Synthesis and characterization of oligonucleotides containing an O^6-deoxyguanosine-alkyl-O^6-deoxyguanosine interstrand cross-link

Jason D. M. Booth

A Thesis

in

The Department

of

Chemistry and Biochemistry

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Science (Chemistry) at
Concordia University
Montreal, Quebec, Canada

January 2008

Canada

Abstract

Synthesis and characterization of oligonucleotides containing an O^6-deoxyguanosine-alkyl-O^6-deoxyguanosine interstrand cross-link

Jason D. Booth

Studies have shown that resistance to bifunctional alkylating agents used in the treatment of chronic myelogenous leukemia can result in diminished effectiveness of these therapeutics. To further understand this process we have developed clinically relevant models of cross-linked DNA duplexes to study the structure and repair of lesions induced by the bifunctional alkylating agents: busulfan and hepsulfam. To emulate the known deoxyguanosine-N^7-alkyl-N^7-deoxyguanosine lesion formed by hepsulfam and the unidentified lesion(s) formed by busulfan, we have designed three synthetic approaches leading to the synthesis of six stable deoxyguanosine dimers. General mono, bi and tripartite syntheses were developed for the preparation of 1,1 and 1,3-deoxyguanosine-O^6-alkyl-O^6-deoxyguanosine cross-links employing 'fast-deprotecting' phenoxyacetyl protection at the N^2 position for facile removal. The O^6-alkyl couplings are performed via a Mitsunobu reaction between a nucleoside and a mono-protected diol. The phosphoramidites of synthesized dimers were incorporated into one helical turn and 'sticky-end' DNA duplexes via solid-phase synthesis employing a DNA synthesizer. All duplexes were of defined structure and sequence for the purpose of further conformational analysis and enzymatic repair studies. Sequence composition was confirmed via enzymatic digestion and purity by subsequent HPLC analysis. A variety of methods including ultraviolet transition thermal denaturation (T_m), circular dichroism (CD) and gel electrophoretic studies were employed to characterize the cross-linked duplexes. CD and T_m studies suggest little deformation from native B-form conformation for the hepsulfam mimics. Sufficient quantities of the mono and bipartate heptyl cross-linked duplexes were obtained for various structural studies with required quantities for biological investigations obtained for the monopartate synthesis.

Acknowledgements

Dedication

I wish to express the incalculable blessing that it has been to know my father and brother. I wish to thank them for their unconditional love and support.

Table of contents

Chapter 3 : Results and Discussion

List of Figures

Chapter 1

List of Schemes

Chapter 1

Chapter 2

Chapter 3

List of Tables

Chapter 3

List of abbreviations

ACN	- Acetonitrile
AGT	- O^6-alkylguanine-DNA alkyltransferase
Alloc-OBt	- Allyl 1–benzotriazolyl carbonate
AMBER	- Assisted Model Building with Energy Refinement
BCNU	- 1,3-bis(2-chloroethyl)-1-nitrosourea
BFAs	- Bifunctional alkylating agents
BSO	- L-buthionone sulfoximine
BU	- Busulfan
CD	- Circular dichroism
CML	- Chronic mylogenous leukemia
CPG	- Controlled pore glass
dA	- Deoxyadenine
dC	- Deoxycytosine
DCC	- Dicyclohexylcarbodimide
DCM	- Dichloromethane
dG	- Deoxyriboguanosine
dG*	- O^6-alkylated deoxyriboguanosine
DIAD	- Disopropylazidodicarboxylate
DIPEA	- Diisopropylethyl amine
DMAP	- Dimethylaminopyridine
DMF	- Dimethylformamide
DMT	- 4-O-dimethoxytrityl
DNA	- Deoxyribonucleic acid
dT	- Deoxythymidine
EtOAc	- Ethylacetate
FAPY	- Formamidopyrimidine
GB/DA	- Generalized Born / Dispersed Area
GSH	- Glutathione
GST	- Glutathione-S-transferase
HM	- Hepsulfam
HPLC	- High performance liquid chromatography
ICL	- Interstrand cross-link
MALDI–TOF	- Matrix assisted laser desorption/ionization – time of flight
MFAs	- Mono-functional alkylating agents
MMR	- Mismatch repair
MS	- Mass spectrometry
NER	- Nucleotide excision repair
NMR	- Nuclear magnetic resonance
NOESY	- Nuclear overhauser enhancement spectroscopy
O^6-BG	- O^6-benzylguanine
Pac-Cl	- Phenoxyacetyl chloride
PAGE	- Polyacrylamide gel electrophoresis
PCR	- Polymerase chain reaction
Pd(PPh$_3$)$_4$	- Palladium tetrakis(triphenylphosphine)

