity, cost and, ease of handling. Unplatinated PNA-aminoethylglycine (PNA-AEG) demonstrated an expectedly high melting temperature of 83˚C with its DNA complement as determined

by temperature dependent UV-Vis spectroscopy, however, no cooperative transition could be observed for PNA aminoethylglycine$PtCl_2$ (PNA-AEG$PtCl_2$) with its complement up to 95˚C. These data are consistent with the presence of Pt-mediated cross-linking to complementary DNA.

Microinjection of *X. laevis* embryos was carried out with PNA-AEG and PNA-AEG$PtCl_2$ in three different concentrations of each along with an injection of cisplatin as a control. Several rounds of injections were carried out; and when the embryos reached stage 40 the then tadpoles were examined for any phenotypic response (Figure 3-2).

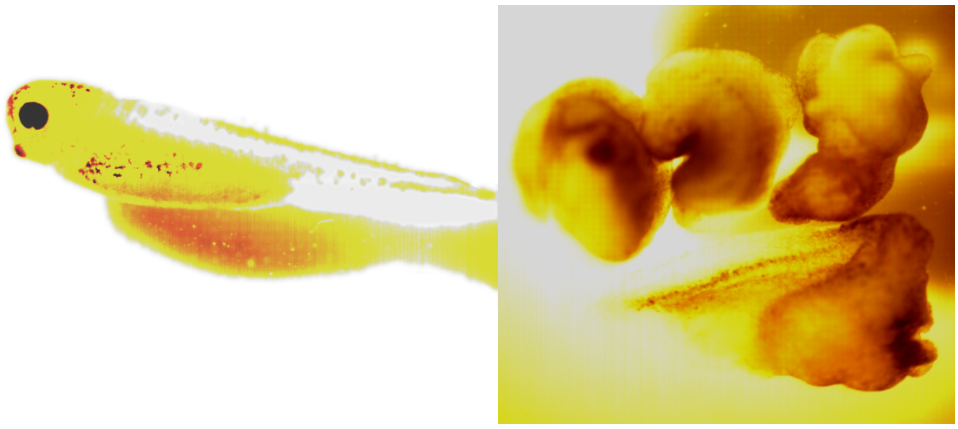


Figure 3-2: Microphotographs of nomal (left) otherwise abnormal (top right) and headless (bottom right) phenotypes.

Data for phenotypic response, shown in table 3-1, indicate that there are no significant phenotypic differences between tadpoles which encountered platinated versus unplatinated oligomer in varying concentrations. High mortality rates can be attributed to the quality of the embryos and also the prominence of unfertilized eggs (62% mortality for uninjected embryos). To avoid this problem in the future, embryos will be injected after the first cleavage (cellular division) so that all embryos injected are known to be fertilized and likely viable. It can be seen that the injection of cisplatin resulted in 100% mortality while other injections resulted in mortalities ranging from 80-86%. It is notable that the PNA-AEG and PNA-AEG$PtCl_2$ treated embryos demonstrated the headless

phenotype, wherein it was completely absent from the control sample, figure 3-2. Additionally, the abnormal phenotype was more abundant in the treated samples compared to the control, and the amount of wild type body morphology was dramatically knocked down.

Table 3-1: Phenotypic analysis of X. laevis embryos at stage 40 after microinjection with PNA-AEG, PNA-AEGPtCl$_2$ and cisplatin along with an uninjected control (wt = wild type). Numbers are reported with the percent of total embryos given in brackets. Total compound injected is denoted by low (1.15×10^{-13} mol), med (5.81×10^{-13} mol) and high (11.62×10^{-13} mol).

asPNA	Total	Dead	Headless	Otherwise abnormal	wt
PNA-AEG low	240	206(86.2)	3(1.3)	15(6.3)	16(6.2)
PNA-AEG med	217	182(83.9)	5(2.3)	12(5.5)	18(8.3)
PNA-AEG high	260	216(83.1)	5(1.9)	18(6.9)	21(8.1)
PNA-AEGPtCl$_2$ low	249	210(84.3)	8(3.2)	15(6.0)	16(6.4)
PNA-AEGPtCl$_2$ med	232	189(81.4)	3(1.3)	28(12.1)	12(5.2)
PNA-AEGPtCl$_2$ high	268	215(80.2)	12(4.5)	26(9.7)	15(5.6)
Cisplatin med	105	105(100.0)	-	-	-
Control (no injection)	362	223(61.6)	0	8(2.2)	131.0(36.2)

There appears to be no pattern in the phenotypes or the mortality between the injection of different PNAs in different amounts. Statistically, the platinated PNA showed no more potent an antisense effect. There may be a slight antisense effect through steric blocking of the mRNA without platination, in which case, a shorter PNA would be desired for further studies. The appearance of wild type individuals can, in part, be explained by failed injection as opposed to authentic lack of antisense action. It was also suspected that our platinated oligomer was contaminated with unplatinated oligomer. PNA-AEG was always observed to a small extent in mass spectra although this could simply have been an artefact of the technique. Regardless, we were driven to find an alternate means of

purification as existing methods were insufficient (this will be described in chapter 4).

3.3 Conclusions

These studies, although encouraging, are only preliminary and further characterization of the system is needed to be carried out *in vitro*. RNA electrophoretic mobility shift assays under both native and denaturing conditions could be used to determine if crosslinks have formed to the Pt-PNA conjugate. Platinated adducts would be further characterized by mass spectrometry. Western analysis of protein levels has yet to be carried out. Microinjection of asPNAs avoids the issue of cell penetration/compartmentalization found in cell culture studies; therefore, we would aim to utilize a eukaryotic cell-free translation system to monitor protein production *in vitro*. The ultimate goal of this project is to provide molecular tools – broadly useful, tunable antisense agents, for the study of developmental biology in model organisms such as *X. laevis* and *D. rerio*.

3.4 Experimental

Peptide synthesis

PNAs were synthesized on a Rink amide resin using an Applied Biosystems 433A peptide synthesizer and commercially available monomers (Applied Biosystems, Warrington, UK.) using standard Fmoc chemistry. Compounds **1** [15] and **2** [10] were prepared according to literature procedures. PNA conjugates were purified by RP-HPLC and characterized by ESI-HRMS. PNA-AEGPtCl_2 HRMS(ESI) Calcd. for $C_{175}H_{230}Cl_2N_{84}O_{53}Pt$ 4621.6943 Found 4621.4878. PNA-AEG HRMS(ESI) Calcd. for $C_{175}H_{230}N_{84}O_{53}$ 4356.7918 Found 4356.5552.

Thermal denaturation

Complementary DNA of the sequence 5'-GAC-TGA-GCA-AGA-AAC-TCA-GGA-G-3' was purchased from The University Core DNA Services at The University of Calgary and used without further purification. UV-Vis dependent thermal denaturation was carried out in aqueous 100 mM NaCl, 10 mM Na_2HPO_4, 0.1 mM EDTA, pH 7.0 at a strand concentration of 8μM. Denaturation was performed between 10 °C to 95 °C at a

rate of 0.5 °Cmin^{-1}. Melting temperature is given as an average of three runs within 1˚C and estimated for cooperative transitions by the first derivative method.

Microinjections

Fertilization and microinjection of *Xenopus laevis* embryos were carried out according to a published procedure.[12] Embryos in 1X Marc's modified ringer (MMR) were injected at the one cell stage with 8 nL of solution. Embryos were injected with either 1.15×10^{-13}, 5.81×10^{-13} or 11.62×10^{-13} mol of PNA-AEG, PNA-AEG$PtCl_2$ or cisplatin. These amounts correspond to a similar study in the literature.[14]After injection, embryos were cultivated in 0.1X MMR until being assayed for phenotypic changes at stage 40.

3.5 References

1. Doyle, D.F.; Braasch, D.A.; Simmons, C.S.; Janowski, B.A.; Corey, D.R. *Biochemistry* **2001**, *40*, 53-64.

2. Dryselius, R.; Aswati, S.K.; Rajaro, G.K.; Nielsen, P.E.; Good, L. *Oligonucleotides* **2003**, *13*, 427-433.

3. Knudsen, H.; Nielsen, P.E. *Nucleic Acids Res*. **1996**, *24*, 494-500.

4. Corey, D.R. *Nat. Chem. Biol.* **1997**, 3, 8-11.

5. Jung, Y.; Lippard, S.J. *Chem. Rev.* **2007**, 105, 1387-1407.

6. Kartalou, M.; Essigmann, J.M. *Mutat. Res.* **2001**, 478, 23-43.

7. Schmidt, K.S.; Boudvillain, M.; Schwartz, A.; van der Marel, G.A.; van Boom, J.H.; Reedijk, J.; Lippert, B. *Chem. Eur. J.* **2002**, *8*, 5566-5570.

8. Boudvillan, M.; Guérin M.; Dalbiès, R.; Saison-Behmoaras, T.; Leng, M. *Biochemistry* **1997**, 36, 2925-2931.

9. Aupeix-Scheidler, K.; Chabas, S., Bidou, L.; Rousset, J.-P.; Leng, M., Toulmé, J.-J. *Nucleic Acids Res*. **2000**, 28, 438-445.

10. Robillard, M.S.; Valentijn, A.R.P.M.; Meeuwenoord, N.J.; van der Marel, G.A.; van Boom, J.H.; Jan Reedijk. *Angew. Chem. Int. Ed.* **2000**, 39, 3096-3099.

11. Robillard, M.S.; van Alphen, S.; Meeuwenoord, N.J.; Jansen, B.A.J.; van der Marel, G.A.; van Boom, J.H.; Reedijk, J. *New J. Chem.* **2005**, *29*, 220-225.

12. Richard-Parpaillon, L.; Héligon, C.; Chesnel, F.; Boujard, D.; Philpott, A. *Dev. Biol.* **2002**, 244, 407-417.

13. Valenzuela D.M.; Economides A.N.; Rojas E.; Lamb T.M.; Nuñez L.; Jones P.; Lp N.Y.; Espinosa R. 3rd.; Brannan C.I.; Gilbert D.J.; Copeland N.G.; Jenkins N.A.; Le Beau M.M.; Harland R.M.; Yancopoulos G.D. *J. Neurosci.* **1995**, 15, 6077–6084.

14. Kuroda, H.; Wessely, O.; De Robertis, E.M. *PLOS Biol.* **2004**, *2*, 623-634.

15. Shi, Y.; Machida, K.; Kuzuya, A.; Komiyama M. *Bioconjugate Chem.* **2005**, *16*, 306-311.

Chapter 4

4 Denaturing Polyacrylamide Gel Electrophoresis of Peptide Nucleic Acids as a Purification and Analytical Technique

A flexible and convenient protocol for the analysis and purification of peptide nucleic acid oligomers and peptide nucleic acid-peptide chimeras by denaturing polyacrylamide gel electrophoresis is described. Vertical slab gel electrophoresis, 26% in polyacrylamide and 8 M urea at pH 3 was suitable for analysis of oligomers ranging in size from tetramers (4-mers) to tetradodecamers (24-mers). Single–base resolution of oligomers was achieved and separations are generally superior to those given by standard reversed-phase HPLC techniques. The separation of a related series of PNA oligomers showed the distance migrated was linearly dependent on the logarithm of the molecular weight. The migration of oligomers through the gel is dependent the number of basic functional groups present, such as amino groups, and the A and C content of the oligomer. Excerpts from this chapter are taken from Dodd and Hudson *Electrophoresis* **2007**, *28*, 3884-3889.

4.1 Introduction

Peptide nucleic acid (PNA) is a polyamide-based oligonucleotide analogue possessing a high affinity and sequence selectivity for complementary DNA and RNA as well as being chemically robust and resistant to enzymatic degradation (for more on PNA and its properties refer to chapter 1.3).[1,2] These properties have led to its use as a hybridization probe and investigation into its potential as an antisense or antigene agent.

The synthesis of PNA oligomers lends itself well to automation and is performed most often by one of the two major approaches: the fluorenylmethyloxycarbonyl- (Fmoc) route or *tertiary*-butyloxycarbonyl- (Boc) route using traditional solid-phase peptide chemistry.[3] After cleavage from the resin, purification of the oligomer is typically carried out by reversed-phase high pressure liquid chromatography (HPLC) and the products are subsequently characterized by mass spectrometry. As with other uncharged DNA analogues, PNA has lowered solubility at neutral pH and shows a tendency to aggregate

oftentimes making chromatographic separation more difficult than with the corresponding DNA oligomers.[3] Indeed, we have encountered problems with the purification of certain sequences. The platinum bearing PNAs mentioned in the previous chapter were particularly difficult in this regard. In lower yielding syntheses there is often still enough of the ideal sequence present to allow its study, however, it is oftentimes impossible to isolate in pure form by HPLC.

As an alternative to reversed-phase HPLC, Nielsen and Orgel have reported anion-exchange chromatography at a pH of 10 as an efficient means of separation.[4] This technique was found to be sensitive to the G and T content of the PNA oligomer as these bases become deprotonated at high pH. However, drawbacks of this system are the use of specialized/expensive equipment and also the tendency of PNAs to rearrange and depurinate under basic conditions.[5]

Polyacrylamide gel electrophoresis of PNA complexed with DNA or RNA is well described in the literature wherein the electrophoretic mobility is imparted to the complex by the natural anionic charge of the natural oligonucleotides.[6] PNA has also been polymerized into polyacrylamide gels for use in affinity electrophoresis of DNA oligomers.[7] Despite the utility and ubiquity of electrophoresis in nucleic acid chemistry, there exists no protocol for the electrophoresis of PNAs themselves. Due to the typical inclusion of a terminal lysine residue, PNA oligomers are cationic under acid conditions and it was therefore reasoned that a protocol that was suitable for denaturing electrophoresis of basic proteins may find success in the purification and analysis of PNA oligomers. Additionally, given the basicity of the nucleobases, under acidic conditions the charge of a PNA oligomer becomes dependent on the C and A content and skeletal rearrangement is less common.[5]

4.2 Results and Discussion

General

It was found that the charge state of the oligomer was the dominant factor in determining the rate of migration of the PNA however, size also plays an important role (*vide infra*).

An increase in size will generally correspond to an increase in charge state (due to the incorporation of A and C monomers). Unlike in DNA PAGE the concentration of acrylamide and the degree of cross-linking need not be adjusted to accommodate differently sized oligomers, it was found that the same system was convenient for tetramers up to tetradodecamers and presumably beyond. Time and voltage become the only experimentally determined variables necessary. The antagonism of size and charge should not be construed as a drawback of the technique concerning the purification of crude mixtures. The side products in a crude mixture will generally be of the same charge state and are therefore separable based on molecular mass. PNAs used in this study are noted below (Table 4-1).

Table 4-1: Sequence identity of PNAs used in the present work.

PNA	Sequence[a] N→C	Charge State[b]
1	H_2N-T_6-Lys-NH_2	+2
2	H_2N-Gly-GTAGATC(HMU)CT-Lys-NH_2	+6
3	Ac-Gly-GCGCAAC-Lys-NH_2	+6
4	Ac-Gly-GCGCAA(PhpC)-Lys-NH_2	+5
5	Ac-Lys-T_{12}Gly-NH_2	+1
6	Ac-Lys-T_{13}Gly-NH_2	+1
7	Ac-Lys-T_{14}Gly-NH_2	+1
8	Ac-Lys-T_{15}Gly-NH_2	+1
9	H_2N-Lys-T_{12}Gly- NH_2	+2
10	H_2N-Lys-T_{13}Gly- NH_2	+2
11	H_2N-Lys-T_{14}Gly- NH_2	+2
12	H_2N-Lys-T_{15}Gly- NH_2	+2

[a] HMU: hydroxymethyluracil; PhpC: phenylpyrrolocytosine. [b]Charge state is calculated based on protonation of *N*-terminal amino groups, lysine sidechains, and adenine and cytosine nucleobases.

Sensitivity and resolution

PNA 1 (Table 4-1) was loaded on a 1.5 mm thickness gel in successive increments of 0.025 optical density units (ODU) from 0.025 - 0.2 ODU (0.23 – 1.84 μmol). After electrophoresis the bands in the gel were visualized by UV shadowing over a fluorescent TLC plate (Fig. 4-1). The lower detection limit of this technique was qualitatively determined to be below 0.025 (ODU). Depending on the sequence up to 2 ODU can be run without excessive streaking or band broadening making this technique amenable to both analytical and preparatory work. Additionally, thicker gels (up to 2 mm) behave in the same way and offer increased capacity for preparative scale purifications. The same gel when subjected to rapid Coomassie blue staining gave a comparable detection limit (Fig. 4-1), the dye was found to effectively bind to PNA at concentrations below the detection limit for a typical protein.[8] Although there exist many examples of staining PNA duplexes in the literature,[9] the staining of single stranded PNA has not been hitherto reported.

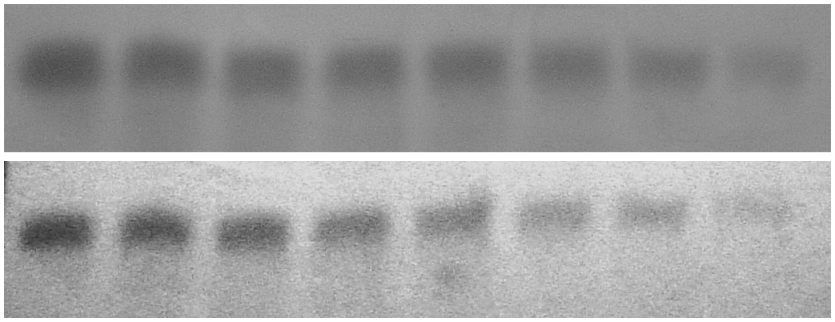


Figure 4-1: Top: detection of PNA 1 by Coomassie blue. Bottom: detection of PNA1 by UV shadowing. Decreasing amount of PNA, from left to right: 6.99, 6.12, 5.24, 4.37, 3.50, 2.62, 1.75, 0.87 μg.

The resolution of the polyacrylamide gel as compared to HPLC can be seen in Figure 4-2. This particular example illustrates that peaks observed in the HPLC trace could not be resolved under optimized conditions, yet in the gel the major band can be easily separated from the minor by-products.

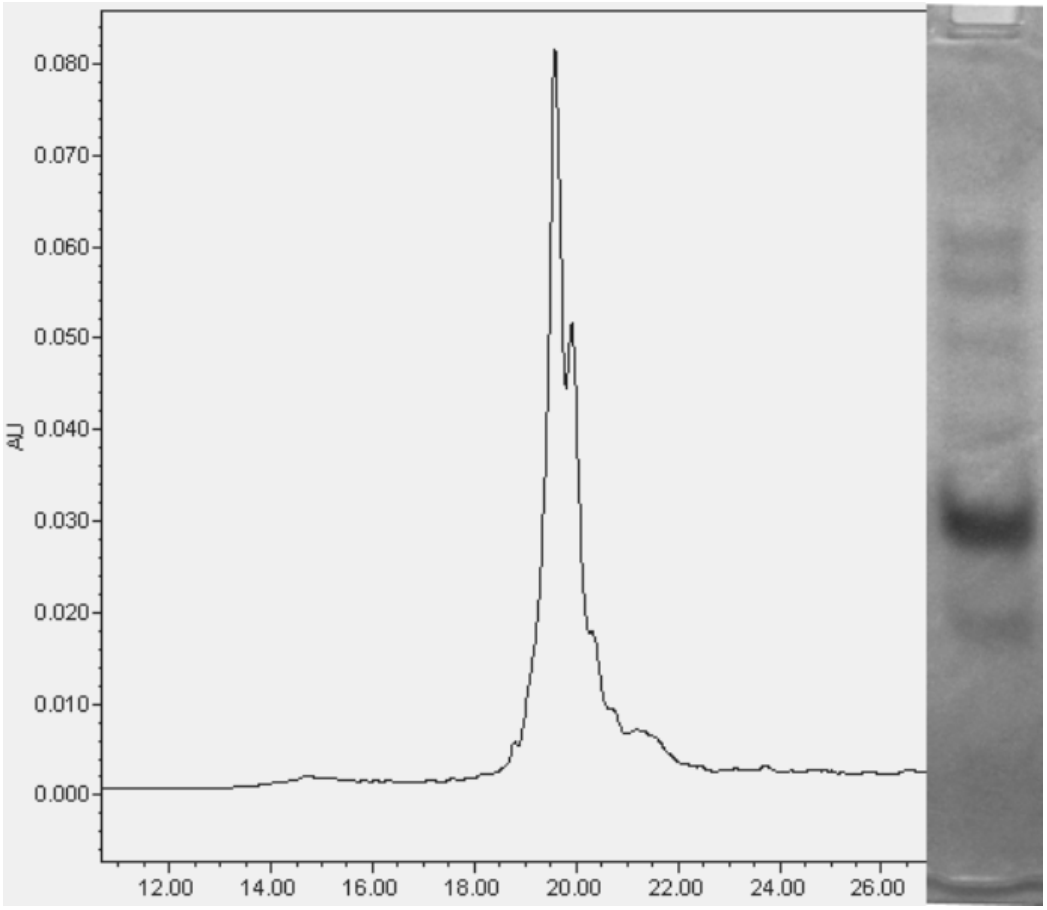


Figure 4-2: Comparison of polyacrylamide gel electrophoretogram and HPLC trace of crude PNA 2.

Dependence of gel mobility on molecular weight

Single base resolution is observed between PNAs 9-12 as highlighted in Figure 4-3, these oligomers differ by a mass of < 300 Da. When the inverse log of the molecular weight is plotted against migration distance a linear relationship is observed between oligomers of comparable sequence identity (Fig. 4-3). As is the case for PAGE of DNA, the absence of charge differences gives an apparent linear relationship between the partition coefficient and log molecular weight.[10] Acetylation of the *N*-terminal amino group greatly reduces the mobility of the PNA as this reduces the overall charge of the molecule by one unit. From these data it is clear that loss or gain of a positive charge results in a

remarkably large change in migration. Since acetylation (or lack thereof) is a common source of impurities in PNA synthesis, the separation of these products is essential and readily achieved by PAGE.

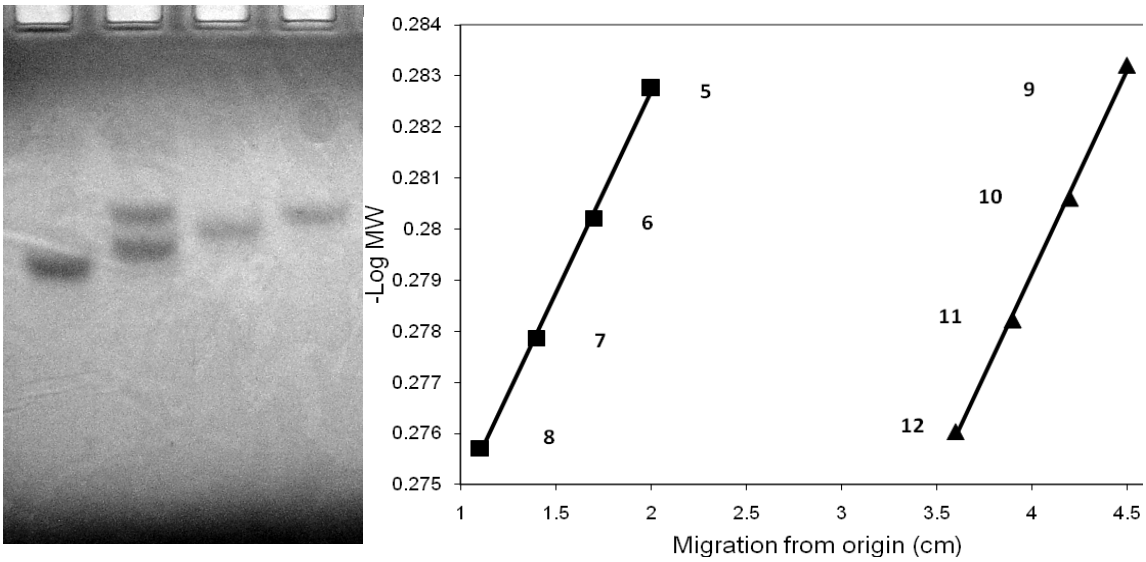


Figure 4-3: Left: Single base resolution of PNAs 9-12 (dicationic) from left to right. Note: lane 2 contains both PNA 10 and 12. Right: Relationship between distance travelled through the gel and oligomer molecular weight. Solid squares: monocationic PNAs. Solid triangles: dicationic PNAs.

Dependence of gel mobility on PNA sequence

The environment of the gel matrix was found to be approximately at pH 3 by electrophoresis of indicator dyes; this value is supported in the literature.[7] Under these acidic conditions, based on the reported pKa values of the nucleosides, three of the four natural bases would be expected to bear a partial positive charge. In order to confirm the contribution of charge to the oligomer from protonation of the nucleobases, the 4 natural deoxynucleosides were run through the gel under identical conditions to the PNAs and their mobility was found to be C>A>>G>T (Fig. 4-4). Therefore, the amount of charge on an oligomer depends not only on pendant basic groups but also on the sequence and is strongly dependent on A and C content while being insensitive to T and G content.