PPh$_3$	- Triphenylphosphine
SAR	- Structure activity relationship studies
SAX	- Strong anion exchange
SVPDE	- Snake venom phosphodiesterase
TBAF	- Tetrabutylammonium fluoride
TBDPS-Cl	- t-Butyldiphenylsilyl chloride
TBS	- t-Butyldimethylsilyl
TCA	- Trichloroacetic acid
TEA	- Triethylamine
THF	- Tetrahydrofuran
TLC	- Thin layer chromatography
Tm	- Thermal melt
TpA	- Thymine phosphate adenine sequence
UV	- Ultraviolet
WC	- Watson-Crick

Chapter 1 : Introduction

1.0 Background

Cancer is second in the developed world only to heart disease in the number of deaths that are caused. In spite of the enormous thrust of research to understand and treat cancer in its various forms, it remains one of the most frequent ailments in Western society. A central tenet of the problem is a fundamental lack of understanding into the mechanisms of initiation and cessation of the disease. Some of the most common medications employed are toxic to both healthy and cancerous cells. A severe limitation to the effectiveness of many therapeutics is the requirement to selectively kill neoplastic cells must be balanced against preserving healthy cells in the patient. This results in a limitation to effective treatment being the amount of damage that can be tolerated by normal cells. Unfortunately, many treatments commonly also present a latent risk of future complications due to the non-specific nature of lesions inflicted upon DNA in all cells.[1] This presents the impetus to better understand the role of cellular repair mechanisms as it pertains to increasing the therapeutic window of existing anti-cancer treatments or the development of new medications possessing higher specificity.

Molecular biology has come to the forefront as one of the most effective tools in our arsenal to discern enzymatic repair pathways critical to understanding the origins of the development of resistance to existing therapies. From this enormous body of research, many valuable insights have been gleaned into the mechanism of initiation of cancer and cellular repair mechanisms. The increased anti-tumour activity of bifunctional alkylating agents (BFAs) over mono-functional alkylating agents (MFAs) has been known for more than one half of a century.[2] The toxicity that results from the ability to form an interstrand

1

cross-link (ICL) in DNA has made them some of the most commonly employed chemotherapeutic agents still in use.[3, 4, 5, 6] The diminished toxicity of MFAs has been related to single strand breaks or damage caused to individual bases whereas BFA have been shown to induce double strand breaks.[2] ICLs can also serve to act as an absolute barrier prohibiting strand separation thus arresting DNA transcription and/or replication leading to apoptosis and inhibition of tumour growth.[5]

It is commonly held that alkylating agents are non-specific with respect to the cell-cycle phase, occurring in both dividing and non-dividing cells.[7] Alkylation of DNA has deleterious biological consequences during DNA replication. More specifically, ICL are highly toxic to proliferating cells due to their ability to induce mutations and chromosomal rearrangements causing gross chromosomal alterations such as double strand breaks or single strand gaps resulting in apoptosis.[1,5,8] The ability of ICL agents to obstruct normal cellular metabolism has resultantly placed them among the most toxic sources of DNA damage. A survey of 234 genotoxic agents demonstrated that 12 of the 20 most toxic were BFAs capable of forming ICLs.[9] In like fashion a single ICL can kill a repair-deficient bacteria or yeast cell,[3] where about 20 can be lethal to mammalian cells that lack the ability to remove the cross-link.[5] The acute toxicity of ICLs can be demonstrated by their potential to induce tumour formation in rodents where the TD_{50} (the total life time dose of carcinogen required to increase the probability of tumour formation to 50%) is 10-1000 fold less for cross-linking agents than for MFAs.[1] In addition to exogenous sources of ICL agents such as psoralens present in many plants and cosmetics, endogenous agents can be generated during lipid peroxidation such as malondialdehyde also forming ICL.[3,10] The relevance of lipid peroxidation is not considered significant due to their low stability and

2

relatively low reactivity under physiological conditions.[11, 12] The inherent toxicity that both exogenous and endogenous ICL agents pose to normal cellular metabolism has consequently led cells to necessarily develop mechanisms to repair these absolute blocks to strand separation in their DNA.