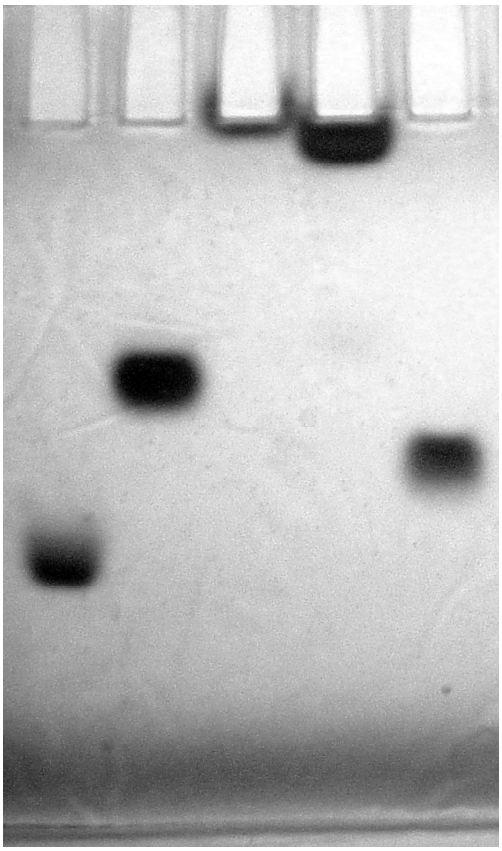


Figure 4-4: Mobility of deoxynucleosides. From left to right: Methylene blue; 2'-deoxyadenosine; 2'-deoxythymidine; 2'-deoxyguanosine; 2'-deoxycytidine.

Substitution of C with the fluorescent analogue phenylpyrrolocytosine (PhpC, 6-phenyl-3H-pyrrolo[2,3-d]pyrimidin-2(7H)-one) slows the PNA beyond what could be explained by the modest increase in molecular weight alone (Fig. 4-5). It is postulated that this hetereocycle is not protonated at this pH under these conditions due decreased pKa caused by the fusion of the aromatic pyrrole ring to cytosine. The ability to visualize the fluorescently-labeled PNA while in the gel is quite useful for aiding identification and separation from nonfluorescent by-products.

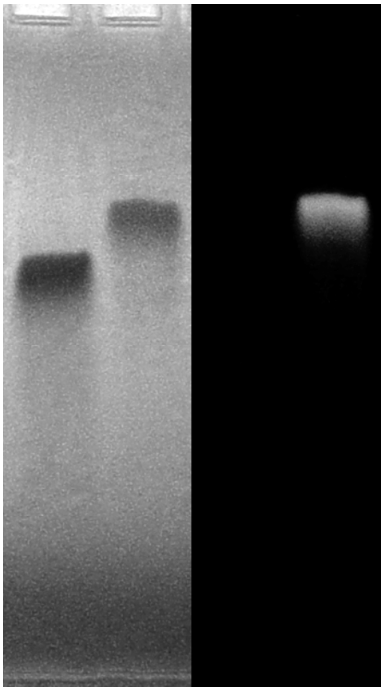


Figure 4-5: Illumination at 260 nm (detection by UV shadowing, left) of PNAs 3 and 4 compared to fluorescence detection (illumination at 360 nm) of PNAs wherein only 4 is visible due to the presence of the fluorescent nucleobase replacement phenylpyrrolocytosine (PhpC).

Extraction of PNA oligomers from the polyacrylamide matrix

Using the rapid freeze/thaw technique (see experimental section) recovery of the PNA was above 80%. This is somewhat less than the near quantitative extraction of DNA from polyacrylamide gels using a comparable technique and may be due to the lower overall charge on PNA and the corresponding lower aqueous solubility of the PNA oligomers. Sonication of the gel fragments was found to be requisite. In order to ensure the PNAs were not degraded in the elution process the samples were lyophilized and run again on a gel against the original sample, identical migration was observed (Fig. 4-6). The identity of the PNA extracted from gels has also been confirmed by mass spectral analysis and HPLC traces of the eluted fractions also confirmed the PNAs were not degraded. Although the PNA recovered was contaminated with urea and salts, these are easily removed by size exclusion chromatography.

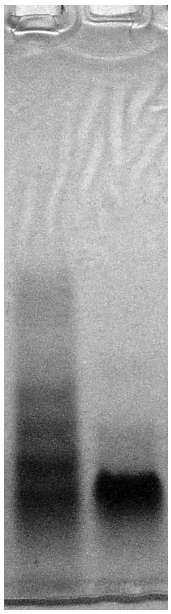


Figure 4-6: A crude mixture (left) run adjacent to a previously excised and extracted band (right).

As a means of recovering PNA, electroelution was found to be less efficient (40%). In an attempt to increase the recovery of PNA from the gel, chemical depolymerization of the gel matrix using aqueous $CuSO_4/H_2O_2$ was performed followed by reversed-phase solid phase extraction of the supernatant.[11] This aggressive approach yielded only unidentifiable low molecular weight compounds implying degradation of the PNA.

4.3 Conclusions

In summary, a simple, accessible and inexpensive protocol for the analysis and purification of PNA oligomers by denaturing polyacrylamide gel electrophoresis has been developed as for PNA chemistry. The electrophoretic mobility was quite sensitive to the A, C content and overall charge on the PNA oligomer and single-base resolution was achieved. This technique operates under conditions that are more mild (pH 3) than traditional reversed-phase HPLC (0.1 % trifluoroacetic acid in water, pH 1) and would be equally amenable to other non-ionic nucleic acid analogues such as morpholinos[12] or methylphosphonates.[13]

4.4 Experimental

Gel polymerization

All gels discussed are 26% polyacrylamide, 1.6% bisacrylamide. The polyacrylamide gels were cast according to a modified literature procedure.[14,15] In a typical experiment, 7.4 g urea (MP Biomedicals, Irvine, Ca., USA) were dissolved in 6.5 mL of 60% (w/v) aqueous acrylamide and 1.1 mL 2.5% (w/v) aqueous *N,N*'-methylenebisacrylamide (purchased as the solids from Biorad, Richmond, Ca., USA). Glacial acetic acid (925 μL, Caledon Laboratories, Georgetown, On, Canada) and 50 μL concentrated aqueous ammonia (BDH, Toronto, On, Canada) were added and the final volume adjusted to 15 mL with 18 MΩ deionized water (Waters Milli Q system). The solution was degassed by application of vacuum and polymerization was initiated by the addition of 300 μL of 10% (w/v) aqueous ammonium persulfate (APS, BDH) and 90 μL of *N,N,N',N'*-tetramethylethylenediamine (TEMED, Aldrich, Milwaukee, Wisconsin, USA). Plate margins were taped or sealed with agarose and clamped and the solution poured between plates measuring 8×10 cm with 0.15 cm spacers (Mini-Protean Vertical Slab Gel System, Bio-Rad Laboratories, Hercules, CA). The running buffer used was 1 M acetic acid, 0.1 M glycine and the gels were pre-electrophoresed at a constant voltage of 10 V/cm overnight at room temperature or until the distortion front was seen to pass entirely through the gel. To prevent distortion during pre-electrophoresis the wells were overlaid with a 10 % (w/v) polyethylene glycol (PEG - 8000, BDH, Toronto, On), 8 M urea, 1 M acetic acid and 50 mM NH_4OH solution according to a literature procedure.[16] This procedure was found to be scalable as a larger apparatus worked equivalently (Protean II xi Vertical Slab Gel, Bio-Rad Laboratories, Hercules, CA).

Sample preparation

PNA oligomers were quantitated at 260 nm on a Cary 300 Bio UV-Vis spectrophotometer using the nearest neighbour approximation.[17] Specific volumes of aqueous PNA samples of known concentration were lyophilized and subsequently redissolved in 10 μL of 8 M urea acidified with 5 % acetic acid (v/v), heated to 90 ˚C for 1 minute, allowed to cool to room temperature and then loaded into the gel.

Electrophoresis and visualization

Methylene blue (J. T. Baker Inc., Phillipsburg, NJ, USA) was used as a tracking dye; a 2 % aqueous solution was mixed in equal proportions to loading buffer and 1 μL was either added to the PNA samples after heating or loaded directly into an adjacent lane so as not to obscure rapidly migrating bands. Samples were run under a constant voltage of 20-30 V/cm. Generally, no PNA ran ahead of the methylene blue, and often the experiment was run until the dye had passed completely through the length of the gel. After electrophoresis the bands were visualized by placement of the gel upon a fluorescent TLC plate and illuminated with a handheld UV lamp at 260 nm. Bands could also be visualized by rapid Coomassie blue staining carried out according to a modified literature procedure.[8] The concentration of the fixing solution was adjusted to 40% isopropanol, 10% acetic acid and 50% H_2O. All other steps were carried out as previously reported.

PNA recovery

To recover PNA from the polyacrylamide matrix, the appropriate bands were excised and placed in 0.5-1.0 mL of deionized Milli Q water. The samples were frozen in either liquid nitrogen or dry ice and thawed rapidly at 90 ˚C. After reaching 90 ˚C, the samples were then sonicated at 60 ˚C for 1 h. Pulverizing the gel did not result in improved sample recovery and therefore was deemed unnecessary. The extraction of PNA from the gel was monitored by visualization of the excised gel fragment over a fluorescent TLC plate and was stopped when there was no evidence of PNA remaining in the gel fragment. The abovementioned procedure resulted in extraction of PNA to below visual detection limits. The PNA samples were then passed through a Millex®-LG 0.2 μm syringe driven filter (Millipore, USA) to remove any particulate matter. The samples were then desalted using either Bio-Gel or Sephadex-based size exclusion chromatography and isolated by lyophilization.

PNA synthesis and characterization

The PNA oligomers used in this study were synthesized according to a standard literature procedure from the Fmoc-protected monomers supplied by ABI, Foster City, CA, USA.

Alternatively, monomers synthesized by literature procedure were used,[18] table 4-1. Oligomers incorporating unnatural nucleobases (PNA 2 and 4) were prepared according to literature procedures.[19,20] Crude PNAs were purified by RP-HPLC performed using a Waters 600 instrument using a Microsorb reversed-phase column (5 µm, 300 Å, C-18 bonded-phase from Varian Analytical Instruments, Walnut Creek, CA) using a linear gradient of 0.1 % trifluoroacetic acid (TFA) in water: 0.1 % TFA in acetonitrile from 100:0 to 5:95 over 40 minutes, at 45 ºC. The oligomer identity was confirmed by ESI-TOF mass spectral analysis.

4.5 References

1. Nielsen, P.E. *Mol. Biotechnol.* 2004, *26,* 233-248.

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

3. Nielsen, P.E.; Egholm, M. *Peptide Nucleic Acids Protocols and Applications*, Horizon Scientific Press, Wymondham, Norfolk, UK **1999**.

4. Shmidt, J.; Nielsen, P.E.; Orgel, L.E.; *Anal. Biochem.* **1996**, *235*, 239-241.

5. Eriksson, M.; Christensen, L.; Shmidt, J.; Haaima, G.; Orgel, L.E.; Nielsen, P.E. *New J. Chem.* **1998**, *22*, 1055-1059.

6. Bentin, T.; Larsen, H.J.; Nielsen, P.E. *Peptide Nucleic Acid Targeting of Double Stranded DNA in Peptide Nucleic Acids Protocols and Applications*. Horizon Bioscience, Wymondham, Norfolk, U.K. **2004**, pp. 117-121.

7. Gabor, I.L. *BioTechniques* **1999**, *27*, 798-808.

8. Sasse, J.; Gallagher, S.R. *Detection of Proteins in Current Protocols in Molecular Biology,* John Wiley & Sons, **2003**, 10.6.1-10.6.25.

9. Datta, B.; Wang, M.; Armitage, B. A. *PNA Beacons and Fluoroprobes in Peptide Nucleic Acids Protocols and Applications*. Horizon Bioscience, Wymondham, Norfolk, U.K. **2004**, pp. 141-152.

10. Chrambach, A.; Rodbard, D. *Science* **1971**, *172*, 440-451.

11. Donato, H.; Doig, M.T.; Priest, D.G. *J. Biochem. Biophys. Met.* **1998**, *15,* 331-335.

12. Summerton, J. *Biochim. Biophys. Acta,* **1999**, *1489*, 141-158.

13. Miller, P.S. *Biotechnology* (N Y). **1991**, *9*, 358–362.

14. Panyim, S.; Chalkey, R. *Arch. Biochem. Biophys.* **1969**, *121*, 337-346.

15. Waterborg, J.H. *Acetic Acid-urea Polyacrylamide Gel Electrophoresis of Basic Proteins in The Protein Protocols Handbook, 2nd Edition*., Humana Press Inc., Totowa, NJ., **2002**, pp. 103-111

16. Paulson, J.R.; Higley, L.L. *Anal. Biochem.* **1999**, *268,* 157-159.

17. Puglisi, J.D.; Tinoco, I., Jr. *Methods Enzymol. RNA Process. Pt. A.* **1989**, *180,* 304-325.

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

19. Hudson, R.H.E.; Liu Y.; Wojciechowski F. *Can. J. Chem*. **2007**, *85*, 302-312.

20. Dabenieks, A.K. *M.Sc. Thesis*, University of Western Ontario, 2006.

Chapter 5

5 Reaction of 5-ethynyldeoxycytidine with 1,3-dipoles as a route to heterocycle-appended, blue fluorescent base analogues.

Fluorescently-labeled nucleobases are frequently used for the reporting of nucleic acid structural phenomena. Post-synthetic modification is particularly attractive as groups that are sensitive to solid-supported nucleic acid chemistries may be appended. Herein we report the participation of 5-ethynyldeoxycytidine in the Huisgen 1,3-dipolar cycloaddition with azides and nitrile oxides for the modular generation of useful fluorophores. The reaction may proceed both with the monomer and, in the future, within the context of an oligomer. Synthesis and photospectroscopic characterisation of intrisically fluorescent triazole- and isoxazole-appended cytidines is reported. Fluorescence was found to be highly dependent on solvent conditions and X-Ray crystallographic data show the proton of the exocyclic amine of the nucleobase and the triazole *N* engaged in a *H*-bond. Excerpts of this chapter are taken from Dodd, D.W., Swanick, K.N., Price, J.T; Brazeau, A.L.; Ferguson, M.J.; Jones, N.D; Hudson, R.H.E. *Org. Biomol. Chem*. **2010**, *8*, 663-666.

5.1 Introduction

Modified nucleosides have been used to elucidate nucleic acid structural dynamics for over half a century. In this pursuit, luminescent derivatives have enjoyed the most success, and the literature is replete with reports of analogs that respond fluorometrically to hybridization within the context of an oligomer (see chapter 1.1 for further detail).[1-7] Fluorescent probes may be divided into two general classes: (1) those bearing a remotely appended fluorophore, separated electronically and in space from the base-pairing face; and (2) those in which the nucleobase is integral to the fluorophore. Members of the latter class are more attractive because their fluorescent responses are more clearly indicative of specific hybridization events. We aim to synthesize compounds that may be used as base discriminating fluorophores, created pre- or post-synthetically. The following paragraphs describe some existing techniques for the fluorescence labeling of nucleic

acids.

Selection of methods for nucleic acid fluorescence labeling.

As the inherent fluorescence properties of nucleic acids are limited, for many years it has been sought to label nucleic acids for fluorescence-based applications. These applications are wide ranging; fluorescence microscopy, pertaining to chromosome analysis and fluorescence in situ hybridization; gel staining, for better detection limits resulting from lower background readings without the need of radiolabeling; and quantitative PCR, measuring amplification of cDNA in real time, are just a few. Tagging of DNA with affinity probes, biotin, and thiols for attachment to surfaces is also desirable. Below is a very small segment of the array of chemistries available for these purposes (Fig. 5-1). For

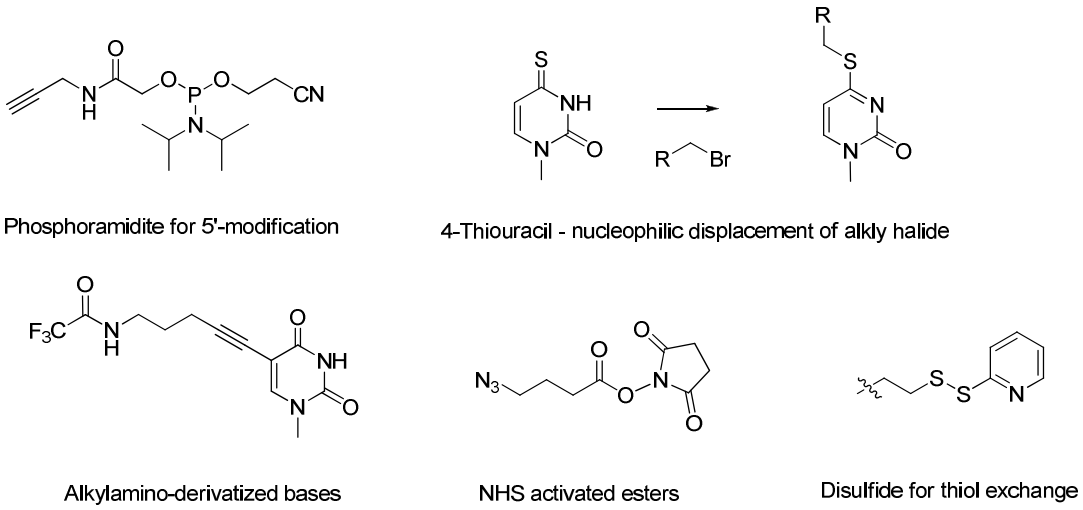


Figure 5-1: A small selection of functionalities used for the labeling of nucleic acids.

modification of phosphate groups modified phosphoramidites may be used. Pictured above is an alkyne functionalized moiety for use in cycloaddition reactions. It has been recently been used for templated ligation of oligonucleotides and lariat formation but can also be used for the attachment of conventional fluorophores.[8] One of the first base modifications for labeling was described in the early 1970s by Leonard's group. 4-thiouracil can be reacted with a multitude of alkyl halides, the first of which was a bromomethyl-derivatized coumarin[9] yielding a fluorescent base analogue which is unfortunately no longer base-pairing competent. Masked alkyl amino groups may be

attached to bases. After protecting group removal the exposed amine may react with a variety of active esters or alkyl halides.[10] Disulfides can also be incorporated and are of great importance for biosensors as they through their affinity to gold surfaces. They may also undergo thiol exchange or be unmasked and reacted with Michael acceptors such as maleimides or iodoacetamides. The maleimide functionality may also be used in a Diels-Alder reaction with dienes.[11]

Azides have multiple uses as a reporter groups, they are absent from biological systems, unreactive to most functionalities, and are readily synthesized (typically by S_N2 or S_NAr mechanisms). They have been used as photoaffinity probes since the late 1960s[12] and more recently, in an adaptation of the classic Staudinger reduction,[13] termed the Staudinger ligation (Fig. 5-2).[14] Although this reaction was originally designed for protein labeling, it has also become quite popular with nucleic acid chemists and therefore warrants discussion.[15] Phoshines react with azides to form iminophosphoranes which are readily hydrolyzed; however, in Bertozzi's design, a proximal ester may act as an electrophilic trap for the aza-ylid. Subsequent hydrolysis then gives a stable amide bond and a non-toxic phosphine oxide.

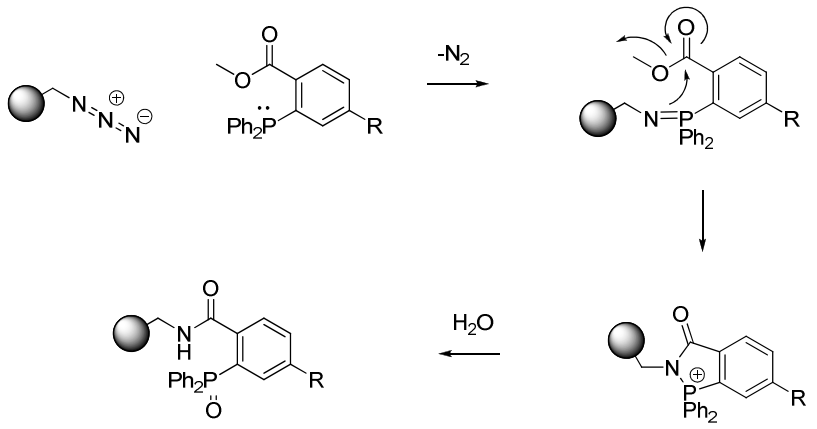


Figure 5-2: The Staudinger ligation developed by Bertozzi for protein labeling, since adopted for use with nucleic acids.

Copper-catalyzed Huisgen cycloaddition of azides and alkynes (CuAAC).

The Cu(I)-catalyzed Huisgen cycloaddition of azides and alkynes, to give 1,4-disubstituted 1,2,3-triazoles,[16,17] has been exploited for rapid drug discovery,[18] for conjugation of modified biological macromolecules,[19] for derivatization of nanoparticles[20] and (more relevent to our work) for the modification of nucleosides and DNA.[21,22] This cycloaddition is particularly well suited to nucleic acid chemistry for several reasons. First of all, azides and alkynes are unreactive towards the functional groups present in DNA, indeed to almost all functional groups present in living systems (azides will react slowly with thiols).[23] Azides and alkynes are small and do not perturb the natural structure of nucleic acids to any great extent and the triazole generated is non-toxic

Ethynyl- and octadiynyl-modified uracil are the most commonly used substrates for this "click" chemistry, with coumarin dyes, benzyl groups and biotin labels being attached at the 5-position. Generally, the 5-postion of pyrimidines is a useful site for manipulation due to its inherent reactivity and the propensity for substituents appended here to project into the major groove while leaving the Watson-Crick face unhindered.[1-7] The purine analogs, 7-deazaadenosine and 7-deazaguanosine, have been used to similar ends as the 7-position can also be nicely accommodated in the major groove (Fig. 5-3).[24-27]

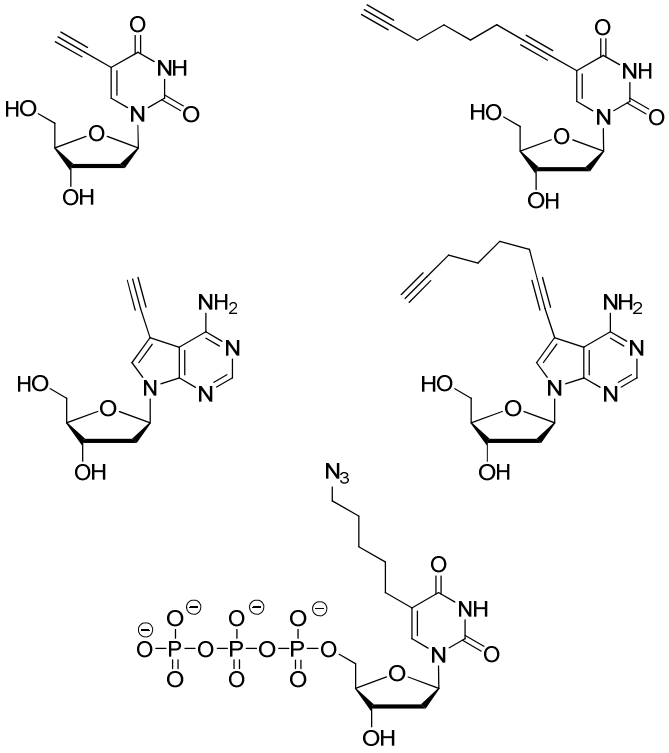


Figure 5-3: Nucleoside(triphosphate)s for labeling of nucleic acids.

Azidonucleosides may not be used for phosphoramidite chemistry as they undergo a Staudinger reduction in the presence of phosphorus (III) but may still be accommodated using enzymatic techniques such as PCR or by being post-synthetically incorporated.[28]

Seela has reported the synthesis of a pyrrolodeoxycytidine bearing a pendant alkyne which can undergo a cycloaddition with a non-fluorescent azidocoumarin to give a fluorescent triazole (Fig. 5-4).[25,29] Many such azides are commericially available that fluoresce on reaction with an alkyne. These are desirable as there is little interference (background fluorescence) from leftover starting materials; however, for us these belong to the first category of base-discriminating fluorophores described: those bearing an appended, traditional chromophore.

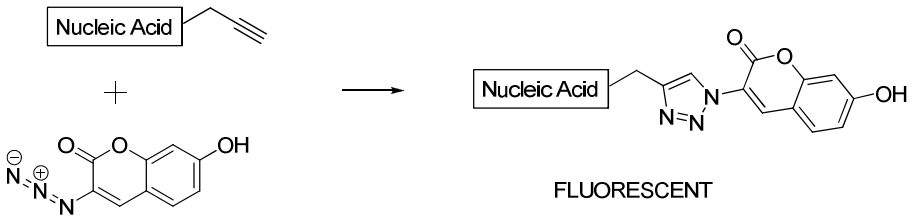


Figure 5-4: Non-fluorescent starting materials to give a fluorescent triazolylcoumarin.

Another popular position of attachement is the 2'-*O* of ribose. This is easily done by reaction of the suitably protected nucleoside with propargyl bromide. Wagenknecht and others have successfully applied this strategy for the generation of fluorescent oligonucleotides.[30] Although sometimes destabilizing towards duplex formation, large chromophores attached at the 2'-postion are oftentimes useful reporter groups.[30] Again, however, this method does not satisfy our goals of creating minimally modified base-discriminating fluorophores.

Many more examples of click-modified nucleosides exist. But, however popular this method is for uracil and purine analogs, the corresponding cytidine analogues are notable by their absence. Prior to the present work, there had been no reports of the reactivity of 5-ethynylcytidine towards azides, nor have there been studies of the fluoresence of any conjugated triazolylpyrimidines in the absence of an extraneous fluorophore. It has been found that 8-ethynyldeoxyadenosine undergoes reaction with azides and the resultant triazoles were highly fluorescent (as already described in chapter 1.1). Therefore, we deemed fluorescent triazolyldeoxycytidines to be prime targets for potential use in the various facets of nucleic acid chemistry.

"Metal-free" cycloadditions.

Post-synthetic labelling of DNA using the copper(I)-catalyzed azide-alkyne cycloaddition (CuAAC) has been reduced to practice using alkyne-appended nucleobases, of which some are commercially available.[31] Disadvantageous to this procedure is, however, the

use of Cu(I) as catalyst, which must be ligand supported so as to not reductively damage nucleic acids[32] the half life of DNA in the CuAAC when no suitable ligand is present is only a few minutes.[33] Copper-free azide-alkyne cycloadditions can be promoted by using strained alkynes, however, these alkynes are not accessed simply. and are unsuitable to our puroposes as the strained alyknes[34] for the most part would result in unconjugated triazoles which would add little fluorescence benefit to the base (Fig. 5-2). Also, even when conjugated systems do result, a mixture of regioisomers is obtained when no metal catalysis occurs. These regioisomers would presumably result in different binding affinities to a target strand when in the context of an oligomer and would also have different photopysical properties (as highlighted by our work on triazolyadenosines, see chapter 1.1).

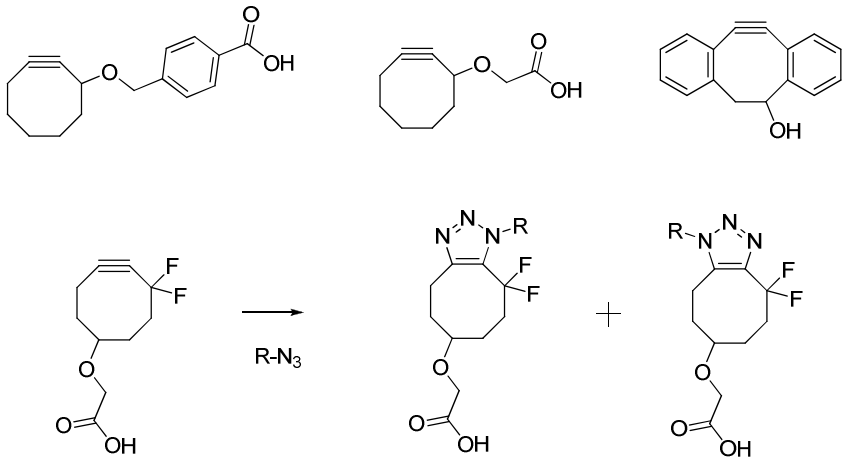


Figure 5-5: Cyclooctynes prepared by the group of Bertozzi that may undergo metal free cycloaddition (top) and the resultant regioisomers obtained though their use (bottom).

There is a wealth of Huisgen 3+2 dipolar cycloadditions available to the synthetic chemist and it is unreasonable to assume that only one may be amenable to solid-supported nucleic acid chemistries. Dehydrochlorination of hydroximoyl chlorides to form one of the now most well known 1,3-dipoles, the nitrile oxide, was carried out as early as 1894 by Werner and Buss.[35] Huisgen studied the cycloaddition of nitrile oxides with dipolarophiles (particularly alkenes) extensively and in the late 1950s and early 1960s and their reactivities allowed Huisgen to formally explain the concept of a dipolar cycloaddition.[36] Since that time, nitrile oxides have been used extensively in natural

product synthesis where they are typically used to generate isoxazoles and isoxazolidines as precursors to amino alcohols and α-hydroxycyclopentanones.[37] Typically, nitrile oxides are not isolable due to their propensity for dimerization so they are therefore prepared in situ. The two most popular methods are shown below in figure 5-6.

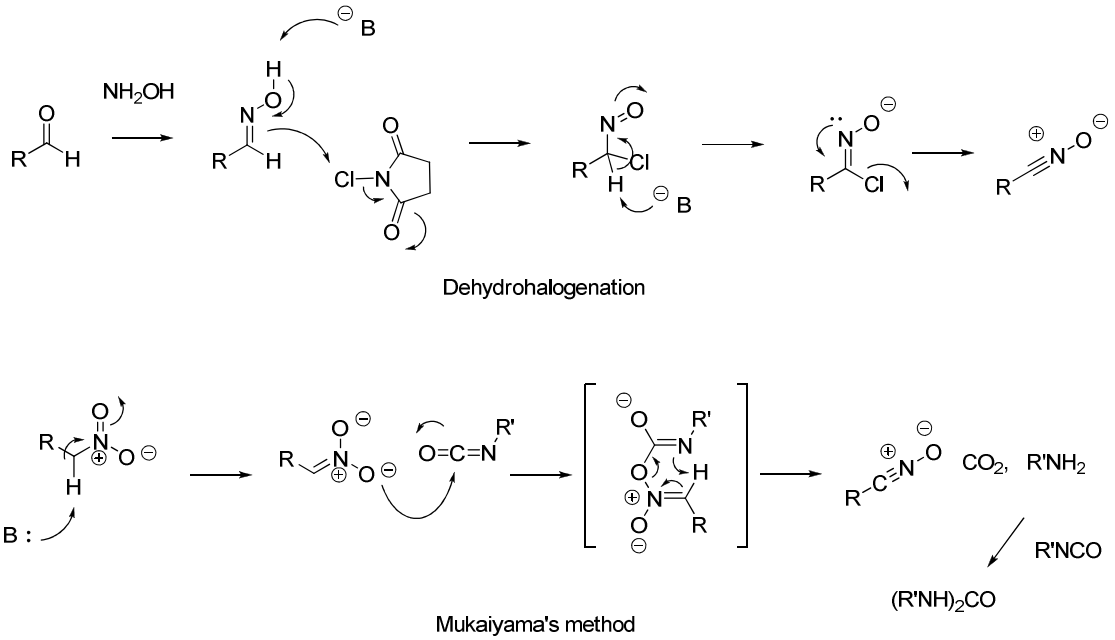


Figure 5-6: Two general methods for the in situ generation of nitrile oxides. Dehydrohalogenation of a hydroximoyl chloride (top) and dehydration of a nitroalkane (bottom).

The first method to be discussed was initially described by Huisgen and Mack in 1961.[38]An oxime is first prepared from the condensation of hydroxylamine and an aldehye, an electrophilic chlorine source is then introduced (shown here is *N*-chlorosuccinimide) to form the hydroximoyl chloride. These hydroximoyl chlorides can either be isolated or subjected to base directly thus yielding the nitrile oxide. The second method utilizes a nitroalkane in the presence of an excess of an arylisocyanate and was developed around the same time as the first by Mukaiyama.[39] Deprotonation of the nitroalkane yields a nucleophile which may attack the isocyanate carbon. Decomposition of the resulting intermediate gives the desired isonitrile and carbamate which promptly releases CO_2. Remaining aryl amine can then be consumed by excess isocyanate to give

an unreactive urea.

In support of the idea that the cycloaddition of nitrile oxides and an alkyne-appended base is amenable to solid-supported chemistries, it has been shown that the conjugation of a solid-supported, propargyl-derivatized DNA and various aryl nitrile oxides proceeds quantitatively and regioselectively through the formation of a 3,5-disubstituted isoxazole.[40,41] The proposed reaction is shown in comparison to the CuAAC in figure 5-7. Because the resultant isoxazole is aromatic, the reaction is stereospecific, and due to the lack of attention paid to the linkage is nucleic acid chemistry we also decided to pursue this family of nucleoside derivatives for further use as base-discriminating fluorophores.

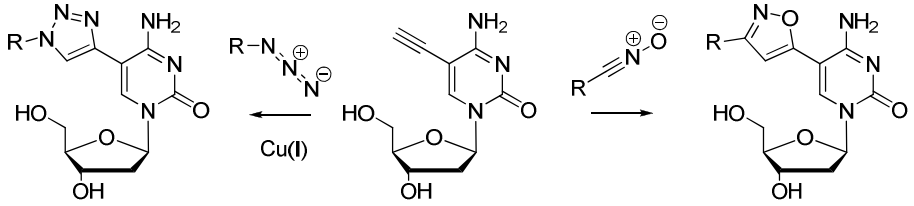


Figure 5-7: The CuAAC (left) as contrasted with the nitrile oxide/alkyne cycloaddition which may proceed without the need of catalysis.

5.2 Results and Discussion

The synthesis 5-ethynyl-2'-deoxycytidine is shown in figure 5-7. This would come to be the common starting material for both cycloaddition reactions. Starting from the HCl salt of 2'-deoxycytidine, the 2'- and 5'-hydroxyl groups were quantitatively acetylated for ease of purification in later steps. The protected nucleoside was then iodinated using I_2/HIO_3 in biphasic $AcOH/H_2O/CCl_4$ in good yield.[42,43] The product was chromatographed at this stage as any impurities were found to drastically reduce yields in the next step. Sonogashira coupling of the iodonucleoside with trimethylsilylacetylene in deoxygenated THF at room temperature using CuI and $Pd(PPh_3)_4$ (0.1 mol %) as catalysts, and subsequent deprotection with K_2CO_3 in methanol yielded 5-ethynyldeoxycytidine in 75% overall yield from deoxycytidine (Fig. 5-8).

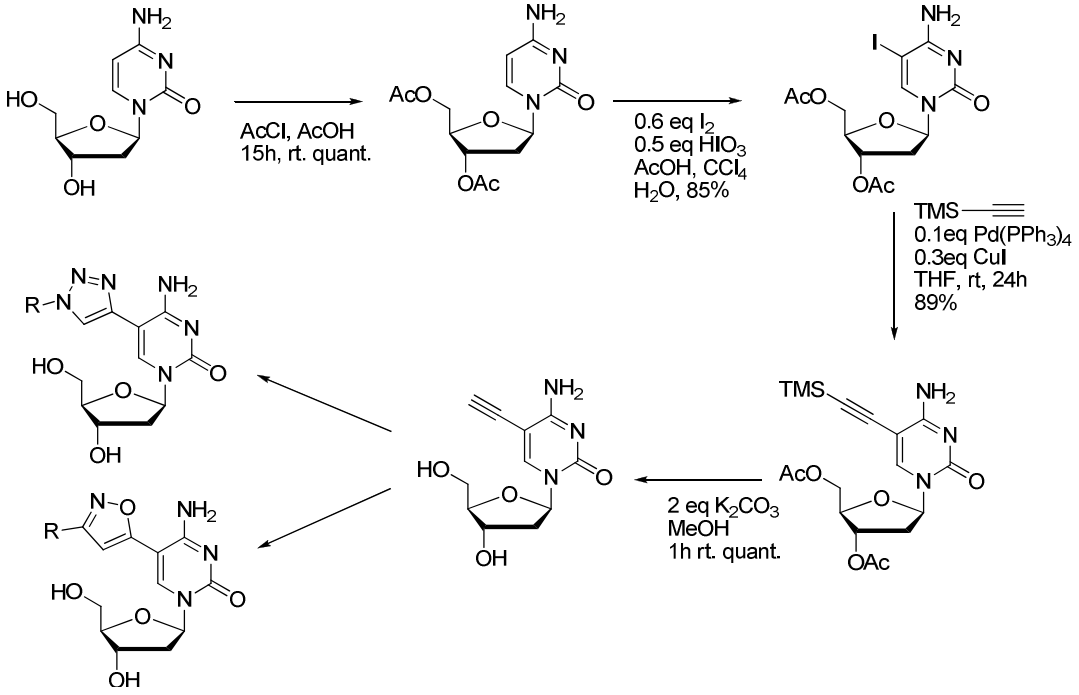


Figure 5-8: Synthesis of 5-ethynyldeoxycytidine to be used as a dipolarophile in Huisgen cycloadditions.

Triazolyldeoxycytidines: syntheses and photoluminescence.

5-ethynylcytidine was stirred with 4 different aryl azides in the presence of Cu(I) (generated *in situ* from $CuSO_4$ and sodium ascorbate) in 1:1 THF:H_2O solution. The resulting triazole-containing compounds **1**–**4** (Fig. 5-9) were characterised by ^{1}H and $^{13}C\{^{1}H\}$ NMR spectroscopies and HRMS. X-ray diffraction data were collected for **1** and **2** (see experimental).

Derivatives **1** and **3** were targeted for their likely visible fluorescence and also for onward electrochemical polymerization and attachment to surfaces. A fluorenyl-substituted nucleoside (not shown) was similarly targeted for fluorene's inherent fluorescence; however, repeated *in situ* oxidation of the intended compound at the doubly benzylic position to give **4** under the synthetic conditions led us instead to develop a rational synthesis of the fluorenonyl derivative from the corresponding azide.

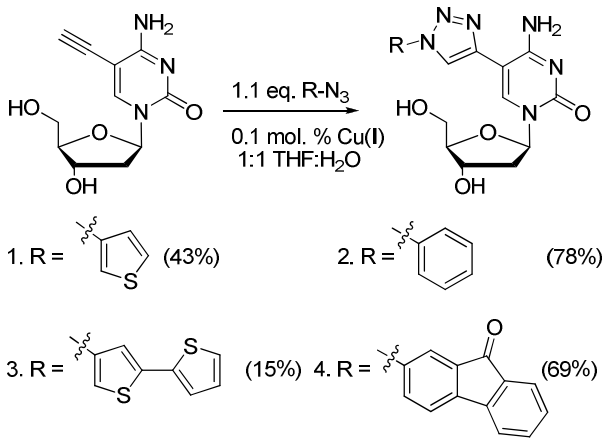


Figure 5-9: Triazolylnucleosides employed in the present work.

All of the deoxycytidine analogues were fluorescent in the blue (375–385nm). Quantum yields were determined by the relative method[44-46] using anthracene and quinine sufate as standards and ranged from low to moderate, the highest being that of **4** for which Φ_F = 0.089 in EtOH (Table 5-1). This figure is almost twice that of the commercially available and ubiquitous methylpyrrolocytosine, which has had continued success in DNA structural elucidation for many years.[47-50]

Table 5-1: Fluorescence quantum yields and excitation and emission maxima of compounds 1-4 in various polar solvents.

	Φ (H_2O)	Φ (EtOH)	Φ (DMF)	λ_{Ex} (EtOH)	λ_{Em} (EtOH)
1	0.002	0.011	0.005	330	375
2	0.003	0.013	0.004	330	380
3	0.006	0.066	0.047	320	380
4	0.006	0.089	0.060	320	385

Unusual, mutually consistent solvatochromatic effects were observed for all compounds. On changing the solvent composition from H_2O to EtOH in *ca.* 10 vol% increments, dramatic variations in emission intensity resulted (Fig. 5-10). Related non-linear solvatochromism has been observed for fluorenone-containing compounds and has been explained by the presence of multiple excited states that are differentially stabilized by

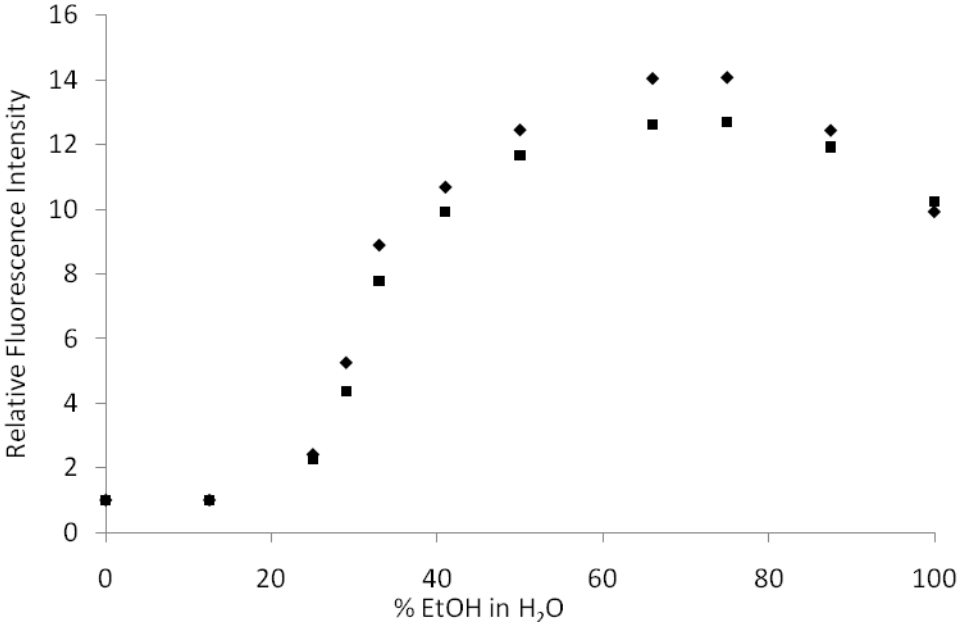


Figure 5-10: Variation in r.t. emission intensity of compounds 3 (♦) and 4 (■) at their respective maxima relative to emission intensity in neat H_2O as a function of solvent composition.

the solvation shell.[51] The compounds had single broad emission bands at r.t. (Fig. 5-11), but at low temperature showed three distinct, closely-separated emission maxima. These indicated multiple excitation states that could in principle be the source of the observed solvatochromism, although more resarch would be required to establish this unequivocally. Only very small shifts in excitation and emission maxima (< 5nm) were observed on varying the solvent composition. The response of a fluorescent nucleobase to changes in solvent is a good predictor of its performance as a base-pairing discriminator because match and mismatch combinations result in differential solvation of the excited state.[1] It is interesting that the triazolylnucleosides synthesised herein react similarly to solvent perturbations regardless of the aryl substituent, which implies that the triazole and the base are an integral part of the fluorophore. Many other solvents were investigated (EtOAc, CH_2Cl_2, MeOH, MeCN, 1,4-dioxane and cyclopentanone) and, to our surprise, it

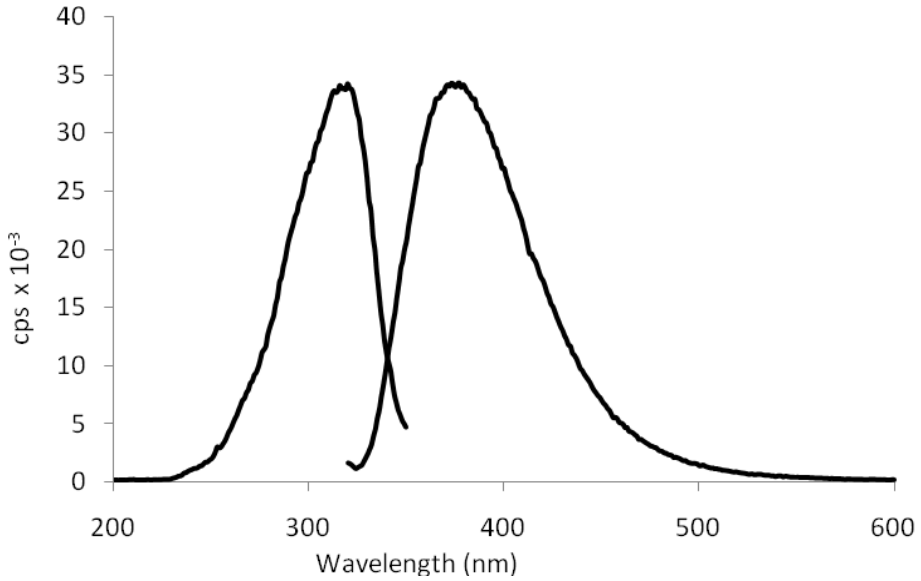


Figure 5-12: Fluorescence excitation and emission sprecta for compound 3 at r.t. in H_2O.

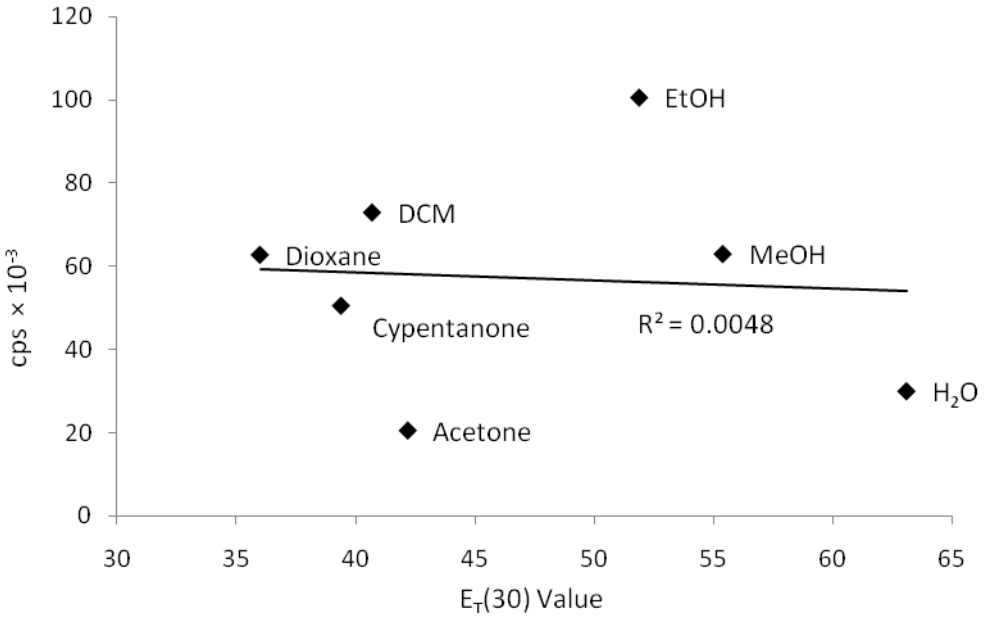


Figure 5-11: Fluorescence emission maxima for compound 3 in various solvents at room temperature plotted against solvent polarity as guaged by Et(30) values.

was found that solvent polarity (guaged by $E_T(30)$ values) was not correlated with emission intensity (Fig. 5-12).

X-ray diffraction revealed that the three aromatic rings are coplanar in the solid state: torsion angles around the ring junctions are < 10 (Fig. 5-13). A proton of the exocyclic amine of the nucleobase and N3 of the triazole ring are separated by *ca.* 2 Å, which implies stabilization of coplanarity by an intramolecular H-bond

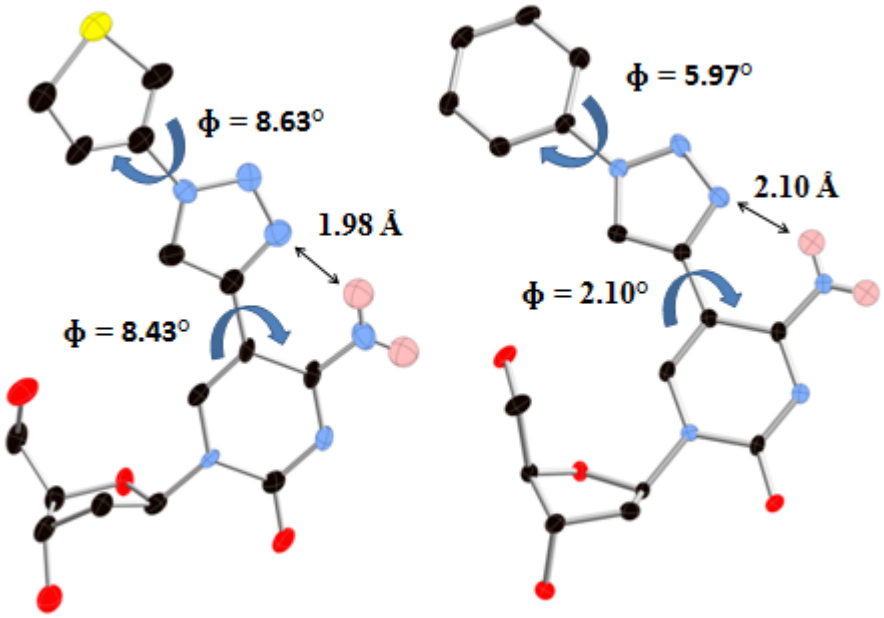


Figure 5-13: ORTEP representations of the molecular structure of compounds 1 (left) and 2 (right) with ellipsoids at 30% probability showing the dihedral angles about ring junctions and and interatomic distance suggesting a hydrogen-bonding arrangement.

Gas-phase DFT calculations were carried out by Alison Brazeau and were used to elucidate the effect of this H-bond on the HOMO-LUMO gap in **1**, which should be sensitive to the coplanarity of the rings, and also to determine the MOs responsible for the observed fluorescence. Mapping these MOs revealed that the nucleobase itself is indeed an integral part of the fluorophore. In all compounds **1**–**4**, the HOMO is derived predominantly from the triazole and nucleobase heterocyles, and in all but one case (**4**),

there is also a small but significant contribution to the LUMO from the base. This implies that these analogues may exhibit greater sensitivity to hybridization events than conventional fluorophore-appended nucleosides.