To cope with the potentially lethal effects of ICLs, cells have evolved complex systems of genetic repair to detect and remove these deleterious lesions. More remains to be discovered about cellular recognition mechanisms if the origins of drug resistance are to be discovered. What is known is that ICLs can induce helical distortion such as unwinding or bending which is thought to be one of the primary sources of lesion recognition. Generalizations are difficult as the structures of ICL show a broad range of diversity which could affect the recognition of ICL lesions.[3] Another possible means of ICL identification is the detection of a blockage in a replication fork or transcription.[3] Whether resistance is acquired via exposure or is inherent to the organism, several pathways commonly are required to work in concert to repair or circumvent ICLs.[3, 7]

Broadly defined, drug resistance for BFAs can typically be described as the capacity to detect and repair ICL lesions in DNA caused by cytotoxic drugs. Currently, many patients that undergo treatment with these medications develop a resistance to existing chemotherapeutics resulting in a decreased efficacy of the drug. Resistance to drug regiments consists of a multifactorial combination of variables that are attributed, but not limited to, enhanced repair of DNA which is strongly correlated to the outcome of the individual patient.[7,13] For example some contributing factors to poor response to therapies consist of a myriad of reasons from enhanced resistance via a number of repair pathways to

3

the response of the patient, the type of drug employed, cellular uptake, the strain of tumour cell treated and multiple unknown factors.[3, 7]

To address the importance of understanding the role of DNA damage in cancer formation and prevention, the study of the repair mechanisms of DNA is of paramount importance to discover new therapeutics that have the potential to act specifically and/or evade repair.[5] Complicating the identification of trends in the study of repair systems is the growing consensus that different cross-links and in some cases different orientations of the same cross-link, are repaired in a different manner.[5] It has been estimated that the most challenging obstacle to discern repair pathways is the existence of multiple overlapping and redundant mechanisms.[5] Many repair studies have been performed on prokaryotic cells which have revealed valuable insights into DNA cross-link repair mechanisms however repair in eukaryotic cells is considerably more complex which is reflected in the increase in the number of proteins involved and the difference in the number of ICL that can be lethal to a cell.[3,5] In addition to more complex repair systems in eukaryotic cells, a BFA must gain entry into the cell followed by access into the nucleus. A serious limitation to the study of DNA repair mechanisms towards ICL is the current lack of substrates of defined structure which can be addressed through synthesis of model sequences.[14]

Due to the relative stability of thymine psoralen dimers when contrasted to the inherent instability of N^7-alkyl-2'-deoxyguanosine adducts which tend to depurinate leaving abasic sites, the majority of studies employing ICL DNA substrates of defined sequence and structure have focused on the use of psoralens (Scheme 1.1).[5,16] Significant aspects of ICL in mammalian cells are comparable to repair in bacteria and yeast,[3] however multiple aspects of repair do not translate to eukaryotic systems despite a significant amount of

4

homology. Additionally, it is highly unlikely that psoralen adducted thymine repair mechanisms are identical to those employed to remove cross-links between N^7-2'-deoxyguanosine adducts. This is supported by pre-clinical studies from Teicher, Frei and co-workers which has demonstrated that multiple cell lines generally exhibit little cross-resistance to different BFAs.[15] This serves to highlight the necessity of DNA segments of defined structure and sequence to allow for isolation of experimental variables involved in the repair process.

Although it is unlikely that ICL agents can be developed that are not liable to cellular repair mechanisms, the potential exists to design systems that can better evade repair or possess higher specificity.[7] Potentially with the identification of the enzymes involved small molecule inhibitors could be synthesized to selectively mitigate repair in tumour cells. The immediate challenge is to synthesize DNA duplexes comprised of stable mimics incorporated into defined sequences and with known structure to act as a diagnostic tool to probe the molecular mechanisms of repair.