Nitrile oxide/alkyne cycloaddition of 5-ethynylcytidine

With the help of Jacqueline Schulman, at the time an undergraduate student in the Hudson laboratory, pyrenylisoxazolylribocytidine **5** was synthesized by reaction of pyrene nitrile oxide (generated *in situ* from the corresponding oxime using NaOCl or chloramine-T) and 2'-3'-5'-*O*-triacetyl-5-ethynylcytidine at room temperature in THF. Yield for the cyclization reaction was below 10%; however, slow addition of NaOCl overnight with a syringe pump allowed for increased yields (35%) that are yet to be fully optimized. Subsequent removal of the protecting groups yielded the ribonucleoside for photophysical characterization (Fig. 5-14).

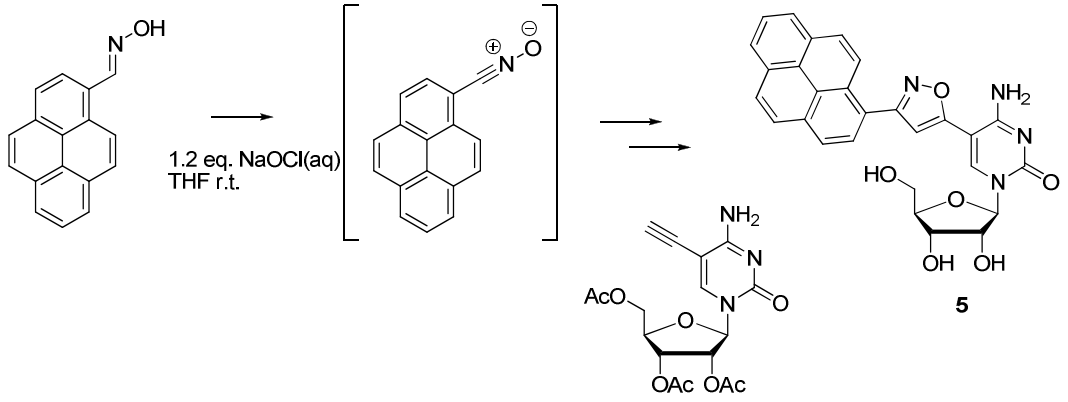


Figure 5-14: Synthesis of pyrene derivatized isoxazolylribocytidine 5.

This work is still preliminary but fluorescence data obtained for **5** show promise. The pyrenyl substituted compound showed a high quantum yield of 0.28 in ethanol. This is demonstrated in figure 5-15. Plotting emission against abrorption can allow for the comparison of newly synthesized compounds to fluorescence standards (see experimental section for greater discussion). Fluorescence was significantly quenched (up to 20 fold) in water/ethanol mixtures, as shown in figure 5-16. It appears that in more polar environments excimer emission may be being observed as is indicated by the broad nature of the emission spectrum in 5:1 H_2O:EtOH. This solvatochromism is a good

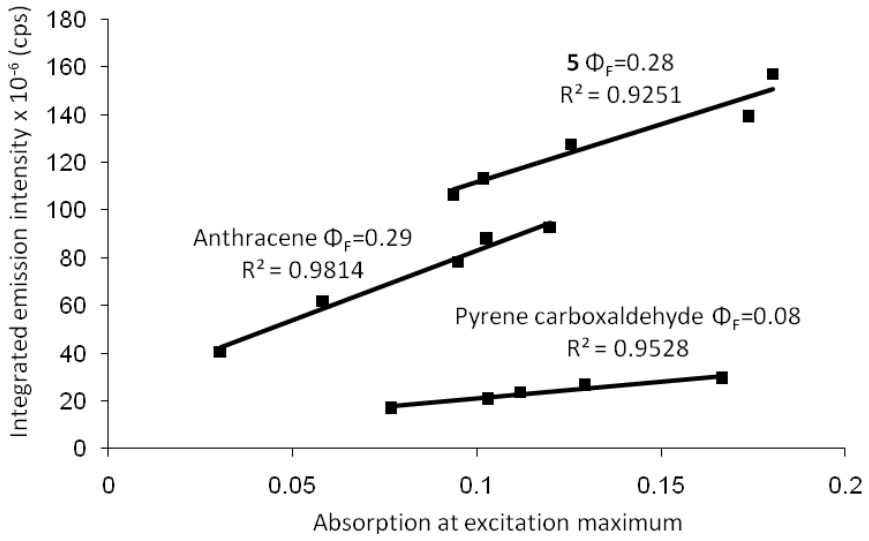


Figure 5-15: Plot of integrated emission versus absorption to determine the quantum yield of 5 (see experimental).

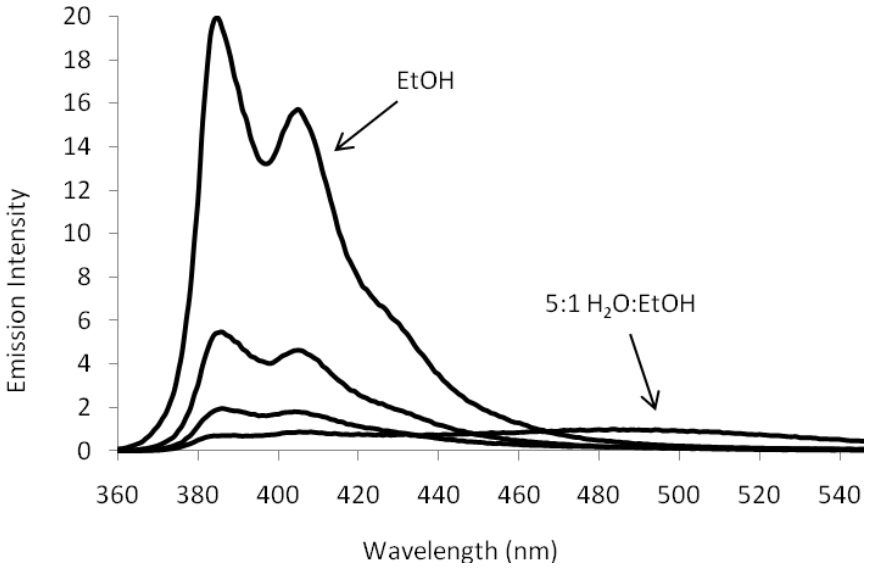


Figure 5-16: Solvatochromism of compound 5 in H_2O/EtOH mixtures. Drastic reduction in emission occurs upon increase in solvent polarity.

indication that the analogue may report on binding events when in the context of an oligomer. More interesting still is the potential application for experiments based on FRET as the photophysical properties of pyrene are well established in this regard. Excimer formation by pyrene has been exploited for hybridization assays[53-55] and we intend to place these analogues proximately in order to gauge hybridization events and attempt to define mismatch sequences.

5.3 Conclusions

In conclusion, a family of fluorescent nucleoside analogues has been made, and their potential for use in DNA structural analysis has been evaluated through study of their photophysical properties. Compounds **3** and **4** are more intensely fluorescent than many probes currently being applied in practical biological assays and are unusually sensitive to the nature of the solvent. Crystal structures reveal the presence of an intramolecular H-bond which enhances the coplanarity of the heterocycles. These X-ray diffraction data are the first for any triazolyl-nucleosides despite the great attention paid to the class represented in the literature. Their incorporation into oligomeric nucleic acids will follow shortly. Further work will involve synthesis of the phosphoramidite of 5-ethynyl-2'-deoxycytidine to be incorporated into DNA and its reactivity towards nitrile oxides and azides explored within the context of an oligomer.

5.4 Experimental

General Considerations:

All chemicals were obtained from commercial sources and were used without further purification. Solvents were dried by passing through columns of activated alumina. Flash column chromatography (FCC) was performed on Merck Kieselgel 60, 230-400 mesh and thin layer chromatography (TLC) was performed on Merck Kieselgel F-60 plates. Chemical shifts are reported in parts per million (δ), were measured from tetramethylsilane (0 ppm) and are referenced to the residual proton in the deuterated solvent: $CDCl_3$ (7.26 ppm), DMSO-*d*6 (2.48 ppm), D_2O (4.75 ppm) for ^1H NMR and $CDCl_3$ (77.0 ppm), DMSO-*d*6 (39.5 ppm) for ^{13}C NMR. Multiplicities are described as s

(singlet), d (doublet), t (triplet), q (quartet), m (multiplet) and br s (broad singlet). Coupling constants (*J*) are reported in Hertz (Hz). Exchangeable protons were identified by their disappearance when the sample was shaken against D_2O. High resolution mass spectra (HRMS) were obtained using electrospray ionization time of flight methods (ESI-TOF). Solution ^{1}H NMR spectra were collected using a Varian 400 spectrometer (400.09 MHz for 1H and 100.61 MHz for ^{13}C) at r.t. unless otherwise noted.

Synthesis of 3', 5'-diacetyl-2'-deoxycytidine

2'-Deoxycytidine HCl salt (0.95 g, 3.6 mmol) was dissolved in AcOH (10 mL) and stirred at 35°C until dissolution at which time the reaction was allowed to come to rt. A mixture of $CHCl_3$ (8 mL) and AcCl (3 mL) was added and the reaction continued to be stirred overnight after which time the colourless solution was concentrated to *ca.* 4.0 mL and placed under high vacuum. Subsequently, the residue was dissolved in MeOH (50 mL) and evaporated to yield a white foam which was left under high vacuum overnight and was of sufficient purity to be carried forward to the next step. Crude yield: 1.22 g (quant.). Spectroscopic data matched that of the literature compound.[56] ^{1}H NMR (DMSO-d^6) δ: 2.03 (s, 3H), 2.06 (s, 3H), 2.42 (dd, 2H, 2*J* = 2.4, 3*J* = 6.4), 4.22 (m, 1H), 5.18 (m, 1H), 6.08 (t, 1H, 3*J* = 6.8), 6.13 (d, 2H, 3*J* = 8.0), 7.95 (d, 2H, 3*J* = 7.6), 8.51 (s, 1H, N*H*), 9.56 (s, 1H, N*H*).

Synthesis of 5-iodo-3', 5'-diacetyl-2'-deoxycytidine

To a stirred solution of 3',5'-diacetyl-2'-deoxycytidine (3.42 g, 11.0 mmol) in H_2O (17.0 mL) was added I_2 (1.65 g, 6.50 mmol), HIO_3 (0.48 g, 2.73 mmol), CCl_4 (17 mL) and AcOH (25 mL). The resultant mixture was heated to 40°C and stirred vigourously for 24 h. Solvent was removed *in vacuo* to yield an orange solid. Crude product was purified by FCC (EtOAc:Acetone 1:1 – 1:2) to yield 4.14 g (86%) of an off-white foam. Spectral data conformed to literature precedent.[43] ^{1}H NMR (DMSO-*d*6) δ: 2.06 (s, 3H), 2.09 (s, 3H), 3.65 (m, 1H), 4.20 (m, 2H), 4.25 (m, 2H), 5.17 (m, 1H), 6.10 (t, 1H, 3*J* = 4.0), 7.24 (br s, 1H,), 8.01 (s, 1H), 8.16 (br s, 1H,).

Synthesis of 5-trimethylsilylethynyl-3',5'-diacetyl-2'-deoxycytidine

Solid 5-iodo-3',5'-diacetyl-2'-deoxycytidine (1.20 g, 2.75 mmol), and $Pd(Ph_3)_4$ were dissolved in dry deoxygenated THF (8mL) under N_2 to which were added solid CuI (0.23 g, 1.21 mmol), trimethylsilylethyne (1.07 mL, 8.25 mmol) and Et_3N (1.15 mL, 8.25mmol). The solution was stirred in the dark for 24 h at rt. After removal of solvent *in vacuo* the product was purified by FCC (EtOAc:Acetone 2:1 – 1:1) to yield 997 mg (89%) of an off-white solid. ^{1}H NMR ($CDCl_3$) δ: 0.15 (s, 9H), 2.00 (s, 3H), 2.04 (s, 3H), 2.32 (m, 2H), 4.22 (overlapping, m, 3H), 5.13 (m, 1H), 6.31 (t, 1H, 3*J* = 7.0), 6.75 (br s, 1H), 7.86 (br s, 1H), 7.91 (s, 1H). ^{13}C NMR (DMSO-*d6*) δ: -0.5, 8.4, 20.3, 45.6, 63.2, 73.6, 81.3, 85.3, 90.0, 96.1, 99.6, 144.6, 152.9, 163.5, 169.6, 169.7. HRMS (ESI): Calcd. for $C_{18}H_{25}N_3O_6Si$ 407.1513 Found 407.1520.

Synthesis of 5-ethynyl-2'-deoxycytidine

To a stirred mixture of K_2CO_3 (0.65 g, 4.72 mmol) and MeOH (15 mL)**4** was added 5-trimethylethynyl-3',5'-diacetyl-2'-deoxycytidine (0.55 g, 1.34 mmol). The mixture was stirred at 0°C for 2 h then filtered through Celite. The Celite plug was washed with MeOH (3 × 3 mL) and the filtrate was concentrated and purified by flash column chromatography (MeOH:EtOAc 20:80 Rf=0.2) to yield 306 mg (91%) of an off white foam. ^{1}H NMR (D_2O) δ: 2.21 (m, 1H), 2.38 (m, 1H), 3.46 (s, 1H), 3.69 (dd, 1H, 2*J* = 3.6, 3*J* = 12.4), 3.78 (dd, 1H, 2*J* = 5.2, 3*J* = 12.4), 3.97 (m, 1H), 4.33 (m, 1H), 6.13 (t, 1H, 3*J* = 4.0), 8.14 (s, 1H). ^{13}C NMR (DMSO-*d6*) δ: 40.4, 61.3, 70.5, 81.3, 82.9, 89.1, 94.7, 110.3, 142.4, 155.1, 158.9. HRMS (ESI): Calcd. for $C_{11}H_{13}N_3O_4$ 251.0906 Found 251.0911.

General Procedure for the Huisgen cycloaddition

To a 1:1 THF/H_2O stirred solution of $CuSO_4$ (0.1 eq) and sodium ascorbate (0.2 eq) was added 5-ethynyl-2'-deoxycytidine (1 eq) and an organic azide (1.1 eq) at 0 °C. After these additions the solution was brought to rt and stirred until total consumption of 5-ethynyldeoxycytine as determined by TLC analysis. After evaporation of solvent, the residue was dissolved in THF, adsorbed onto silica gel, and purified by FCC.

Synthesis of 2'-deoxy-5-(1-(phenyl)-1*H*-1,2,3-triazol-4-yl)cytidine

The title compound was synthesised by following the general procedure. Thus, to a stirred 1:1 THF/H_2O solution (5 mL) was added 5-ethynyl 2'-deoxycytidine (250 mg, 0.99 mmol), azidobenzene (129 mg, 1.09 mmol), $CuSO_4$ 5 H_2O (22.4 mg, 0.09 mmol) and sodium ascorbate (39 mg, 0.18 mmol) yielding 335 mg (74%) of a light yellow powder after FCC. This material was recrystallized from 1:1 MeOH/H_2O to yield tan coloured needles suitable for X-ray diffraction studies. ^{1}H NMR (DMSO-*d6*) δ: 2.10 (m, 1H), 2.22 (m, 1H), 3.60 (m, 1H), 3.67 (m, 1H), 3.81 (m, 1H), 4.25 (m, 1H), 5.18 (dd, 1H, *J* = 4.78, 4.96), 5.24 (d, 1H, 3*J* = 4.27), 6.18 (t, 1H, 3*J* = 6.32), 7.52 (t, 1H, 3*J* = 7.35), 7.63 (overlapping, dd, 3*J* = 8.03, 7.35), 7.67 (overlapping, br s, 1H), 7.83 (br s, 1H,), 7.91(d, 2H, *J* = 8.03), 8.48 (s, 1H), 8.85 (s, 1H). ^{13}C NMR (DMSO-*d6*) δ: 40.8, 48.7, 60.8, 69.7, 85.5, 87.4, 119.4, 120.5, 129.1, 129.9, 136.4, 140.4, 153.8, 162.6. HRMS (ESI): Calcd. for $C_{17}H_{18}N_6O_4$ [+Na+] 393.1277 Found 393.1287.

Synthesis of 2'-deoxy-5-(1-(thiophen-3-yl)-1*H*-1,2,3-triazol-4-yl)cytidine

Solid $CuSO_4 \cdot 5\ H_2O$ (35 mg, 0.14 mmol) and sodium ascorbate (84 mg, 0.43 mmol) was added to 5-ethynyl-2'-deoxycytidine (214 mg, 0.85 mmol) dissolved in THF:H_2O (1:1, 5 mL). Liquid 3-azidothiophene (165 mg, 1.70 mmol) was added to the mixture which was stirred for 24 h at 35°C after which time the solvent was removed. Purification by flash column chromatography (TLC MeOH:DCM 1:1 Rf=0.8) gave the product as an off-white solid. Yield: 140 mg (43%). ^{1}H NMR (DMSO-*d6*) δ: 2.11 (m, 1H), 2.20 (m, 1H), 3.64 (d, 2H, 3*J* = 16.0), 3.82 (m, 1H), 4.26 (m, 1H), 5.19 (m, 1H), 5.25 (m, 1H), 6.18 (s, 1H), 7.59 (br s, 1H), 7.64 (s, 1H), 7.81 (br s, 1H), 7.84 (s, 1H), 8.06 (s, 1H), 8.46 (s, 1H), 8.77 (s, 1H). 13C NMR (DMSO-*d*6) δ: 48.6, 60.8, 69.7, 85.4, 87.3, 96.2, 115.5, 119.8, 120.8, 119.8, 120.8, 128.6, 135.1, 140.3, 142.1, 153.8, 162.5. HRMS (ESI): Calcd. for $C_{17}H_{16}N_6O_4$ [+H+] 377.1041 Found 377.1032.

2'-Deoxy-5-(1-(2,2'-bithiophen-3-yl)-1*H*-1,2,3-triazol-4-yl)cytidine

Solid $CuSO_4$ 5 H_2O (58.2 mg, 0.23 mmol), sodium ascorbate (142 mg, 0.72 mmol), 5-ethynyl-2'-deoxycytidine (360 mg, 1.43 mmol), and 4-azido-2,2'-bithiophene (594 mg, 2.87 mmol) were combined and stirred at r.t. for 48 h in a solution of THF:H_2O (1:1, 10 mL). The solvent was removed and the crude product was purified by flash column

chromatography to give the title compound as a light brown solid. Yield: 100 mg (15%). ^{1}H NMR (DMSO-*d*6) δ: 2.09 (m, 1H), 2.22 (m, 1H), 3.60 (d, 1H, 3*J* = 12.5), 3.68 (d, 1H, 3*J* = 11.8), 3.83 (dt, 1H, 3*J* = 3.6, 7.0), 4.26 (m, 1H), 5.18 (s, 1H), 5.26 (s, 1H), 6.18 (t, 1H, 3*J* = 6.3), 7.15 (dd, 1H, 3*J* = 3.4, 5.1), 7.47 (dd, 1H, 3*J* = 3.6, 4*J* = 1.0), 7.63 (dd, 1H, 3*J* = 5.1, 4*J* = 1.0), 7.84 (d, 1H, 4*J* = 1.5), 8.00 (d, 1H, 4*J* = 1.4), 8.45 (s, 1H), 8.83 (s, 1H). 13C NMR (DMSO-*d*6) δ: 40.8, 60.8, 69.7, 85.4, 87.4, 96.2, 114.0, 116.9, 119.8, 125.2, 126.7, 128.6, 135.0, 135.2, 138.2, 140.4, 142.1, 153.9, 162.4. HRMS (ESI): Calcd. for $C_{19}H_{18}N_6O_4S_2$ [+Na+] 481.0732 Found 481.0729.

Synthesis of 2'-deoxy-5-(1-(9*H*-fluoren-9-on-2-yl)-1*H*-1,2,3-triazol-4-yl)cytidine

Solid $CuSO_4 \cdot 5H_2O$ (39.6 mg, 0.16 mmol), sodium ascorbate (63.4 mg, 0.32 mmol), 5-ethynyl-2'-deoxycytidine (400 mg, 1.59 mmol), and 2-azido-9H-fluorenone (387 mg, 1.75 mmol) (prepared from 2-aminofluorenone *via* a literature procedure)[57] were combined and stirred at r.t. for 48 h in a solution of THF:H2O (1:1, 10 mL). 1H NMR (600MHz, DMSO-*d*6) δ: 2.11 (m, 1 H), 2.23 (m, 1 H), 3.63 (m, 1 H), 3.67 (m, 1 H), 3.84 (m, 1 H), 4.27 (m, 1 H), 5.14 (t, 1 H, 3*J* = 5.0), 5.26 (d, 1 H, 3 *J* = 4.1), 6.21 (t, 1 H, 3 *J*= 6.4), 7.46 (t, 1 H, 3*J* = 7.6), 7.54 (br s, 1 H), 7.70 (m, 2 H), 7.77 (br s, 1 H), 7.93 (d, 1 H, 3 *J* = 7.0), 8.09 (d, 1 H, 3 *J* = 8.2), 8.13 (s, 1 H), 8.21 (d, 1 H, 3 *J* = 8.2), 8.45 (s, 1 H), 9.04 (s, 1 H). 13C NMR (DMSO-*d*6) δ: 41.1, 61.4, 70.4, 85.8, 87.8, 116.2, 117.7, 120.2, 122.2, 124.8, 125.7, 127.0, 130.4, 134.0, 135.2, 136.3, 137.5, 140.9, 144.2, 145.1, 162.9, 192.3. HRMS (ESI): Calcd. for C24H22N6O4 [+Na+] 495.1393 Found 495.1392.

X-Ray Crystallographic Data

Selected crystal data for **1**·$2H_2O$, $C_{15}H_{20}N_6O_6S$, M = 412.23, orthorhombic, $P2_12_12_1$, a = 4.8637(6) Å, b = 12.1128(18) Å, c = 30.967(5) Å, α = β = γ = 90°, U = 1814.2(5) Å^3, Z = 4, T = 150(2) K, D_c = 1.510 Mg/m^3, μ = 0.227 mm^{-1}, F(000) = 864, GoF = 1.086, R_1 and wR_2 were 0.0752 and 0.1196, respectively, for 279 parameters and 6165 reflections [I > 2σ(I)]. Crystal data have been deposited with the CCDC under reference number 751468.

Photophysical Data, Quantum Yields

Method

Photoluminescence quantum yields (Φ) were found by the relative method[44-46] using anthracene (Φ_F=0.32) and quinine sulphate in 0.1M H_2SO_4 (Φ=0.55) as a reference standard. The quantum yield of the unknown Φ(x) can be calculated by the following equation:

$$\Phi(x)= \Phi(ST)\ (AST/AX)\ (FX/FST)\ (\eta 2X/\eta 2ST)$$

Where Φ(ST) is the quantum yield of the standard, A is the absorbance at the excitation wavelength, F is the integrated area in the emission curve, the subscripts X and ST refer to unknown and standard and η is the refractive index of the solvent. When measuring a series of diluted solutions with various absorbance readings the following equation may be used:

$$\Phi(x)= \Phi(ST)\ (GradX/GradST)\ (\eta 2X/\eta 2ST)$$

Where Grad is the gradient from the plot of integrated area in the emission curve versus absorbance at the excitation wavelength. Prior to measuring the quantum yield of the unknown samples, the validity of the methodology was confirmed by measuring these characterized compounds: anthracene (Φ_F=0.29, in ethanol),[58] dichloroanthracene (Φ_F=0.58, in ethanol)8 and deuterated anthracene (Φ_F=0.32, in cyclohexane),[59] and gave the following values 0.27, 0.58, 0.34 that are in very good agreement with the literature values.

5.5 References

1 Okamoto, A.; Saito Y.; Saito, I. *J. Photochem. Photobiol.* **2005**, *6*, 108-122.

2 Okamoto, A.; Tainaka, K.; Unzai T.; Saito, I. *Tetrahedron* **2007**, *63*, 3465-3470.

3 Wilson, J.N.; Kool, E.T. *Org. Biomol. Chem.* **2006**, *4*, 4265-4274.

4 Millar, D.P. *Curr. Opin. Struct. Biol*. **1996**, *6*, 322-326.

5 Silverman, A.P.; Kool, E.T. *Chem. Rev*. **2006**, *106*, 3775-3789.

6 Sinkeldam, R.W., Greco, N.J., Tor, Y. *Chem. Rev.* **2010**, *110*, 2579-2619.

7 Dodd, D.W.; Hudson, R.H.E. *Mini-Rev. Org. Chem.* **2009**, *6*, 378-391.

8 Kumar, T.; El-Sagheer, A.; Tumpane, J.; Lincoln, P.; Wilhelmsson, L.M.; Brown, T. *J. Am. Chem. Soc*. **2007**, *129*, 6859-6864.

9 Secrist, J.A.; Barrio, J.R.; Leonard, N.J. *Biochem. Biophys. Res. Commun*. **1971**, *45*, 1262-1270.

10 Gierlich, J.; Burley, G.A.; Gramlich, P.M.E.; Hammond, D.M.; Carell, T. *Org. Lett*. **2007**, *8*, 3639-3642.

11 Waggoner, A. *Method. Enzymol*. **1995**, *246*, 362-373.

12 Fleet, G.W.J.; Porter, R.R.; Knowles, J.R. *Nature* **1969**, *224*, 511-512

13 Staudinger, H.; Meyer, J. *Helv. Chim. Acta* **1919**, *2*, 635-646.

14 Saxon, E.; Bertozzi, C.R. *Science* **2000**, *287*, 2007-2010.

15 Weisbrod, S.H.; Marx, A. *Chem. Commun*. **2008**, 5675-5685.

16 Tornoe, C.W.; Christensen C.; Meldal, M. *J. Org. Chem.* **2002**, *67*, 3057-3064.

17 Rostovtsev, V.V.; Freen, L.G., Fokin, V.V.; Sharpless, K.B. *Angew. Chem.* **2002**, *114*, 2708; *Angew. Chem. Int. Ed*. **2002**, *41*, 2596-2598.

18 Tron, G.C.; Pirali, T.; Billington, R.A.; Canonico, P.L.; Sorba, G.; Genazzani, A.A. *Med. Res. Rev*. **2008**, *28*, 278-308.

19 Lutz, J.F.; Zarafshani Z. *Adv. Drug Deliv. Rev*. **2008**, *60*, 958-970.

20 Hawker, C.J.; Wooley, K.L. *Science* **2005**, *309*, 1200-1205.

21 Kele, P.; Mezö, G.; Achatz, D.; Wolfbeis, O.S. *Angew. Chem. Int. Ed*. **2009**, *48*, 344.

22 Gramlich, P.M.E.; Wirges, C.T.; Manetto, A.; Carell, T.; *Angew. Chem. Int. Ed*. **2008**, *47*, 8350-8358.

23 Cartwright, I.L.; Hutchinson, D.W.; Armstrong, V.W. *Nucleic Acids Res*. **1976**, *3*, 2331-2339.

24 Kocalka, P.; Andersen, N.K.; Jensen, F.; Nielsen, P. *ChemBioChem* **2007**, *17*, 2106-2116.

25 Seela, F.; Sirivolu V.R.; Chittepu, P. *Bioconjugate Chem.* **2008**, *19*, 211-224.

26 Park, S.M.; Lee Y.S.; Kim, B.H. *Chem. Commun.* **2003**, 2912-2913.

27 Park, S.M.; Shen Y.; Kim, B.H. *Org. Biomol. Chem.* **2007**, *5*, 610-612.

28 Gierlich, J.; Gutsmiedl, K.; Gramlich, P.M.E.; Schmidt, A.; Burley, G.A.; Carell, T. *Chem. Eur. J.* **2007**, *13*, 9486-9494.

29 Seela, F.; Sirivolu, V.R. *Org. Biomol. Chem*. **2008**, *6*, 1674-1687.

30 Berndl, S.; Herzig, N.; Kele, P. Lachmann, D.; Li, X.H.; Wolfbeis, O.S.; Wagenknecht, H.-A. *Bioconjugate Chem*. **2009**, *20*, 558-564.

31 Alkynyl monomers and synthetic services by: Baseclick GmdH, Tutzing, Germany. http://www.baseclick.eu.

32 El-Sagheer, A.H.; Brown, T. *Chem. Soc. Rev*. **2010**, *39*, 1388-1405.

33 Kanan, M.W.; Rozenman, M.M.; Sakurai, K.; Snyder, T.M.; Liu, D.R. *Nature* **2004**, *431*, 545-549.

34 Sletten, E.M.; Bertozzi, C.R.; *Angew. Chem. Int. Ed*. **2009**, *48*, 6974-6998.

35 Werner, A.; Buss, H. *Berichte* **1894**, *27*, 2193-2201.

36 Huisgen, R. *Naturwiss. Rundschau* **1961**, *14*, 63.

37 Jäger, V.; Colinas, P. A. Synthetic Applications of 1,3-Dipolar Cycloaddition Chemistry Toward Heterocycles and Natural Products; Padwa, A.,Pearson, W. H., Eds.; John Wiley and Sons: New Jersey, NJ, 2003; p 361.

38 Huisgen, R.; Mack, W. *Tetrahedron Lett*. **1961**, 583-589.

39 Mukaiyama, T.; Hoshino, T. *J. Am. Chem. Soc*. **1960**, *82*, 5339-5342.

40 Singh, I., Vyle, J.S., Heaney, F. *Chem. Commun.*, **2009**, 3276-3278.

41 Singh, I., Heaney, F., *Org. Biomol. Chem.*, **2010**, *8*, 451-456.9,10

42 Matsuda, A.; Shinozaki, M.; Suzuki, M.; Watanabe, K.; Miyasaka, T. *Synthesis,* **1986**, 385-386

43 Chang, P.; Welch, A.D. *J. Med. Chem.* **1963**, *4*, 428-431.

44 Williams, A.T.R.; Winfield, S. A. *Analyst* **1983**, *108*, 1067.

45 Lavabre, D.; Fery-Forgues, S. *J. Chem. Educ.* **1999**, ***76***, 1260.

46 Morris, J. V.; Mahaney, M. A.; Huber, J.R. *J. Phys. Chem.* **1976**, *80*, 969-972.

47 Marti, A.A.; Jockusch, S.; Li, Z.M.; Ju, J.Y.; Turro, N.J. *Nucleic Acids Res*. **2006**, *34*, 50

48 Zang, H.; Fang, Q.; Pegg, A.E.; Guengerich, F.P. *J. Biol. Chem*. **2005**, *280*, 30873.

49 Berry, D.A.; Jung, K.Y.; Wise, D.S.; Sercel, A.D.; Pearson, W.H.; Mackie, H.; Randolph, J.B.; Somers, R.L. *Tetrahedron Lett.* **2004**, *45*, 2457; (*d*)

50 Liu C.H.; Martin, C.T. *J. Mol. Biol*. **2001**, *308*, 465.

51 Józefowicz, M. *Spectrochim. Acta A-M* **2008**, *71*, 537.

52 Reichardt, C. *Chem. Rev.* **1994**, *94*, 2319.

53 Tong, G.; Lawlor, J.M.; Tregear, G.W.; Haralambidis, J. *J. Am. Chem. Soc.* **1995**, *117*, 12151 –12158.

54 Paris, P.L.; Langenhan, J.M.; Kool, E.T. *Nucleic Acids Res.* **1998**, *26*, 3789–3793.

55 Conlon, P.; Yang, C.Y.J.; Wu, Y.R.; Chen, Y.; Martinez, K.; Kim, Y.M.; Stevens, N.; Marti, A.A.; Jockusch, S.; Turro, N.J.; Tan, W.H. *J. Am. Chem. Soc.* **2008**, *130*, 336–342.

56 Piton, N.; Mu, Y.; Stock, G.; Prisner, T. F.; Schiemann, O.; Engels, J. W. *Nucleic Acids Res.* **2007**, *35*, 3128-3143.

57 Nimura, S.; Kikuchi, O.; Ohana, T.; Yabe, A.; Kondo, S.; Kaise, M. *J. Phys. Chem. A* **1997**, *101*, 2083-2088.

58 Guilbault, G. G. Practical Fluorescence, Second Edition, Marcel Dekker, Inc., New York, **1990**. p. 14-16.

59 Guilbault, G. G. Practical Fluorescence, Second Edition, Marcel Dekker, Inc., New York, **1990**. p. 14-16. 9 Berlman, I. B. Handbook of Fluorescence Spectra of Aromatic Molecules, Second Edition, Academic Press Inc., New York, **1971**, p. 357.

Chapter 6

6 Hydrogelation abilities of nucleobase-modified cytidines possessing substituted triazoles

Nucleoside-derived hydrogelators have been sought for their potential biomedical applications, such as are found in tissue engineering and drug delivery. By judiciously adding a degree of hydrophobicity certain analogues are able to form micelles, bi-layers and gels in water. Research in this area has yet to lay down solid ground rules for the rational design of novel nucleoside gelators making further studies necessary. The synthesis and examination of a series of aryl-substituted 5-triazolylcytidines yielded an analogue that gelates water. 5-(1-(2,2'-bithiophen-3-yl)-1*H*-1,2,3-triazol-4-yl)-2'-deoxycytidine was found to form gels in water down to 0.3 wt%. The ribocytidine analogue failed to form gel in aqueous solution; but was able to form a hydrogel in the presence of guanosine. Images obtained by SEM show the different architectures of the gel; varying from cribriform, to fibrous, to lamellar. The present gelating compound studied may have potential as a component of a controlled-release drug delivery system. Excerpts of this chapter have been taken from Dodd, D.W.; Hudson, R.H.E. *Artificial DNA*, **2010**, *1*, 90-95.

6.1 Introduction

Polymeric hydrogels have been used as medical implants and show promise as drug delivery systems. Due to their high water content, the gels are typically biocompatible and can be designed to respond to certain external parameters such as, pH, electric field and temperature.[1] While polymeric hydrogels based on polyacrylates and peptides remain the most prevalent, interest in small-molecule gelators has continued to grow.[2-4] In the case of low molecular weight gelators, interactions which hold together the framework and trap water within interstitial spaces are necessarily non-covalent. This generally makes the gel more prone to dissociation which is not necessarily deleterious for applications in biomedical engineering. Scissile, non-covalent gelling interactions are

even desirable for the application of controlled drug release as the scaffold may break down easily and diffuse after outliving its primary function. Furthermore, the pharmacophore can be incorporated into the gelator itself, forgoing the need to capture a drug within the gel.[5]

Necessary to their biological functions, nucleobases form interactions in perpendicular directions. Hydrogen bonding in the plane of the heterocycle coupled with perpendicular π-π stacking hold together higher-order nucleic acid structures found in nature and also make nucleobases a suitable scaffold for the development of hydrogelators. Indeed, modified nucleosides have been found to gelate water and this, taken together with the possibility of nucleoside analogues being pharmacologically relevant as anti-tumour and/or anti-viral drugs, makes them attractive targets. For many years it has been known that certain guanosine derivatives form hydrogels at high concentration in a K^+ dependent manner.[6] From this starting point, a number of nucleoside-lipid conjugates have been investigated and found to gel water with varying success.[7] Attachment of hydrophobic moieties *via* the sugar are common; however, a rarer strategy employed by Kim and co-workers aimed at derivatizing the nucleobase at the 5-position of deoxy- and ribouridine.[8,9] Huisgen cycloaddition of an alkyne functionalized nucleobase and an azide provides a modular means of accessing a variety of analogues and, reasoning that aryl moieties would increase the potential for stacking interactions, they synthesized a series of alkylbenzyltriazole-appended uridines. Some of the nucleosides were found to form stable gels down to the low concentration of 0.2 wt%, with the deoxy-series being more effective than the corresponding ribonucleosides. Nielsen and coworkers have previously reported the synthesis of 5-triazolylydeoxycytidines, however, the study lacks characterization data and the resultant compounds were not investigated for gelation properties.[10] As a postlude to our work with triazolylcytidines as potential base-discriminating fluorophores,[11] and the recognition of the paucity of examples of cytidine-based hydrogelators,[7,12] we have investigated the gelation properties of 1*H*-substituted-5-triazolyl(deoxy)cytidines (Fig. 6-1).

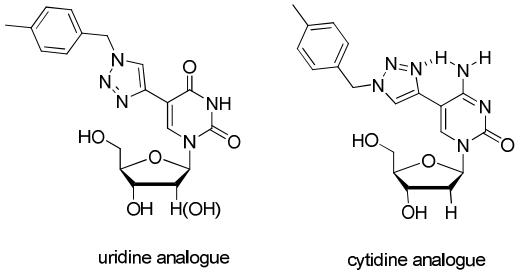


Figure 6-1: Kim's alkylbenzyltriazole-appended uridine compared to the same triazoly(deoxy)-cytidine.

6.2 Results and Discussion

The nucleoside derivatives were accessed *via* the sugar-protected 5-ethynylcytidine as a key intermediate. The ethynyl nucleoside was obtained in good yield by Sonogashira coupling between trimethylsilylacetylene and the acetylated iodo-nucleoside, followed by desilylation with TBAF. Huisgen cycloaddition of the ethynyl nucleoside with a variety of alkyl- and arylazides afforded sugar-protected conjugates which gave the final compounds after methanolysis to remove the acetate groups. The synthetic approach used to access compound **4b** is shown in figure 6-2.

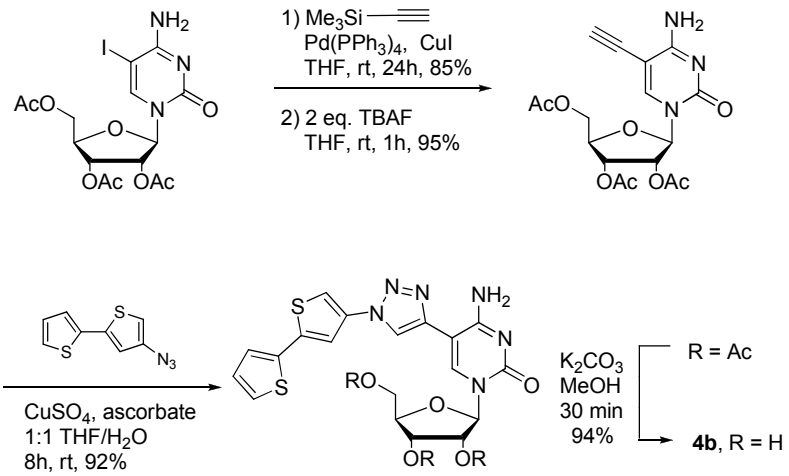


Figure 6-2: Synthesis of bithiophenyl-triazolylcytidine 4b using Sonogashira coupling and the Huisgen cycloaddition

The ribonucleoside analogues shown in figure 6-3 were synthesized by this route in good yields, while the deoxy-series were synthesized by a previously reported procedure.[11]

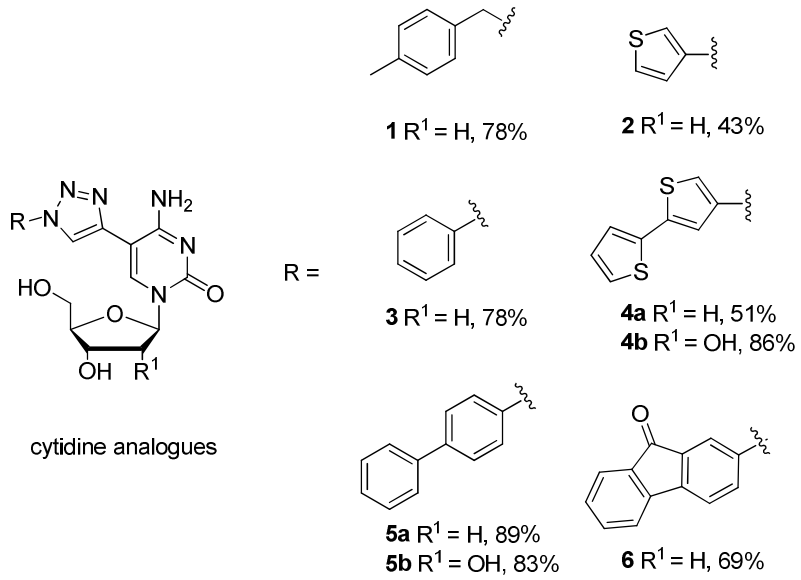


Figure 6-3: 5-Triazolycytidines prepared and examined for the present work.

Examination of their properties revealed that compound **1** unfortunately did not form a gel with water as the corresponding uridine analogue did. Since this structural modification did not translate from deoxyuridine to deoxycytidine, compounds with different hydrophobic substituents were investigated. Compounds with slightly smaller hydrophobic groups such as **2** (thiophene) and **3** (phenyl) tended toward forming

crystalline solids. Although both solid-state structures (**2** from ref. 11, **3** see chapter 5) showed inclusion of a large number of solvent molecules (water or methanol), no guiding structural principles were discerned. It was posited that the main driving force for crystal formation was *H*-bonding due to the extent of *H*-bonding displayed in the solid state. Thus, in an attempt to offset the amount of *H*-bonding relative to hydrophobic interactions, we prepared compounds **4**-**6** with larger hydrophobic substituents. Happily, we found that the bithiophenyl compound **4a** formed a gel during an attempted recrystallization from water. Guided by this observation we prepared compounds with structurally similar aryl moieties varying the nature of the ring, size, and rigidity/conformation such as biphenyl (**5a,b**) and fluorenonyl (**6**),[11] but none of these compounds possessed the ability to form hydrogels. As we continued to synthesize aryl derivatives, it was becoming increasingly evident that the resultant compounds were predominantly crystalline and **4a** was the only compound which gelled water *per se*. Dissolution of **4a** in water under heating followed by allowing the solution to come to rt resulted in an opaque gel down to a minimum gelling concentration (MGC) of 0.3 wt. %, persisting in that state for 1-2 h at ambient temperature. More robust, albeit more opaque, gel (persisting for *ca.* 24 h) formed from solutions of higher concentration (> 3 wt %). After these findings, the ribo-analogue **4b** was pursued.

In contrast to **4a**, the ribo-analogue **4b** would not form gel at any concentration, the upper limit being determined by solubility. However, in the case of a solution containing **4b** and its base pairing partner guanosine in a 1:1 ratio, gelation occurred down to a MGC of 0.5 wt % compound **4b**. The introduction of guanosine into a solution of **4a** had no significant effect on the MGC (gels formed down to 0.4 wt%) and it is important to note that guanosine alone does not form a hydrogel at the concentration tested.

Scanning electron microscopy of desiccated hydrogels (xerogels) and the non-gelating **4b** revealed interesting supramolecular architectures. In the case of **4b**, where gelation did not occur, some fibrous areas along with amorphous regions are visible; however, the sample was predominantly crystalline (Fig. 6-4). Images from gel formed from the mixture of **4b** with guanosine are in stark contrast to those of **4b** alone. No crystalline structures remain visible and a porous, cribrifom or sponge-like material dominates.

Lamellar modes of aggregation are observed for **4a**. Ribbons on the order of 1-5 μM in width and < 1 μM in depth are clearly observed running in an organized, parallel arrangement for lengths of greater than 50 μM.

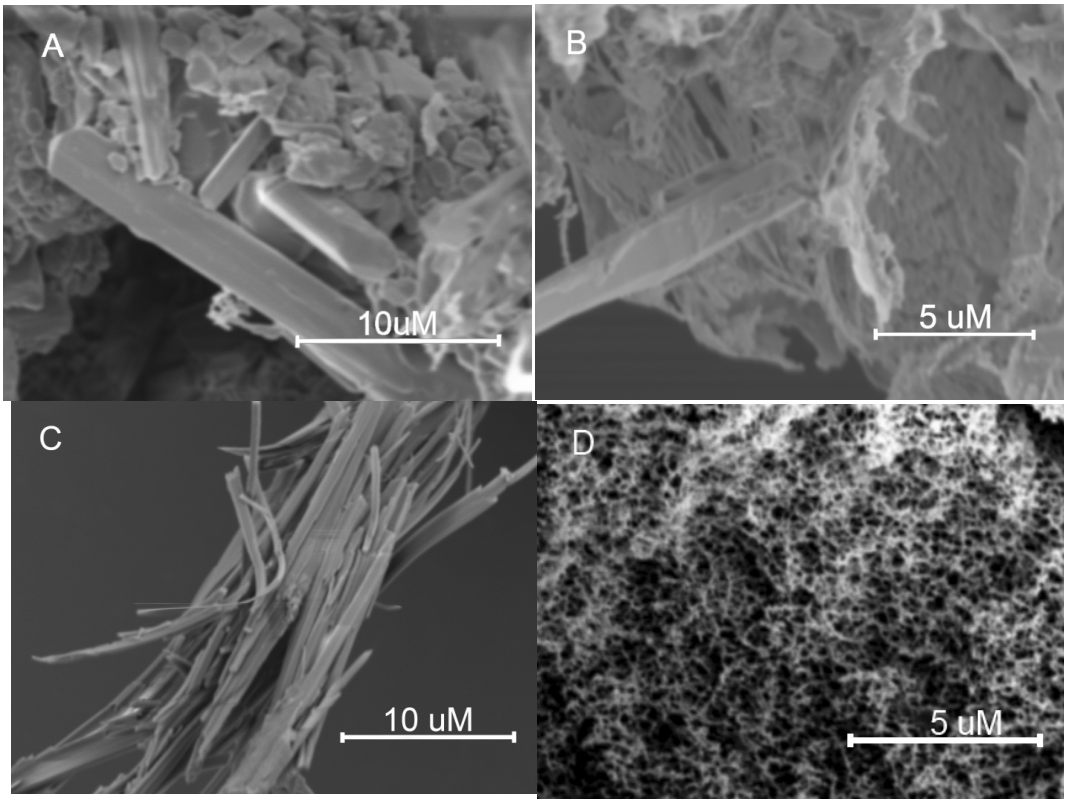


Figure 6-4: Scanning electron micrographs of xerogels or powder of 4b (A, B), 4a derived from a 1 wt% gel (C) and 4b at 1 wt% with one molar equivalent of guanosine (D).

In their analysis of intermolecular bonding of the uridine-based hydrogelators, Kim and co-workers implied the importance of both hydrogen bonding involving the nucleobases and the hydroxyl groups of the (deoxy)ribose. A balance must also be struck in the hydrophobicity/hydrophilicity of the compound to favor hydrogelation over complete solubilization or crystallization. While solid-state structures were not reported for the uridine compounds, *ab initio* calculations[13] have revealed that the energetically preferred conformation was as depicted in figure 6-1, wherein the *C-H* of the triazole ring is cis to

the *C4=O* of the nucleobase in a nearly coplanar arrangement. The *N3* of the triazole is greater than 3Å from the 5'-*OH* group making the likelihood of H-bonding between them improbable in an aqueous environment. A solid state structure has been reported for a cytidine bearing a triazole substituent at *C5* wherein the conformation about the triazole-nucleobase bond is reversed due to the existence of an intramolecular *N···H* bond.[11] The resultant trajectory of the benzyl substituent is consequently different between the two nucleosides and such subtle structural differences likely contribute to their ability to hydrogelate.

Cytidine derivatives with smaller hydrophobic substituents tended to crystallize readily. From the observation of the network of H-bonding in the solid state structures of **2** and **3**, it was decided to increase the hydrophobicity of the compounds by increasing the number of aryl groups. At this point, it was found that the bithiophenyl-triazolyldeoxycytidine formed a hydrogel. To investigate the tolerance to structural modification and possibly discover other gelators as well as potentially reveal some underlying structural principles for hydrogelating competent analogues, two other derivatives were targeted. The biphenyl derivative **5** possesses unfused aromatic groups, but unlike the triazole-thiophene system which is nearly coplanar (in the solid state), many studies indicate that the biphenyl system exhibits significant twist (ca. 40 °).[14] Since this modification led to compounds (**5a,b**) that did not form hydrogels, compound **6** in which the phenyl rings are bridged by a carbonyl group and thus fused in coplanar arrangement was prepared in order to better mimic the bithiophene moiety. This compound also failed to hydrogelate and serves to highlight the capricious nature of this type of supramolecular chemistry.

The previously reported uridine hydrogelators showed that deoxynucleosides outperformed the ribonucleosides for reasons unknown. The effect of an added hydroxyl group was more drastic in our case with our ribonucleoside **4b** not displaying any gelation properties and simply precipitating out as a fine, microcrystalline suspension. Whether the extra hydroxyl group interferes by disadvantageously H-bonding and/or affecting the sugar conformation is not known.

SEM images of a xerogel formed from **4a** show an organized framework of lamellae

running in parallel. Similar to structures obtained by single crystal X-ray diffraction, one must be cautious when interpreting these data as drying of the sample has been known to occasionally introduce artifacts.[15] We suspect that the tendency for **4b** to crystallize interferes with its ability to form a gel. Fibrous networks can be seen (**Fig. 3A,B**), however, crystals predominate. The introduction of guanosine to a solution of **4b** results in gelation. This could be due to guanosine interfering with crystallization or by becoming involved in the molecular assembly responsible for gelation itself. The gel formed in this case is very different from that of **4a** and appears much less organized on the micron scale. It is of interest to note that no significant change was observed in **4a**'s gelation ability on the introduction of guanosine.

Uridine, in interaction with itself, can form slip base pairs as can cytidine; however, hemi-protonated cytosine, as found in the i-motif may form 3 H-bonds, the effect of pH was therefore investigated and acidification was found to hamper the gelation ability of **4a**. A 0.5 wt% gel of compound **4a** was gradually acidified with microlitre titres of 0.1 M HCl. At pH 5.0 the gel became unstable to inversion and at pH 4.2 nothing but a fine precipitate remained. Presumably, this behaviour is derived from the protonation of the *N*3 of cytosine and may allow for development of a pH responsive drug delivery system.

There are inherent benefits of the photophysical and electrochemical properties of the incorporated bithiophene. **4a** is known to be fluorescent and fluorometrically responsive to its environment.[11] This could easily allow for visualization of gel decomposition. Electrochemical reactivity offers the potential for electropolymerisation of **4a** in the gel state and this work is currently under study.