Scheme 1.1 Formation of a stable [2 + 2] cycloaddition psoralen mono (**1b, 2b**) and cross-linked (**3b**) adducts in DNA from UVA activation of a psoralen intercalator with specific affinity for d(TpA)$_2$. Both furan and pyrone rings of psoralen react separately giving two possible products as only addition to the furan ring first results in a low enough barrier of activation permitting the second pyrone addition resulting in an ICL.[16]

Psoralen

320-410 nm

1b

320-410 nm

2b

>320 nm

3b
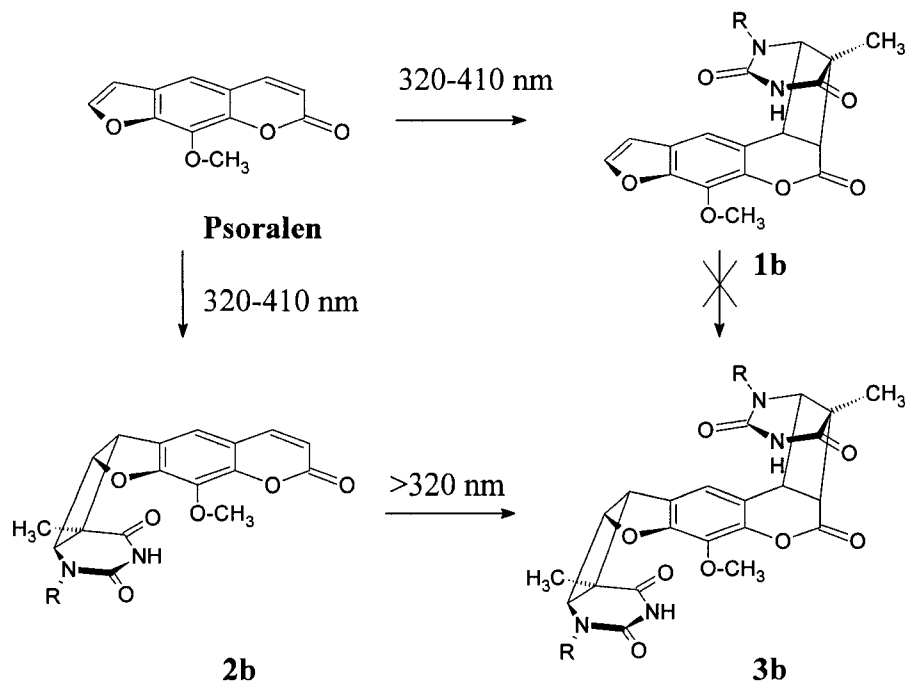
1.1 Busulfan and Hepsulfam

The first clinical usage of busulfan (BU) was reported over 40 years ago for the treatment of chronic mylogenous leukemia (CML).[17] By contrast the first pre-clinical use of hepsulfam (HM) was published in 1988 as a replacement of BU in treatment in CML.[18] BU and HM are both part of a class of drugs which contain polar sulfonic acid ester leaving groups (Figure 1.1). Both exert the majority of their toxicity through irreversible alkylation of DNA forming ICLs. HM being a derivative of BU, was designed to improve the antitumour efficacy of the parent compound by the introduction of a more polar leaving group.[19] HM demonstrated increased efficacy against leukemia cell lines that were relatively resistant to BU.[19,20] After decades both medications still find application as chemotherapeutic agents for the treatment of CML with the study of their repair mechanisms remaining an active area of research.
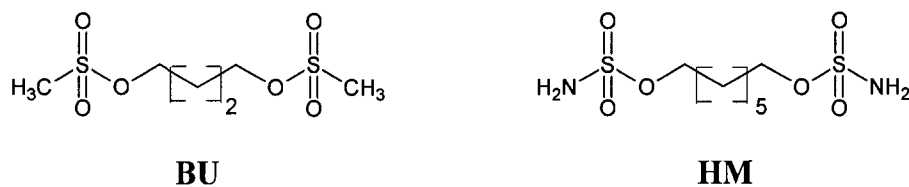
Figure 1.1 Structures of sulfonic acid ester cross-linking agents busulfan (BU) and hepsulfam (HM).

1.2 Origins and Mechanism of Action of Alkylating Agents

DNA undergoes various types of damage as a result of alkylation which involves nucleophilic attack upon electrophilic atoms. It is generally observed that electrophiles preferentially add at ring and exocyclic nitrogens, and to a lesser extent exocyclic oxygens, resulting in alkylated or Schiff base derivatives.