6.3 Conclusions

In conclusion, we have synthesized a variety of nucleoside analogues in a modular fashion from 5-ethynyl(deoxy)cytidine and various azides. This class of molecules contained **4a** which was able to gelate water down to a MGC of 0.3% and may have application in controlled drug release.

6.4 Experimental

General Considerations

Chemicals were obtained from commercial sources unless otherwise noted and used without further purification. Flash column chromatography (FCC) was performed on Silicycle Siliaflash 60, 230-400 mesh and thin layer chromatography (TLC) was performed on Merck Kieselgel F-60 plates. Chemical shifts are reported in parts per million (δ), were measured from tetramethylsilane (0 ppm) and are referenced to the residual proton in the deuterated solvent: $CDCl_3$ (7.26 ppm), DMSO-d_6 (2.48 ppm), D_2O (4.75 ppm) for ^1H NMR and $CDCl_3$ (77.0 ppm), DMSOd_6 (39.5 ppm) for ^{13}C NMR. Multiplicities are described as s (singlet), d (doublet), t (triplet), q (quartet), m (multiplet), br s (broad singlet) and ov (overlapping). Coupling constants (*J*) are reported in Hertz (Hz). Exchangeable protons were identified by their disappearance when the sample was shaken against D_2O. High resolution mass spectra (HRMS) were obtained using electrospray ionization time of flight methods (ESI-TOF). Solution ^1H NMR spectra were collected using a Varian 400 spectrometer (400.09 MHz for ^1H and 100.61 MHz for ^{13}C) at rt.

Scanning Electron Microscopy

Hydrogels were frozen in liquid nitrogen and lyophilized prior to imaging. Coarse powder samples were adhered to the surface and sputtered with a Au/Pd alloy using a Hummer VI apparatus to a thickness of 12-15 nm prior to imaging on a Hitachi S-3400 N scope.

Gelation test

The compound was heated in water at the appropriate concentration until dissolution. After allowing the solution to come to room temperature the gelation ability was determined by the “stable-to-the-inversion-of-the-container” method.[16]

Synthesis

5-Trimethylsilylethynyl-2',3',5',-tri-*O*-acetyl-cytidine

To dry deoxygenated THF (20 mL) was added 2',3',5',-tri-*O*-acetyl-5-iodocytidine (3.00 g, 6.06 mmol), $Pd(PPh_3)_4$ (705 mg, 0.61 mmol) and CuI (349 mg, 1.89 mmol) under an atmosphere of N_2. On dissolution, Et_3N (2.53 mL, 18.2 mmol), and trimethylsilylacetylene (1.09 mL, 18.2 mmol) were added and the solution was stirred under N_2 for 24 h or until deemed complete by TLC. After solvent removal *in vacuo* the product was purified by FCC eluting with a gradient of MeOH in CH_2Cl_2 (0:1 to 1:9). Evaporation of the appropriate fractions yielded 2.04 g (85%) of an off white powder. ^{1}H NMR ($CDCl_3$) δ: 0.21 (s, 9H), 2.05 (s, 3H), 2.10 (s, 3H), 2.19 (s, 3H), 4.36 (ov m, 2H), 4.39 (ov m, 1H), 5.29 (dd, 1H, *J* = 5.2 and 5.3), 5.36 (dd, 1H, *J* = 3.5 and 5.3), 5.84 (br s, 1H), 6.09 (d, 1H, *J* = 3.52), 7.89 (s, 1H), 8.38 (br s, 1H). ^{13}C NMR δ: -0.2, 8.6, 46.1, 20.4, 20.9, 62.2, 69.0, 73.9, 79.1, 88.4, 92.2, 95.2, 101.8, 143.3, 153.7, 164.4, 169.3, 169.5, 170.0. HRMS (ESI): Calcd. for $C_{20}H_{27}N_3O_8Si$ [$+H^+$] 466.1645 Found 466.1579.

5-ethynyl-2',3',5',-tri-*O*-acetyl-cytidine

Solid 5-trimethylsilylethynyl-2',3',5',-tri-*O*-acetyl-cytidine (1.00 g, 2.15 mmol) was dissolved in THF at 0°C and 2 equivalents of TBAF were delivered as a 1.0 M solution in THF (430 μL, 4.30 mmol). After the solution was allowed to come to rt stirring was continued for 1 h at which time the reaction was complete as judged by TLC. Solvent was removed to reveal yellow oil which, after FCC (MeOH:CH_2Cl_2, 0:1 to 1:9), gave 804 mg (95%) of a white powder. ^{1}H NMR ($CDCl_3$) δ: 2.07 (s, 3H), 2.11 (s, 3H), 2.19 (s, 3H), 3.35 (s, 1H), 4.37 (ov m, 2H), 4.38 (ov m, 1H), 5.29 (dd, 1H, *J* = 5.3 and 5.9), 5.37 (dd, 1H, *J* = 3.5 and 5.3), 5.80 (br s, 1H), 6.10 (d, 1H, *J* = 3.5), 7.13 (br s, 1H), 7.96 (s, 1H). ^{13}C NMR δ: 20.4, 20.5, 20.8, 62.3, 69.2, 73.9, 74.9, 79.3, 84.0, 88.5, 90.5, 144.3, 153.7, 164.4, 169.4, 169.5, 170.2. HRMS (EI): Calcd. for $C_{17}H_{19}N_3O_8$ 393.1172 Found 393.1177.

2',3',5',-tri-*O*-acetyl-5-(1-(2,2'-bithiophen-3-yl)-1*H*-1,2,3-triazol-4-yl)cytidine

To a stirred 1:1 THF/H_2O solution (10 mL) was added 5-ethynyl-2',3',5',-tri-*O*-acetylcytidine (250 mg, 0.64 mmol), 4-azido-2,2'-bithiophene[11] (145 mg, 0.70 mmol),

$CuSO_4 \cdot 5\ H_2O$ (22.4 mg, 0.09 mmol) and sodium ascorbate (39 mg, 0.18 mmol). Stirring was continued for 8 h at rt. Solvent was then removed and 353 mg (92%) of a tan coloured powder was obtained after FCC. ^{1}H NMR ($CDCl_3$) δ: 2.08 (s, 6H), 2.10 (s, 3H), 4.32 (dd, 1H, *J* = 12.1 and 2.4), 4.38 (m, 1H), 4.61 (dd, 1H, *J* = 12.1 and 7.0), 5.33 (dd, 1H, *J* = 5.5 and 5.3), 5.47 (dd, 1H, *J* = 5.3 and 4.8), 6.20 (d, 1H, *J* = 4.8), 7.00 (dd, 1H, *J* = 5.1 and 3.7), 7.19 (d, 1H, *J* = 3.7), 7.24 (d, 1H, *J* = 5.1), 7.45 (br s, 1H), 7.54 (ov s, 1H), 7.55 (ov s, 1H), 7.99 (s, 1H), 8.27 (br s, 1H), 8.52 (s, 1H). ^{13}C NMR ($CDCl_3$) δ: 20.4, 20.9, 30.8, 63.3, 70.1, 73.6, 76.7, 77.3, 79.8, 88.4, 98.1, 112.5, 116.3, 118.0, 124.6, 125.5, 128.0, 135.2, 135.8, 138.6, 139.0, 142.3, 154.2, 162.8, 169.6, 171.0. HRMS (ESI): Calcd. for $C_{25}H_{24}N_6O_8S_2$ [$+H^+$] 601.1175 Found 601.1180.

5-(*p*-methylbenzyl)-1*H*-1,2,3-triazol-4-yl)-2'-deoxycytidine **1**

Prepared via literature procedure[11] in 78% yield: ^{1}H NMR (DMSO-d_6) δ: 2.07 (m, 1H), 2.16 (m, 1H), 3.15 (s, 3H), 3.59 (dd, 1H, *J* = 3.4 and 11.9), 3.67 (dd, 1H, *J* = 3.4 and 11.9), 3.78 (dd, 1H, *J* = 3.4 and 7.0), 4.24 (m, 1H), 6.14 (app t, 1H, J = 6.2), 7.17 (d, 1H, J = 8.4), 7.24 (d, 1H, J = 8.4), 7.80 (br s, 1H), 7.92 (br s, 1H), 8.39 (s, 1H), 8.48 (s, 1H). HRMS (ESI): Calcd. for $C_{19}H_{22}N_6O_4$ [$+Na^+$] 421.1595 Found 421.1589.

5-(1-(2,2'-bithiophen-3-yl)-1*H*-1,2,3-triazol-4-yl)cytidine **4b**

To a slurry of K_2CO_3 in MeOH (280 mg, 2.0 mmol in 15 mL) was added 2',3',5',-tri-*O*-acetyl-5-(1-(2,2'-bithiophen-3-yl)-1*H*-1,2,3-triazol-4-yl)cytidine (300 mg, 0.50 mmol) and the mixture was stirred for 30 min or until reaction completion as determined by TLC. After filtration over Celite the filtrate was evaporated and the residue adsorbed onto silica gel and purified by FCC (< 15 cm silica, isocratic CH_2Cl_2:MeOH, 8:2) yielding 223 mg (94%) of a tan coloured powder. ^{1}H NMR (DMSO-d_6) δ: 3.49 (m, 1H), 3.65 (m, 1H), 3.96 (m, 1H), 4.16 (dd, 1H, *J* = 2.15 and 5.28), 5.05 (dd, 1H, *J* = 5.28 and 6.84), 5.81 (d, 1H, *J* = 6.84), 5.52 (br s, 3H), 7.16 (dd, 1H *J* = 3.5 and 5.9), 7.46 (d, 1H, *J* = 3.5), 7.53 (br s, 1H), 7.62 (d, 1H, *J* = 5.9), 7.81 (ov br s, 1H), 7.83 (s, 1H), 7.98 (s, 1H), 8.43 (s, 1H), 8.42 (s, 1H). ^{13}C NMR (DMSO-d_6) δ: 40.8, 60.8, 69.7, 85.4, 87.4, 96.2, 114.0, 116.9, 119.8, 125.2, 126.7, 128.6, 135.0, 135.2, 138.2, 140.4, 142.1, 153.9, 162.4 HRMS

(ESI): Calcd. for $C_{19}H_{18}N_6O_5S_2$ 475.0831 Found 475.0855.

5-(1-(biphenyl-4-yl)-1*H*-1,2,3-triazol-4-yl)-2'-deoxycytidine **5a**

Prepared via literature procedure[11] in 89% yield. ^{1}H NMR (DMSO-d_6) δ: 2.10 (m, 1H), 2.20 (m, 1H), 3.63 (m, 2H), 3.81 (m, 1H), 4.29 (m, 1H), 6.18 (app t, 1H, *J* = 6.8), 7.41 (t, 1H, *J* = 7.3), 7.50 (dd, 2H, *J* = 7.0 and 7.3), 7.64 (br s, 1H), 7.74 (d, 2H, *J* = 7.5), 7.85 (ov br s, 1H), 7.93 (d, 2H, *J* = 7.9), 8.01 (d, 2H, *J* = 7.9), 8.49 (s, 1H), 8.91 (s, 1H). ^{13}C{^{1}H} NMR (DMSO-d_6) δ: 40.8, 60.8, 69.7, 85.4, 87.4, 96.3, 119.3, 120.8, 126.79, 120.9, 129.1, 135.6, 138.7, 140.4, 140.7, 142.7, 153.9, 162.4. LRMS (ESI): Calcd. for $C_{23}H_{22}N_6O_4$ [+H^+] 447.5 Found 447.2., Calcd. for $C_{23}H_{22}N_6O_4$ [+Na^+] 469.4 Found 469.2.

5-(1-(biphenyl-4-yl)-1*H*-1,2,3-triazol-4-yl)cytidine **5b**

Prepared via literature procedure[11] in 83% yield. ^{1}H NMR (DMSO-d_6) δ: 3.52 (m, 1H), 3.68 (m, 1H), 3.99 (m, 1H), 4.21 (dd, 1H, *J* = 2.1 and 5.3), 5.05 (dd, 1H, *J* = 5.3 and 6.7), 5.91 (d, 1H, *J* = 6.7), 7.41 (t, 1H, *J* = 7.3), 7.50 (dd, 2H, *J* = 7.1. and 7.4), 7.68 (br s, 1H), 7.73 (d, 2H, *J* = 7.1), 7.89 (ov br s, 1H), 8.01 (d, 2H, *J* = 7.8), 8.2 (d, 2H, *J* = 7.8), 8.58 (s, 1H), 9.01 (s, 1H). HRMS (ESI): Calcd. for $C_{23}H_{22}N_6O_5$ [+H^+] 463.1724 Found 463.1791.

6.5 References

1. Kopeček, J. *J. Polym. Sci. A1* **2009**, *47*, 5929-46.
2. Tanaka, T. In Polyelectrolyte Gels: Properties, Preparation, and Applications Harland RS, Prud'homme RK. Eds. Washington D.C. **1992**, *480*, 1-21.
3. Bieser, A.M.; Tiller, J.C. *J. Phys. Chem. B*. **2007**, *111*, 13180-13187.
4. Goyal, N.; Cheuk, S.; Wang, G. *Tetrahedron* **2010**, *66*, 5962-5971.
5. Xing, B.; Yu, C.W.; Chow, K.H.; Ho, P.L.; Fu. D.; Xu, B. *J. Am. Chem. Soc.* **2002**, *124*, 10957.
6. Guschlbauer, M.; Chantot, J.F.; Thiele, D. *J. Biomol. Struct. Dyn.* **1990**, *8*, 491-511.
7. Araki, K.; Yoshikawa, I. *Top. Curr. Chem.* **2005**, *256*, 133-65.

8. Park, S.M.; Shen, Y.; Kim, B.H. *Org. Biomol. Chem.* **2007**, *5*, 610-2.

9. Park, S.M.; Lee, Y.S.; Kim, B.H. *Chem. Commun*. **2003**, 2912-3.

10. Andersen, N.K.; Spáčilová, L.; Jensen, M.D.; Kočalka, P.; Jensen, F.; Nielsen, P. *Nucl. Acid. S.* **2008**, *52*, 149-150.

11. Dodd, D.W.; Swanick, K.N.; Price, J.T.; Brazeau, A.L.; Ferguson, M.J.; Jones, N.D.; Hudson, R.H.E. *Org. Biomol. Chem.* **2010**, *8*, 663-666.

12. Venkatesan, N.; Seo, Y.J.; Bang, E.K.; Park, S.M.; Lee, Y.S.; Kim, B.Y. *Bull. Korean. Chem. Soc.* **2006**, *27*, 613-630.

13. Kočalka, P.; Andersen, N.K.; Jensen, F.; Nielsen, P. *ChemBioChem* **2007**, *8*, 2106-16.

14. Grein, F. *J. Mol. Struct: THEOCHEM* **2003**, *624*, 23-28.

15. Kölbel, M.; Menger, F.M. *Chem. Commun.* **2001**, 275.

16. Menger, F.M.; Caran, K.L. *J. Am. Chem. Soc.* **2000**, *122*, 11679-11691.

Chapter 7

7 Hoogsteen recognition by pyrrolocytosines in DNA and PNA

Fluorophores acting as base surrogates that respond fluorometrically to nucleic acid structural events have found uses across the biological sciences. However, the problem of the design and synthesis of these molecules remains a chemical one. Based on the known scaffold of pyrrolocytosine, two different fluorescent base analogues were exploited for two different applications in the present work. Firstly, modified PNAs containing phenylpyrrolocytosine bases designed to engage guanine with an additional hydrogen bond were tested in cell culture. Elevated melting temperatures, localization to cellular compartments, and allele-selective inhibition of expression of mutant huntingtin protein were observed. Second, driven by the historical usage of tryptophan in the fluorescence study of proteins and the commercial availability of indole derivatives, we appended a suitably-substituted indole ring system to the known pyrrolocytidine scaffold. This resulted in the base analogue (indole-3-acetamide)pyrrolocytosine which was incorporated into oligodeoxynucleotides and displayed fluorometric response to hybridization. Incorporation of this monomer was, however, in low yield and the modification was found to be mildly destabilizing to duplex formation. Some results from this chapter appear in Hu, J.; Dodd, D.W.; Hudson, R.H.E.; Corey, D.R. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 6181-6184.

7.1 Introduction to the pyrrolocytosine scaffold

Pyrrolocytosine

The heteroannulation of 5-alkynyluridine, resulting in furanouridine was described more than a quarter of a century ago. It is of limited use as its Watson-Crick base-pairing face is not complementary to any natural nucleobase; however, substitution of the furan oxygen for nitrogen yields the base-pairing competent pyrrolocyctosine (pC) (Fig. 7-1). The story of pyrrolocytosine (pC) begins in 1987 with a paper from Inoue and co-workers[1] where they synthesized both the 6-unsubstituted and methyl-subsituted (R=H,

Me. Fig 7-1) monomers and incorporated them into dodecadeoxyribonucleotides *via* the phosphotriester approach; they found that the modified bases were able to base-pair with G and did not affect duplex stability appreciably. Almost 10 years later Gamper and co-workers prepared the cyanoethylphosphoramidite of the unsubstituted (R = H) pC but found that degradation occurred during the oxidation step in the automated synthesis. They were however able to incorporate a furanouridine monomer without observing unwanted degradation during the oxidation step.[2]

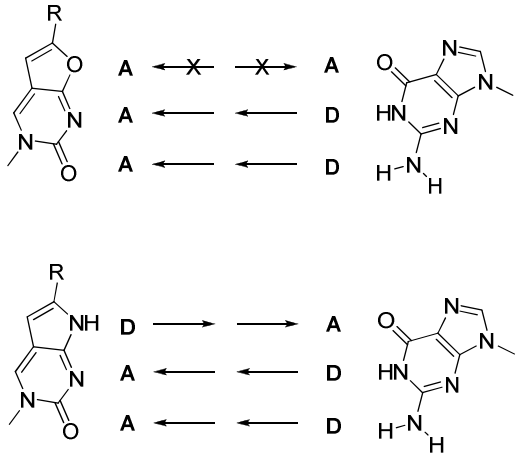


Figure 7-1: Hydrogen bond acceptor (A) and donor (D) sites for furanouridine (top left) and pyrrolocytosine (bottom left) showing (in)compatibility for hybridization with guanine.

Gamper did find though that during the deprotection and cleavage of the oligodeoxynucleotide from the resin with concentrated aqueous ammonia conversion of furanouridine to pyrrolocytosine resulted. Pyrrolocytosine is one of the rarer isosteric pyrimidine analogues and has been used as an emissive C analogue in mismatch detection,[3] and also in the study of RNA secondary structure.[4] 6-Methylpyrrolocytosine is the most commonly used derivative and is only modestly emissive as the monomer and is quenched successively on incorporation into single-stranded and double-stranded oligonucleotides.[5] Able to base-pair effectively with guanine, it can be stabilizing or destabilizing to the duplex depending on the identity of the substituent at the 6-position of the pyrrole and the sequence context. 6-Methylpyrroloctosine, first synthesized in 1987,[1] remained unexploited as a BDF until its rediscovery in 2001.[5] It was also not made in

phosphoramidite form until 2004.[6] Since that time, it has been increasingly investigated due to its quenching upon hybridization and more recent commercial availability despite its low fluorescence efficiency ($\Phi_F = 0.05$). Although initially synthesized by an alternate route, it is now commonly synthesized from uridine *via* Sonogashira coupling chemistry after iodination at *C*5 as shown in figure 7-2.[6]

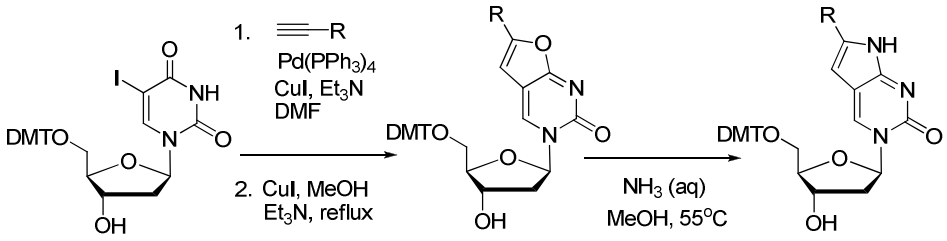


Figure 7-2: Synthetic route to pyrrolocytosine by Sonogashira chemistry followed by a 5-endo-dig cyclization and O→N atom exchange in aqueous ammonia.

Although the 6-methyl-substituted pyrrolocytosine has found success as a base-discriminating fluorophore, by variation of R group identity, greater emission and increased fluorescence response can be achieved. Substituting for a phenyl group yields a compound that is much more emissive than methylpyrrolocytosine and has a higher degree of quenching upon hybridization (85% *versus* 50% although this varies with sequence context). The overall greater luminescence of 6-phenylpyrrolocytosine can be demonstrated by the visual discrimination of mismatches which was not possible at comparable concentrations of the conventional luminophore (Fig. 7-3).[3] The greater emission intensity and increased responsiveness may translate to lower detection limits in SNP analysis.

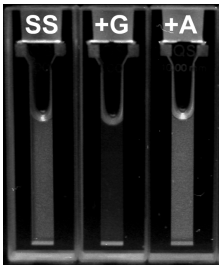


Figure 7-3: SNP analysis using a 6-phenylpyrrolocytosine containing oligonucleotide. From left to right, single-stranded, G match sequence, A mismatch. Note: quenching occurs only on hybridization to the perfectly matched sequence. Even though the A-mismatch exists as a duplex at the analysis temperature, no quenching is observed.[3]

Phenylpyrrolocytosine has been incorporated into DNA and RNA for use as a probe for HIV-1 reverse transcriptase RNase H activity. This application would be highly useful in drug-screening assays as HIV reverse transcriptase is a desirable target for small molecules. Upon cleavage of the heteroduplex up to 14-fold emission increases were observed and the modified base acted as a superior substrate to the standard FRET-base technology previously used.[7]

Mechanistically similar to the "G-clamp" (outlined in chapter 1.1), [bis-*o*-(aminoethoxy)phenyl]pyrrolocytosine (boPhpC) was designed to interact with the Hoogsteen face of guanine (Fig. 7-4).

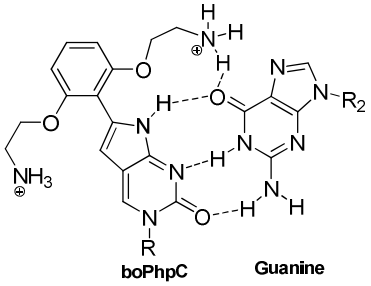


Figure 7-4: The proposed interaction of [bis-o-(aminoethoxy)-phenyl]pyrrolocytosine to guan(os)ine with an additional hydrogen bond to the Hoogsteen face.[8]

This base has been incorporated into PNA and was shown to increase the thermal stability of duplexes with natural nucleic acid oligomers by *ca.* 10°C although this does vary with sequence context.[8] The compound, like the "G-clamp", has a very high quantum yield ($\Phi = 0.32$ for boPhpC in H_2O); however, unlike the "G-clamp" boPhpC is highly responsive to microenvironmental conditions. Upon hybridization to a match sequence the emission intensity is decreased by greater than 50%. The high quantum yield, responsiveness and stabilizing properties of this base replacement give it potential for antisense/antigene applications and as a reporter group.

Although boPhpC showed remarkable binding properties the Hudson group sought to further optimize binding. In the previous work, they found that one substitution with an aminoethoxy arm at the *ortho* position of the phenyl ring showed very similar thermal stability to the bis-substituted compound with its complement sequence. This was investigated to find if accumulated positive charge (from the cationic aminoethoxy substituents) on the PNA oligomer increased binding to the DNA target by electrostatic means. As electrostatics were not deemed vital to binding the question arose of whether the linker was of appropriate length and also whether perhaps a meta-substituted compound would show increased binding due to a greater lack of steric interference and electrostatic interference with the *O*6 of the opposite guanine. Compounds shown in figure 7-5 were therefore synthesized, incorporated into PNA and their binding efficiency tested against DNA and RNA targets.[9]

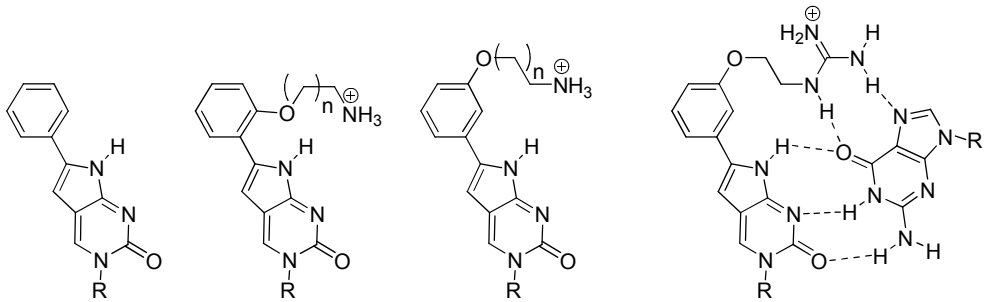


Figure 7-5: Varying substitutions of the phenyl ring of PhpC.[8,9]

The guanidinium substituted pC, mono-*meta*-guanidinum-pC (mmguaPhpC), showed the strongest binding to both RNA and DNA targets putatively because of the formation of 2 hydrogen bonds to the Hoogsteen face of guanine. It was also shown that meta-substituted compounds were stronger binders and that the first choice of linker length was appropriate as the propyl linker had slightly less affinity towards its complement. These compounds also reported fluorometrically on binding and are good candidates for future structural elucidation or antigene/antisense applications. The Hudson Laboratory therefore sought out a collaborative project with Dr. David Corey's Laboratory in order to test the efficacy of the modified base boPhpC in regards to modifying gene expression. In particular, the binding properties of the modified oligomers seemed well suited towards treatment of Huntington's disease, an area of expertise in the Corey Laboratory.