In vivo there are three major sources of alkylation which originate from exogenous and endogenous sources.[21] The first is endogenous enzymatic methylation in a sequence specific manner for control of various cellular processes that typically results in a relatively minor perturbation of the DNA helix structure.[21,22,23] The second source of DNA alkyl lesions can be generated from endogenous compounds such as acrolein and malondialdehyde.[3,21] Alkylation with these compounds can result in the formation of a MFA or BFA where the BFA can either form a cyclic adduct or an ICL.[3] The third mechanism of DNA alkylation involves direct attack of nucleobases on small electrophilic exogenous compounds such as nitrogen mustards, sulfonates, and nitroso compounds which are present in the environment as pollutants or employed as chemical warfare and chemotherapeutic agents.[21]

BU and HM are BFAs belonging to the previously mentioned third class of alkylating agents containing two reactive sulfonic acid ester moieties at either end of a linear alkyl chain of four and seven carbons respectively producing two electrophilic carbon atoms at either end. Both BU and HM's primary mechanism of action to exert a cytotoxic effect is to prevent DNA strand separation thus precluding transcription and/or replication through the formation of a covalent bond between the drug and the N^7-position of two 2'-deoxyguanosines with no observed reactivity towards other nucleosides (Figure 1.2).[24] It is

8

thought that the increased nucleophilicity of the N^7-position can be attributed to high electron density produced by base stacking and charge transfer within duplex DNA.[25] High nucleophilicity coupled with the N^7 lone pair of electrons being freely accessible in the major groove could account in part for the observed selectivity.[24]

R = deoxyribosyl linkage to DNA backone
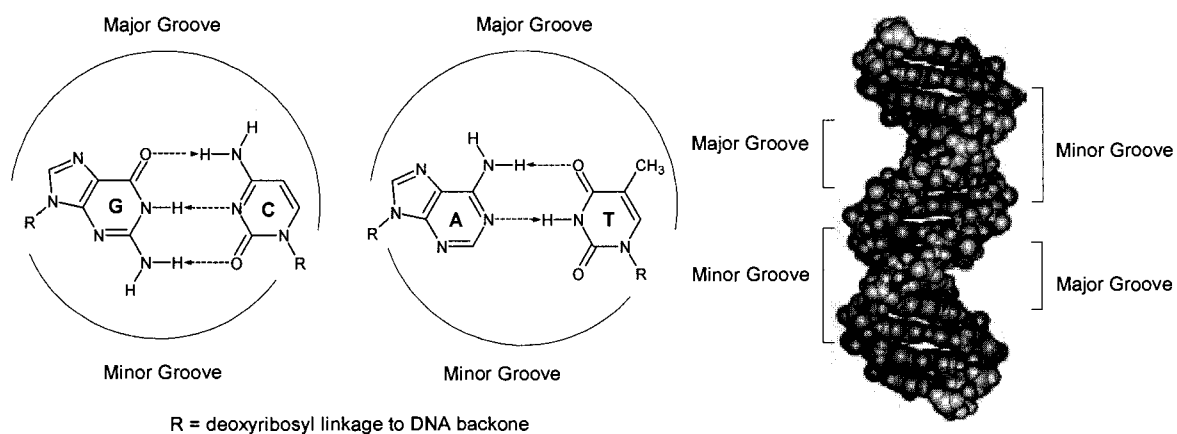
Figure 1.2 Representation of the major and minor grooves within B-form DNA as viewed at the base pairing level. Grooves are defined with respect to the glycosyl linkage of each base to its respective 2'-deoxyribose.[26]

Cross-link formation through bifunctional alkylation of the major observed N^7-2'-deoxyguanosine adduct can arise either by reaction with two adjacent 2'-deoxyguanosines in the same DNA strand or with 2'-deoxyguanosines in opposite strands of the sequence. Initiation of cross-link formation occurs when one of the electrophilic α-carbons to the sulfonic acid ester of BU or HM undergoes a second order reaction leading to nucleophilic attack by DNA resulting in loss of the sulfonic acid ester moiety which is a good leaving group (Figure 1.3). This leads to formation of a covalent bond with the N^7-position where positive charge is dispersed throughout the heteroaromatic ring. With the first alkylation event in place, the 2'-deoxyguanosine mono-adduct can then react with a variety of nucleophilic compounds in the body undergoing a second nucleophilic attack due to the bifunctional nature of BU and HM. The second sulfonic acid ester can be hydrolyzed to an