7.2 Introduction to Huntington's Disease

Pathophysiology of Huntington's Disease

Huntington's disease (or Huntington's chorea) (HD) is an autosomal dominant neurodegenerative disorder that affects 4-10 out of 100,000 people of European descent. Records of what are thought to be Huntington's disease date back as far as 1374 and it was Paracelcus in the early 1500s that coined the term chorea due to the dance-like involuntary movements of the sufferers. In the 16th and 17th centuries the cause of the disease was, like many other diseases, thought to be possession by the devil.[10] It has been postulated that one or more of the excecuted Salem "witches" in 1690s Massachusetts suffered from the disease, although this is likely conjecture it is a fitting explanation; progression of the disease is marked by (sometimes violent) mood swings and uncontrolled, jerking movement. George Huntington, a physician working in Long Island, wrote the classical description of the disease's symptomology in 1872.[11] Huntington recognized that the diseased organ was the nervous system, but beyond that… Well, these are the treatments he suggested below.

"*The treatment of chorea now most generally adopted is by purgatives, tonics, counter-irritants, and anti-spasmodics. The first indication is, if possible, to remove the exciting*

cause and it will probably be different in each individual case. Bleeding used to be employed, and it is said with good results, but it is rarely used at present, except in cases when there is much pain in the head, or along the spine, when it may be taken moderately by cups or leeches." – Huntington, G. 1872 from "*On Chorea*" ref. 11.

Then, as now, there was no cure. Drugs for the disease such as setraline, fluoxatine, and tetrabenazine are wholly insufficient and treat only the symptoms.[12-14] The disease usually manifests itself at the age of 35-40 yrs and progresses rapidly. Cognitive and psychiatric disturbances, coupled with crippling motor dysfunction lead to patients that eventually lose the ability to walk, talk and even sleep. Always fatal, the patient will often fall to general infection, heart disease or aspiration pneumonia brought about by the inability to swallow.[15]

Milton Wexler, who started Hereditary Disease Foundation with the help of Woody Guthrie's widow, knew that his two daughters had a 50:50 chance of suffering the same fate as his wife Lenore who succumbed to HD at age 53. One of Milton Wexler's daughters, Nancy, would be instrumental in identifying the genetic basis of HD (a first for any disease) thereby increasing our knowledge of the disease and the potential for a cure. To identify the gene responsible for the disease, genetic samples from a highly affected community in Venezuela were subjected to restriction fragment length polymorphism (RFLP) and linkage analyses.[16] These analyses showed a genetic marker linked to the disease on the fourth chromosome and, 10 yrs later, a truly massive collaborative project found the gene responsible for HD and the nature of the mutation responsible for the disease state.[17] The mutation is a CAG trinucleotide repeat expansion in an intragenic region towards the N-terminus of the resultant protein and increasing severity in cases of HD is correlated to an increase of these repeats. The disease is therefore termed a polyglutamine or poly-Q disorder as the mutation results in an elongated stretch of glutamine residues. The gene product of this sequence is a 348 KDa protein called huntingtin (htt), it acts in many capacities in the cell and many of its functions, although intensely researched, remain unknown. The gene is expressed ubiquitously; however, the highest levels are found within the central nervous system where HD causes the most destruction. Huntingtin has been shown to have potential

interaction with up to 50 other proteins (mostly towards the N-terminus) and the degree of interaction, in many cases, depends on the length of the CAG repetitive sequence. Different isoforms of htt interact with proteins that are involved in apoptosis, clathrin-mediated endocytosis, vesicle transport, cell signaling, transcriptional regulation and morphogenesis. These interactions have previously been reviewed.[18,19] The gene is important in early embryonic development; Htt knock-out is lethal in mice before gastrulation and is necessary for nervous system development.[20] Mutant htt is neurotoxic mostly thought to be due to aggregation of the poly-Q tracts leading to plaque formation; however, the wild-type protein is also important for neuronal health and may be neuroprotective against the mutant allele. This offers a difficult challenge: to target and block the action of the mutant protein while maintaining function of the wild-type. For reasons outlined in chapter 1.3 the antisense approach is ideal for this. Research groups have selectively targeted SNPs within the mutant mRNA sequence using siRNAs with varying degrees of success.[21-24] The problem with this strategy is that the SNPs are not universal and no one treatment could be used to success in all patients. What all HD patients do have in common is an elongated CAG repeat region. This repeat has been targeted with antisense nucleotides in the past and is also partly the subject of this chapter.[25-29]

Antisense therapies for HD

Mutant Htt transcripts have been targeted with various ASOs. PNA, ENA, LNA, 2'FANA, and PS-DNAs have all been tested *in cellulo* against patient derived cells.[28] The intitial studies concentrated on using PNAs where it was found that when targeted to the CAG repeat very small oligomers were able to decrease mutant htt expression while maintaining most of the wild-type (wt) protein in the cell. PNAs were conjugated to a stretch of 8 lysine residues of D-stereochemistry in order to improve solubility and increase cellular uptake while remaining resistant to peptidases. Effects were seen when using oligomers as small as 7mers and little difference was observed between using oligomers 13 and 21 bases long.[25] Up until quite recently the original successes have been hard to improve upon using a wide range of different ASOs. The highest knockdown and selectivity observed to date has been achieved using RNAi. Although

perfectly matched siRNAs showed very poor selectivity for the mutant allele, central mismatches within the guide stand allow the ASO to operate through a more miRNA like mechanism. In this case, the transcripts are not cleaved but are blocked by the association of AGO2 (and potentially other proteins) and very high selectivity is achieved (> 30 fold).[29]

Based on the perceived value of the steric blockade mechanism and on the findings of the intial report[25] we aimed to introduce boPhpC containing oligomers into PNAs to target the intronic CAG repeat of Htt. It is known that boPhpC increases the stability of PNA/nucleic acid hybrids[8] and we therefore hypothesized that higher thermal stability of the heteroduplexes would result in more potent antisense drugs.

7.3 PNA mediated allele selective inhibition of mutant huntingtin expression. Results and discussion.

Single base substitutions

A series of PNA-peptide conjugates were synthesized with 1 modified base in varying positions in order to determine the optimum placement of boPhpC within the oligomer. Sequences are shown in table 7-1 and contain one C-terminal lysine (K) and 8 *N*-terminal lysines for enhanced uptake by cells. The thermal stability of DNA/PNA-peptide conjugate duplexes was determined by temperature dependent UV-Visible spectroscopy against 13mer and 19mer DNA targets (Table 7-1).

Table 7-1: PNA-peptide conjugates containing one modified boPhpC base (denoted X) synthesized in the present work. Melting temperatures (T_M) and their comparison to unmodified PNA-1 (ΔT_M) are given.

PNA		DNA length	T_M (°C)	ΔT_M (°C)
PNA **1**	K-GCT-GCT-GCT-GCT-G-K_8D	13mer	69.5	-
PNA **2**	K-G**X**T-GCT-GCT-GCT-G-K_8D	13mer	74.5	+5.0
PNA **3**	K-GCT-GCT-G**X**T-GCT-G-K_8D	13mer	74.0	+4.5

PNA **4**	K-GCT-G**X**T-GCT-GCT-G-K_8D	13mer	74.3	+4.8
PNA **1**	K-GCT-GCT-GCT-GCT-G-K_8D	19mer	79.6	-
PNA **2**	K-G**X**T-GCT-GCT-GCT-G-K_8D	19mer	84.4	+4.8
PNA **3**	K-GCT-GCT-G**X**T-GCT-G-K_8D	19mer	83.8	+4.2
PNA **4**	K-GCT-G**X**T-GCT-GCT-G-K_8D	19mer	84.0	+4.4

PNA 1, which had already been examined by Corey's group, gave a melting temperature of 69.5 ˚C against a 13mer complementary sequence. Incorporation of one boPhpC results in roughly a 5˚C stabilization. Although the differences between PNAs 2-4 are minimal, added stability seems to correlate with the proximity of the modification to the *N*-terminus of the PNA as in PNA **2** the modification is closest to the *N*-terminus and in **3** it is the furthest. Against a 19mer DNA target we see an increase in T_M in all PNAs of *ca*. 10˚C. This is likely due to the interaction of the cationic lysine residues with the anionic backbone of DNA. It is therefore important to obtain data on thermal stability using DNAs or RNAs that are of sufficient length to accommodate this peptide tail.

Conjugates 1-4 were tested against HD patient derived fibroblasts in cell culture and mutant and wt Htt protein levels were gauged after 4 days of incubation by western

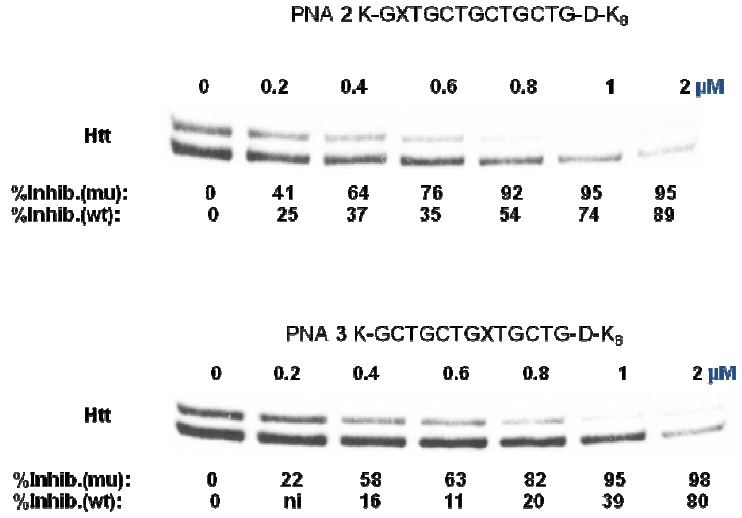


Figure 7-6: Western analysis of protein levels in fibroblasts after exposure to PNAs 2 and 3 showing selective inhibition of the mutant protein (top bands) over wt (lower bands).

analysis (see experimental for details). The results are shown in figure 7-6.

Modified PNA conjugates **2-4** were all shown to selectively inhibit the translation of the Htt mutant transcript over wt. However, one must be cautious when interpreting data from western blots as they are only semi-quantitative. No significant differences were observed between **2-4** and therefore no real conclusions about where optimal placement of the modification might be could be drawn. Moreover, other experiments carried out by Dr. Jiaxin Hu found that, in comparison to PNA **1**, PNAs **2-4** were less selective and of similar overall potency. It was also observed that a 7mer repeat sequence with one modification (PNA **5**, Fig. 7-7) offered little selectivity.

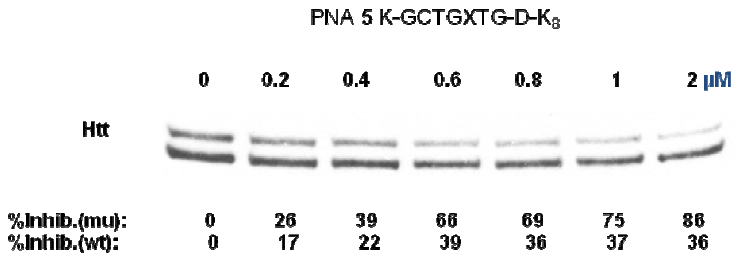


Figure 7-7: Western analysis demonstrating that 7mer conjugate PNA 5 targeted to Htt shows little selectivity towards the mutant allele.

From these data it can be said that 13mers or 7mers containing 1 boPhpC modification offer no advantage over unmodified PNAs as far as inhibition of mutant Htt is concerned. However, they do offer one advantage that will be discussed shortly. We then endeavoured to test oligomers with multiple modifications to further increase the thermal stability of the duplexes and perhaps gain insight into the relationship between binding affinity and selectivity.

Multiple base substitutions

Peptide-PNA conjugates (**6**-**8**) containing multiple base substitutions were synthesized and sent to Dr. Corey's laboratory for testing by Dr. Jiaxin Hu. T_M values were determined against a complementary RNA sequence and the resultant data are shown in table 7-2.

Table 7-2: T_M data for PNA/RNA duplexes and IC_{50} values for inhibition of mutant and wild-type Htt expression in fibroblasts. All sequences contain 1 N-terminal and 8 C-terminal D-lysine residues. X = boPhpC.

PNA Conjugate		T_M(Δ T_M)°C	IC_{50}/mut (μM)	IC_{50}/wt (μM)
PNA **1**	GCTGCTGCTGCTG	82.9	0.047 ± 0.2	>2
PNA **2**	G**X**TGCTGCTGCTG	84.5 (1.6)	0.54 ± 0.05	1.68 ± 0.7
PNA **3**	GCTGCTG**X**TGCTG	85.4 (3.5)	0.71 ± 0.07	1.86 ± 0.1
PNA **6**	G**X**TGCTG**X**TGCTG	83.9 (1)	0.58 ± 0.05	1.3 ± 0.1
PNA **7**	G**X**TG**X**TG**X**TGCTG	> 87 (>4)	0.97 ± 0.02	>4
PNA **8**	G**X**TG**X**TG**X**TG**X**TG	> 87 (>4)	2.6 ± 0.7	> 4

Against an RNA complement the boPhpC modification was stabilizing towards duplex formation albeit to a lesser extent than with DNA. All PNAs were selective to the mutant protein as seen by comparison of IC_{50} vaules for mutant versus wild-type protein; however, the selectivity and potency decreases with the number of substitutions made. PNA **8** has an IC_{50} of 2.6 μM for the mutant allele and greater than 4 μM for the wild-type. At these concentrations the oligomers begin to display toxic effects.

Fluorescence microscopy

Dr. Hu was able to gain some benefit from the fluorescence boPhpC. Through confocal microscopy, it was possible to visualize the oligomer *in cellulo* up to nine days after transfection. The oligomers were found to have a punctuate distribution and colocalized with the endosomal marker transferrin. There are many previous reports of tracking the

location of PNAs within cells and they are found to be concentrated endosomally.[30-32] In these studies PNAs were tagged with extraneous fluorophores which may affect cellular distribution. There are examples in the literature where such tags can direct the trafficking of oligomers in cells. Barton has found that a ruthenium-octaarginine conjugate is directing towards the nucleus[33] and fluorescent dyes have been found to alter the distribution of cell penetrating peptides.[34] This is believed to be the first example of 'label-free' fluorescence imaging of an ASO in cells. The only flaw to this technique lays in the responsiveness of the fluorophore to its environment. It is not known if/when/where the fluorophore is being quenched and we can therefore say little about concentration.

7.4 PNA mediated allele selective inhibition of mutant huntingtin expression. Conclusions.

In conclusion, boPhpC containing PNAs were shown to knockdown the expression of mutant Htt. It is not clear why increased binding of the oligomers to target RNA could result in lower selectivity and potency but factors may include non-specific binding, aggregate formation or endosomal sequestration. Regardless, the modification of duplex stability by the addition of these monomers makes them a useful addition to the chemical toolbox for future work. The real importance of this to the field is that the fluorescence of the monomer can be exploited for imaging purposes with concomitant gene silencing. Such dual purpose oligomers may find a great deal of use across the biological sciences in the near future.

7.4 PNA mediated allele selective inhibition of mutant huntingtin expression. Experimental.

Synthesis of PNA-peptide conjugates.

PNA-peptide conjugates were synthesized on a Rink amide resin using an Applied Biosystems 433A peptide synthesizer and standard FMOC chemistry.[8] Commercially available PNA monomers (Applied Biosystems, Warrington, UK) and Fmoc-D-

Lys(Boc)-OH (Novabiochem, Hohenbruun, Germany) were used with the exception of the boPhpC monomer which was synthesized by Dr. Filip Wojciechowski according to literature precedent.[8] Oligomers were purified by RP-HPLC and characterized by ESI-HRMS as previously described.[35,36]

Cell culture and transfection

Cell culture and transfection was carried out as described.[27] In brief, patient-derived fibroblast cell lines GM04281 were obtained from the Coriell Institute (Camden, NJ) and contained 17 intronic CAG repeats for the wild-type Htt allele and 69 for the mutant. Fibroblasts were kept in the absence of antibiotics at 37˚C with an atmosphere of 5% CO_2 in MEM (sigma) supplemented with 5% heat inactivated FBS and 0.5% MEM non-essential amino acids. Cells were plated at 60,000 cells per plate once reaching confluence in supplemented MEM 2 days prior to transfection. PNA-peptide solutions were heated at 65˚C for 5 min and diluted to the desired concentration with OptiMEM and pipette into wells. After a period of 24 h the media was replaced with supplemented MEM. Cells were harvested 4 days after initial transfection for protein analysis.

SDS-PAGE and western blotting

SDS-PAGE and western blotting were carried out as described.[27] In brief, cells were harvested with trypsin EDTA solution (invitrogen) and lysed and protein concentration was determined by the bicinchoninic acid assay (BCA). SDS-PAGE was carried out with a stacking gel comprised of 4% acrylamide/bisacrylamide (34.7:1) 450 mM Tris-acetate pH 8 and a separating gel comprised of 5% acrylamide/bisacrylamide (34.7:1) 150 mM Tris-acetate pH 6.8 using XT Tricine (Bio-Rad) as running buffer. The gels were run at 100 V for 4 h and were water cooled to prevent overheating. The expression of actin was monitored to ensure equal protein loading. The proteins were then transferred to a membrane (Hybond-C) and blocked. Primary antibodies were anti-htt from Chemicon and anti-actin from Sigma. Horseradish peroxidase (HRP) derivatized anti-rabbit antibody conjugate was used for visualization in conjunction with SuperSignal West Pic Chemiluminescent Substrate and were quantified using ImageJ software in relation to the negative control (mismatch PNA K-GCC-ACT-ACT-GAT-A-D-K_8).

7.5 Indole-3-acetamide substituted deoxypyrrolocytidine. Introduction

Having explored some uses of the pyrrolocytosine scaffold in the context of PNA we then aimed towards making similar substitutions for use in oligodeoxynucleotides (ODNs). Substitutions influence the fluorescence and binding properties when in the context of an oligomer. Tryptophan is used in fluorescence studies of proteins because of its intrinsic fluorescence. Tryptophan undergoes solvatochromic emission, is quenched by nearby protonated amino acids, and can therefore be used to measure conformational states of proteins. It was reasoned that pC substituted with a derivative of indole could also have interesting properties and be used as an environmentally sensitive nucleobase. Also, following from our previous work, we sought to modify the base so that a secondary H-bonding interaction could occur at the Hoogsteen face of an opposing guanine in order to further stabilize the duplex. Therefore the base analogue, (indole-3-acetamide)pyrrolocytosine (IAMpC) was synthesized and is shown in figure 7-8.

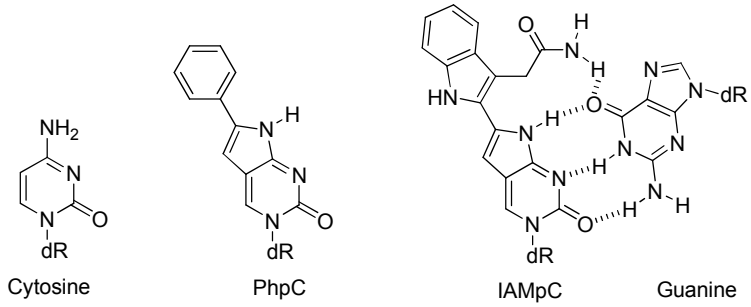


Figure 7-8: Structural comparison of cytosine, PhpC and the indole derivatized IAMpC with proposed Hoogsteen interactions with guanine shown.

The strategy for the synthesis of ODNs containing IAMpC involved a post-synthetic conversion of ODNs containing incorporated phosphoramidite **5** (Fig. 7-9) the synthesis of which was carried out by Dr. Filip Wojciechowski and will not be discussed here.

A test of whether the Boc group could be removed under standard DNA protocols (NH_4OH or NH_3/MeOH) was performed by Dr. Filip Wojciechowski and both the removal of the Boc group and amminolysis of the ethyl ester group was achieved.

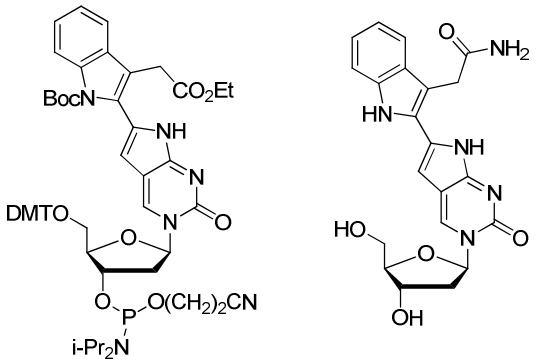


Figure 7-9: Convertible phosphoramidite of IAMpC (left) and IAMpC (right) used for fluorescence studies of the monomer.

The absorption and fluorescence spectra of IAMpC were measured in aqueous buffer and showed an absorption band centred at 375 nm and an emission maximum at 505 nm. This emission wavelength of IAMpC (505 nm) is the beginning of the green region (490-570 nm), and is significantly shifted from previously studied pCs. The compound also has a desirable large Stokes shift (130 nm) (Table 7-3). Study of the photophysical properties of IAMpC revealed that it has a very high fluorescence quantum yield in organic solvents. Remarkably the fluorescence quantum yield approached zero in aqueous buffer or water (Table 7-3). The effect of the medium on the fluorescence was also examined by studying acetone:water mixtures. As the polarity of the medium increases, the fluorescence emission noticeably decreases.

Such a dramatic decrease in emission on solvent polarity increase is a unique property of the (indole-3-acetamide)-substituted pyrrolocytosine among all the pC analogues that we have studied to date. Other pC analogues have demonstratedmore modest fluorometric response (30-50%) to changes in the solvent environment. This decrease in the fluorescence quantum yield has the potential for IAMpC to fluorometrically report on the state of hybridization or report on protein/ligand binding or 8-oxoguanosine detection.

Table 7-3: Fluorescence data for IAMpC in various solvents

Solvent	Exc, λ_{max}(nm)	Em, λ_{max}(nm)	Φ_F
1,4-dioxane	388	454	0.27
acetone	387	455	0.57
ethanol	388	472	0.48
water	375	505	~0.004

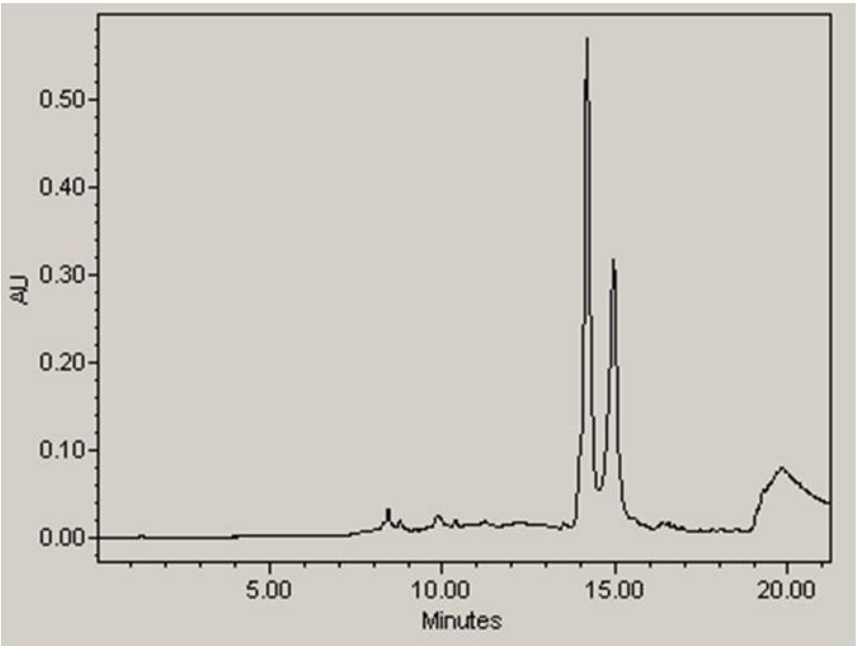
7.6 Indole-3-acetamide substituted deoxypyrrolocytidine. Results and discussion

With the phosphoramidite of IAMpC in hand, it was set out to investigate the hybridization properties and fluorescence response of IAMpC when in the context of an oligomer. The synthesis of oligodeoxynucleotides (ODNs) containing IAMpC was carried out, a selection of which are shown in table 7-4.

Table 7-4: Sequences synthesized for the present work.

ODN-1	5'-TCT-C**X**C-TCT-C-3'
ODN-2	5'-GCG-AAT-**X**TA-A-3'
ODN-3	5'-TCT-G**X**C-TCT-C-3'
ODN-4	5'-TCT-GCC-TCT-C-3'
ODN-5	5'-GAG-AGG-CAG-A-3'

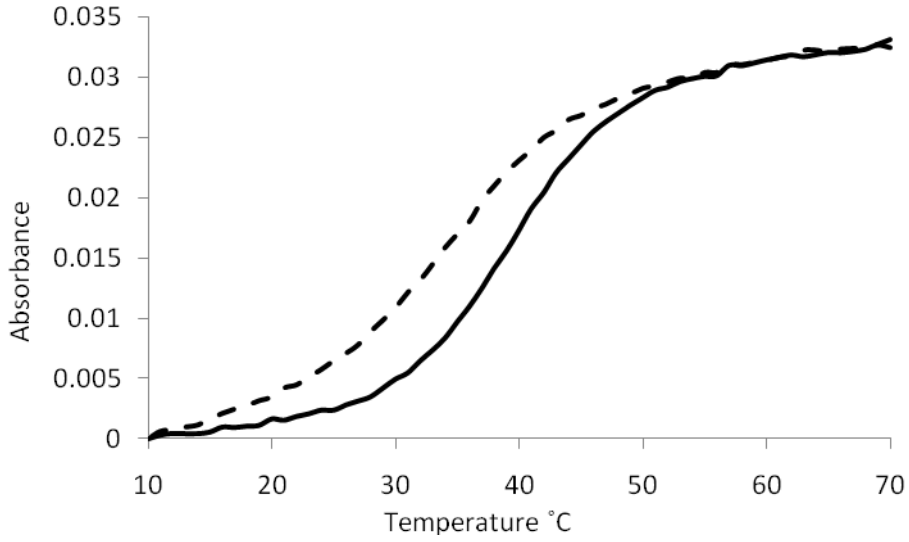
Coupling yields of the IAMpC monomer were good to excellent; however, after cleavage from the resin using concentrated aqueous ammonia overnight at 55˚C HPLC analysis revealed the presence of two peaks. A representative HPLC trace of this is shown in figure 7-10.



7-10: HPLC trace of crude ODN-3 after cleavage from the support using NH_4OH overnight at 55˚C.

The two distinct peaks were collected and subjected to mass spectral analysis. The faster eluting peak is the desired ODN-**3** in high purity (see figure 7-11) while the slower eluting peak is the desired oligomer plus a mass of 100.1 Da. This figure corresponds to the extra mass of the N-Boc protective group. Firstly, the oligomer was subjected to longer exposures to concentrated ammonia at a slightly elevated temperature of 60˚C with no success. It appeared that the ratio of protected to unprotected remains almost unchanged. More aggressive approaches such as 40% methylamine were attempted also to no success and appeared to degrade the oligomer over time before removal of the Boc was complete. This problem must be solved for further development of this work, it is possible that the protecting group is not required during DNA synthesis and this should be attempted.

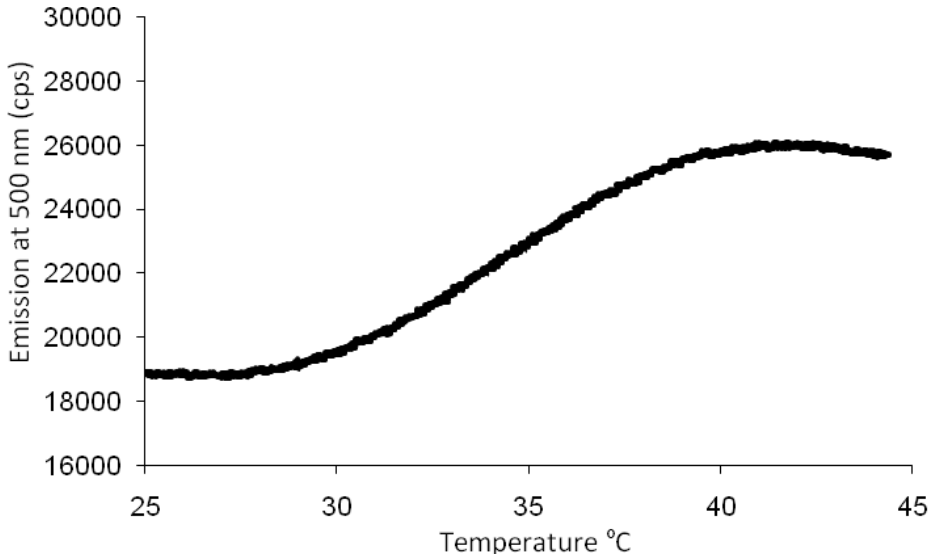
In the case of ODN-**3**, the desired oligomer was purified in sufficient quantity for further study. Thermal denaturation profiles of IAMpC-containing ODN-**3** and unmodified ODN-**4** with their complement ODN-**5** are shown in figure 7-11.



7-11: Melting curves for IAMpC containing ODN-3 (dashed line) and ODN-4 (solid line) with ODN-5 showing decreased thermal stability of the modified duplex.

The T_M of ODN-**3**/ODN-**5** duplex was determined to be 36.5˚C while that of ODN-**4**/ODN-**5** was 41.0˚C. The IAMpC modification therefore destabilizes the duplex by 4.5˚C. This could due to the linker being of inappropriate length. It is also true that an ammonium group (such that is present in boPhpC) is a better hydrogen bond donor than the carboxamide present in IAMpC for Hoogsteen recognition. These data were confirmed by fluorescence dependent thermal denaturation as shown in figure 7-12.

From the fluorescence data collected we can see that IAMpC is quenched by approximately 30 % on binding to its complementary sequence. Although not optimal, the sequence is yet to be tested against mismatch sequences and this modified base may still find use as a BDF. All necessary complementary ODN sequences have been synthesized and purified; however, the IAMpC containing ODNs are lacking due to the problems of Boc group removal that were discussed earlier.



7-12: Thermal denaturation monitored by fluorescence emission confirming a melting temperature of 35˚C and showing fluorescence quenching on binding to a match sequence.

7.7 Indole-3-acetamide substituted deoxypyrrolocytidine. Conclusions.

In conclusion, a new indole-containing pyrrolocytosine analogue has been incorporated into DNA and responds fluorometrically to the presence of a target sequence. IAMpC was found to be mildly destabilizing towards duplex formation. It is postulated that the nature of the extending 'arm' is at fault, previous pC analogues that operate by Hoogsteen recognition have been cationic and 5 atoms in length or longer. IAMpC's 'arm' is 4 atoms in length and electronically neutral. Compounding problems with Boc removal may, in the future, warrant a new synthetic design.

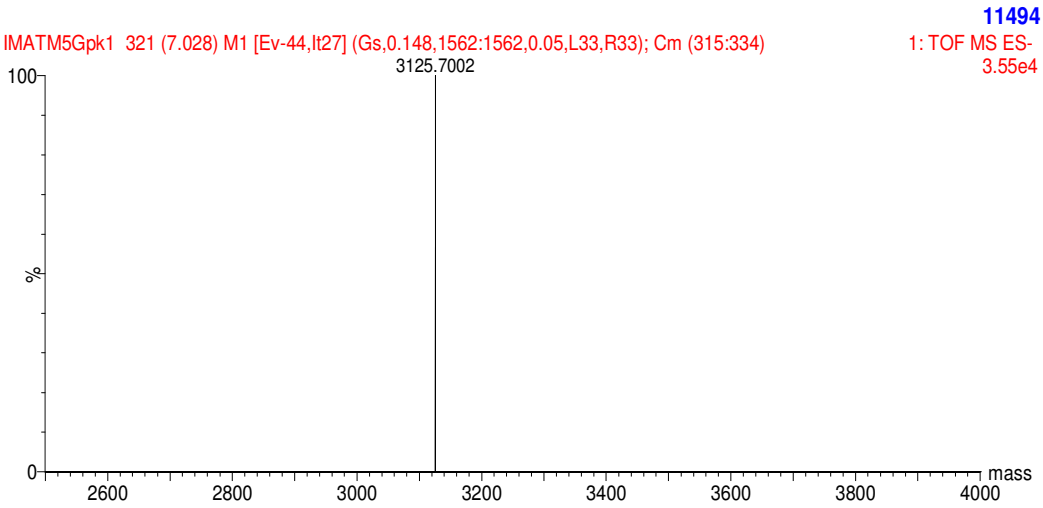
To provide a conclusion for the preceeding chapters, the design and syntheses of fluorescent, in some cases duplex-stabilizing, cytosine analogues was carried out. Successful incorporation of selected analogues into PNA and DNA scaffolds yielded oligomers that could be used for both sequence interrogation and alteration of gene expression. Platinum complexes were synthesized and evaluated for their suitability as

antineoplastic agents. Platinum was also used in an attempt to increase the hybridization strength of PNA to an RNA target. The work contained herein may be expanded in several directions. One of these is the development of on-resin approaches to generate usefully modified oligomers through the use of dipolar cycloadditions. Others would include the testing of the suitability of these monomers for enzymatic incorporation and further photophysical characterization of oligomers used in knockdown experiments.

7.7 Indole-3-acetamide substituted deoxypyrrolocytidine. Experimental.

DNA synthesis and purification

DNA was synthesized at 0.2 µmol scale on an ABI 392 DNA synthesizer using commercially available reagents obtained from ChemGenes (Wilmington, MA) with either 5-ethylthiotetrazole or 4,5-dicyanoimidazole (0.25 M) as activator using the cycles and conditions supplied by the manufacturer. For unmodified bases, coupling times were extended to 180 seconds. IAMpC phosphoramidite was coupled manually at 10-fold excess in 200 µL MeCN with 100 µL of 0.25 M 5-ethylthiotetrazole with a coupling time of 10 min. The efficiency of this method was first investigated using the commercially available C monomer. Oligomers were cleaved from the resin using concentrated aqueous ammonia at 55C overnight, isolated by lyophilisation and first purified 'DMT on' by RP-HPLC on a C-18 300 Å column using 100 mM NEt_3•HOAc, pH 6.5 and MeCN as eluent. Lyophilized 'DMT on' was treated with 80% acetic acid, lyophilized again and subjected to the same HPLC conditions. ODNs were characterized by HRMS. ODN-**3** - HRMS calcd. 3125.0034 found 3125.7002.



Thermal denaturation experiments

ODN concentration was calculated using the nearest neighbour method[37] for unmodified sequences. The absorption coefficient of IAMpC at 260 nm was calculated to be 15450 $M^{-1}cm^{-1}$ by the gradient of a plot of absorbance versus concentration. UV-Vis dependent thermal denaturation was carried out in aqueous 100 mM NaCl, 10 mM Na_2HPO_4, 0.1 mM EDTA, pH 7.0 at a strand concentration of 8µM. Denaturation was performed between 10 °C to 95 °C at a rate of 0.5 $°Cmin^{-1}$. Melting temperature is given as an average of three runs within 1˚C and estimated for cooperative transitions by the first derivative method.

7.8 References

1. Inoue, H.; Imura, A.; Ohtsuka, E. *Nippon Kagaku Kaishi* **1987**, *7*, 1214-1220.

2. Woo, J.; Meyer, R.B., Jr.; Gramper, H.B. *Nucleic Acids Res*. **1996**, *24*, 2470-2475.

3. Hudson, R.H.E.; Ghorbani-Chogamarani, A. *Synlett*, **2007**, 870-873.

4. Tinsley, R.; Walter, N.G. *RNA*, **2006**, *12*, 522-529.

5. Liu, C.; Martin, C.T. *J. Mol. Biol.*, **2001**, *308*, 465-475.

6. Berry, D.A.; Jung, K.-Y.; Wise, D.S.; Sercel, A.D.; Pearson, W.H.; Mackie, H.; Randolph, J.B.; Somers, R.L. *Tetrahedron Lett.*, **2004**, *45*, 2457-2461.

7. Wahba, A.S.; Esmaeili, A.; Damha, M.J.; Hudson, R.H.E. *Nucleic Acids Res*. **2010**, *38*, 1048-1056.

8. Wojciechowski, F.; Hudson, R.H.E. *J. Am. Chem. Soc.*, **2008**, *130*, 12574-12575.

9. Wojciechowski, F.; Hudson, R.H.E. *Org. Lett*. **2009**, *11*, 4878-4881.

10. Zuccato, C.; Valenza, M.; Cattaneo, E. *Physiol. Rev.* **2010**, *90*, 905-981.

11. Huntington, G. *Medical and Surgical Reporter of Philadelphia* **1872**, *26*, 317–321.

12. Peng, Q.; Masuda, N.; Jiang, M.; Li, Q.; Zhao, M.; Ross, C.A.; Duan, W. *Exp. Neurol*. **2008**, *210*, 154-163.

13. Grote, H.E.; Bull, N.D.; Howard, M.L.; van Dellen, A.; Blakemore, C.; Bartlett, P.F.; Hannan, A.J. *Eur. J. Neurosci*. **2005**, *22*, 2081-2088.

14. Hayden, M.R.; Leavitt, B.R.; Yasothan, U.; Kirkpatric, P.; *Nat. Rev. Drug Discovery* **2009**, *8*, 17-18.

15. Sørensen, S.A.; Fenger, K. *J. Med. Genet.* **1992**, *29* 911–914.

16. Gusella, J.F.; Wexler, N.S.; Connealy, P.M.; Naylor, S.L.; Anderson, M.A.; Tanzi, R.E., Watkins, P.C.; Ottina, K.; Wallace, M.R.; Sakaguchi, A.Y. *Nature* **1983**, *306*, 234-238.

17. MacDonald, M.E.; Ambrose, C.M.; Duyao, M.P.; Myers, R.H.; Lin, C.; Srinidhi, L.; Barnes, G.; Taylor, S.L.; James, M.; Groot, N.; MacFarlane, H.; Jenkins, B.; Anderson, M.A.; Wexler, N.S.; Gusella, J.F.; Bates, G.P.; Baxendale, S.; Hummerich, H.; Kirby, S.; North, M.; Youngman, S.; Mott, R.; Zehetner, G.; Sedlacek, Z.; Poustka, A.; Frischauf, A.-M.; Lehrach, H.; Buckler, A.J.; Church, D.; Doucette-Stamm, L.; O'Donovan, M.C.; Riba-Ramirez, L.; Shah, M.; Stanton, V.P.; Strobel, S.A.; Draths, K.M.; Wales, J.L.; Dervan, P.; Housman, D.E.; Altherr, M.; Shiang, R.; Thompson, L; Fielder, T.; Wasmuth, J.J.; Tagle, D.; Valdes, J.; Elmer, L.; Allard, M.; Castilla, L.; Swaroop, M.; Blanchard,

K.; Collins F.S.; Snell, R.; Holloway, T.; Gillespie, K.; Datson, N.; Shaw, D.; Harper, P.S. *Cell* **1993**, *72*, 971-983.

18. Li, S.H.; Li, X.J. *Trends Genet*. **2004**, *20*, 146-154.

19. Harjes, P.; Wanker, E.E. *Trends Biochem Sci.* **2003**, *28*, 425-433.

20. Duyau, M.P.; Auerbach, A.B.; Ryan, A.; Persichetti, F.; Barnes,G.T.; McNeil, S.M.; Ge, P.; Vonsattel, J.F.; Joyner, A.L. *Science* **1995**, *269*, 407-410.

21. Miller, V.M.; Xia, H.; Marrs, G.L.; Gouvion, C.M.; Lee, G.; Davidson, B.L.; Paulson, H.L. *Proc. Natl. Acad. Sci. USA*. **2003**, *100*, 7195-7200.

22. Schwarz, D.S.; Ding, H.; Kennington, L.; Moore, J.T.; Schelter, J.; Burchard, J.; Lindsey, P.S.; Aronin, N.; Xu, Z.; Zamore, P.D. *PLoS Genet*. **2006**, *2*, e140.

23. van Bilsen, P.H.; Jasper, L.; Lombardi, M.S.; Odekerden, J.C.; Burright, E.N.; Kaemmerer, W.F. *Hum. Gen. Ther*. **2008**, *19*, 710-719.

24. Zhang, Y.; Engelman, J.; Friedlander, R.M. *J. Neurochem*. **2009**, *108*, 82-90.

25. Hu, J.; Matsui, M.; Corey, D.R. *Ann. NY. Acad. Sci*. **2009**, *1175*, 24-31.

26. Hu, J.; Dodd, D.W.; Hudson, R.H.E.; Corey, D.R. *Bioorg. Med. Chem. Lett*. **2009**, *19*, 6181-6184.

27. Hu, J.; Matsui, M.; Gaganon, K.T.; Schwartz, J.C.; Gabillet, S.; Arar, K.; Wu, J.; Bezprozvanny, I.; Corey, D.R. *Nat. Biotechnol*. **2009**, *27*, 478-484.

28. Gagnon, K.T.; Pendergraff, H.M.; Deleavey, G.F.; Swayze, E.E.; Potier, P.; Randolf, J.; Roesch, E.B.; Chattopadhyaya, J.; Damha, M.J.; Bennet, C.F.; Montaillier, C.; Lemaitre, M.; Corey, D.R. *Biochemistry* **2010**, *49*, 10166-10178.

29. Hu, J.; Liu, J.; Corey, D.R. *Chem. Biol*. **2010**, *17*, 1183-1188.

30. Kaihatsu, K.; Huffman, K.E.; Corey, D.R. *Biochemistry* **2004**, *43*, 14340-14351.

31. Lebleu, B.; Moulton, H.M.; Abes, R.; Ivanova, G.D.; Abes, S.; Stein, D.A.; Iversen, P.L.; Arzumanov, A.A.; Gait, M.J. Adv. *Drug Delivery Rev*. **2007**, *60*, 517-529.

32. Shiraishi, T.; Pankratova, S.; Nielsen, P.E. *Chem. Biol*. **2005**, *12*, 923-929.

33. Puckett, C.A.; Barton, J.K. *J. Am. Chem. Soc*. **2009**, *131*, 8738-8739.

34. Szeto, H.H.; Schiller, P.W.; Zhao, K.; Luo, G. *FASEB J*. **2005**, *19*, 118-120.

35. Peptide Nucleic Acids, Protocols and Applications, 2nd Ed., Nielsen, P. E. Ed. Horizon Bioscience, Norfolk, UK, 2004. chpt. 11.

36. Hudson, R.H.E.; Liu, Y.; Wojciechowski, F. *Can. J. Chem.* **2007**, *85*, 302-312.

37. Puglisi, J.D.; Tinoco, I. Jr. *Method. Enzymol*. RNA Process. Pt. A, **1989**, *180*, 304.

Curriculum Vitae

Name: David W. Dodd

Post-secondary Education and Degrees:

The University of Western Ontario
London, Ontario, Canada
2002-2006 B.Sc. Chemistry and Biochemistry
Supervised by Dr. N.D. Jones

The University of Western Ontario
London, Ontario, Canada
2006-2011 Ph.D. Chemistry
Supervised by Dr. R.H.E. Hudson

The University of Texas Southwestern Medical Center
Dallas, TX, USA
Starting 2011 as a Postdoctoral Associate
Supervisor Dr. D.R. Corey

Honours and Awards

Nucleosides, nucleotides and nucleic acids travel Award 2010

Graduate Research Thesis Fund 2010

Oral Presentation Award – Nucleic Acids Division, 92nd Canadian Chemistry Conference 2009

NSERC CGS-D Recipient 2009

OGS Recipient 2009 (Declined)

ASPIRE– UT Southwestern Medical Center 2008

Western Graduate Research Scholarship recipient 2007-2008

Publications:

David W. Dodd and Robert H.E. Hudson. "Triazolyldeoxycytidines as low molecular weight hydrogelators." *Artificial DNA* **2010**, *2*, 90-95.

David W. Dodd, Kalen N. Swanick, Jacquelyn T. Price, Alison L. Brazeau, Michael J. Ferguson, Nathan D. Jones and Robert H.E. Hudson. "Blue fluorescent deoxycytidine analogues: convergent synthesis, solid-state and electronic structure, and solvatochromism." *Organic and Biomolecular Chemistry* **2010**, *8*, 663-666.

Jiaxin Hu,* David W. Dodd,* Robert H.E. Hudson and David R. Corey. "Cellular localization and allele-selective inhibition of mutant huntingtin protein by peptide nucleic acid oligomers containing the fluorescent nucleobase [bis-o-(aminoethoxy)phenyl]pyrrolocytosine." *Bioorganic and Medicinal Chemistry Letters* **2009**, *19*, 6181-6184. *Authors contributed equally.

Tsvetan G. Gantchev, Sonia Girouard, David W. Dodd, , Filip Wojciechowski, , Robert H. E. Hudson, and Darel J. Hunting. "gamma-Radiation Induced Interstrand Crosslinks in PNA:DNA Heteroduplexes." *Biochemistry* **2009**, *48*, 7032-7044.

David W. Dodd and Robert H.E. Hudson. "Intrinsically Fluorescent Base-Discriminating Nucleosides." *Mini-reviews in Organic Chemistry* **2009**, *6*, 378-391.

David W. Dodd, Heather E. Toews, Michael J. Trevail, Michael C. Jennings, Robert H.E. Hudson and Nathan D. Jones. "Synthesis and evaluation of the *in vitro* DNA binding properties of chiral *cis*dichloro(pyridyloxazoline)platinum(II) complexes." *Canadian Journal of Chemistry* **2009**, *87*, 321-327.

David W. Dodd and Robert H. E. Hudson. "Analysis and purification of peptide nucleic acids by denaturing polyacrylamide gel electrophoresis." *Electrophoresis* **2007**, *28*, 3884-3889.

David W. Dodd, Heather E. Toews, Florentino d.S. Carneiro, Michael C. Jennings, Nathan D. Jones. "Model intermolecular asymmetric Heck reactions catalyzed by chiral pyridyloxazoline palladium(II) complexes." *Inorganica Chimica Acta* **2006**, *359,* 2850-2858.

Published Conference Proceedings:

David W. Dodd, Sashko Damjanovski and Robert H.E. Hudson. "Peptide nucleic acid Pt(II) conjugates: a preliminary study of the antisense effects in *Xenopus laevis*." *Nucleic acids symposium series* **2010**.

David W. Dodd and Robert H.E. Hudson. "Heterocycle-appended cytidines as base-discriminating fluorophores, hydrogelators and post-synthetic labels." *Nucleic acids symposium series* **2010**.

Filip Wojciechowski, David W. Dodd, André H. St. Amant and Robert H.E. Hudson. "Indole-3-acetamide substituted deoxypyrrolocytidine: unprecedented fluorescence response." *Nucleic acids symposium series* **2010.**

Presentations:

Hamilton, Ontario – 53rd Genetics Society of Canada Meeting 2010 – Oral presentation: David W. Dodd* and Robert H.E. Hudson. "Base-pairing competent fluorophores: gene silencing with concomitant, non-intrusive, fluorescence labelling."

Toronto, Ontario – 93rd Canadian Chemistry Conference and Exhibition 2010 – Oral presentation. Robert H.E. Hudson*, David W. Dodd, Filip Wojciechowski and André H. St. Amant. "Minimally modified fluorescent nucleobases."

Toronto, Ontario – 93rd Canadian Chemistry Conference and Exhibition 2010 – Poster presentation: David W. Dodd* and Robert H.E. Hudson. "5-ethynyldeoxycytidine as a route to small molecule hydrogelators and intrinsic fluorophores."

Hamilton, Ontario – 92nd Canadian Chemistry Conference and Exhibition 2009 – Oral presentation: David W. Dodd*, Kalen N. Swanick, Jaquelyn T. Price, Alison L. Brazeau, Michael J. Ferguson, Zhifeng Ding, Nathan D. Jones and Robert H.E. Hudson. "Intrinsically Photoluminescent and Electrochemiluminescent Cytidine Analogues Accessed *via* the Huisgen Cycloaddition."

University of Toronto - Quebec-Ontario Mini Symposium of Bioorganic and Organic Chemistry 2008 – Poster presentation: Kalen N. Swanick*, David W. Dodd, Robert H.E. Hudson, Zhifeng Ding and Nathan D. Jones. "A New Class of Fluorophore-labelled Nucleobases for Electrogenerated Chemiluminescence."

Université de Montréal - Quebec-Ontario Mini Symposium of Bioorganic and Organic Chemistry 2007 – Poster presentation: David W. Dodd* and Robert H.E. Hudson. "Peptide nucleic acid Pt(II) conjugates and their potential antisense effects in *Xenopus laevis*."

Université de Montréal - Quebec-Ontario Mini Symposium of Bioorganic and Organic Chemistry 2007 – Poster presentation: David W. Dodd* and Robert H.E. Hudson. "Alkylbenzyltriazole modified cytidines as fluorescent nucleoside analogs and potential hydrogelators."

University of Western Ontario - Inorganic Discussion Weekend 2005 - Oral presentation: David W. Dodd* and Nathan D. Jones. "Asymmetric Heck Reactions Catalyzed by Pyridyloxazoline Palladium(II) Complexes."University of Toronto - SOUSCC 2005 - Oral presentation: David W. Dodd* and Nathan D. Jones. "Chiral ligands for asymmetric catalysis."