9

alcohol rendering monofunctionality producing no further reaction. Alternatively, cross-links can form between proteins or intracellular thiols resulting in linkages to DNA.[19] A second 2'-deoxyguanosine on the same or opposite strand of DNA can also be alkylated resulting in the formation of an intra or interstrand cross-link respectively.
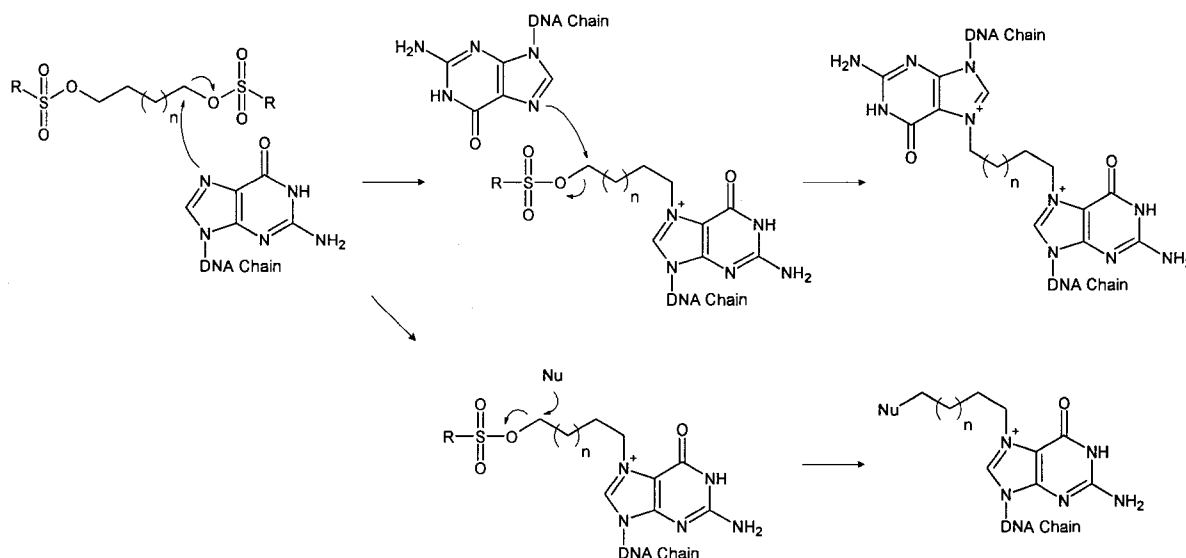
Figure 1.3 Proposed mechanisms of alkylation of 2'-deoxyguanosine by BU (n=1, R=CH$_3$) and HM (n=4, R=NH$_2$) followed by cross-linking of DNA or attack by a nucleophile (protein, water, or other biomolecule).[24]

Overall, the reactivity of methylating agents towards 2'-deoxyguanosine is well known. It has been shown that the N^7-position of dG and backbone phosphate oxygens account for about 85% of all methylation products.[27] Despite the major lesions formed by small electrophilic alkylators (including methansulfonic acid) being N^7-MedG and phosphate backbone adducts, some studies have correlated electrophilic alkylators' carcinogenicity with the number of minor O^6-alkyldG and O^4-alkyldT lesions formed.[28,29] Although not reported for BU or HM, potential remains to elicit a toxic effect through, not only N^7-alkyl-N^7 ICL formation, but also through more stable O^6-2'-deoxyguanosine or O^4-2'-deoxythymine adducts.

10

1.3 Derivatization and Stability

Medications that alkylate at the N^7-position such as BU and HM form unstable adducts which invariably decompose during purification. To prevent decomposition via a depurination mechanism which results in the loss of the actual cross-link itself leaving an abasic site, an alkaline treatment is required. The use of alkaline derivatization results in a formamidopyrimidine (FAPY) derivative of increased stability making purification possible. Base treatment can be performed with methods involving NaOH or piperidine methods, each having different mechanisms, conditions and thus outcomes. Treatment with NaOH at pH 13 at physiological temperature results in an addition of a hydroxide ion to the C^8-position of an N^7-alkyl-2'-deoxyguanosine resulting in a ring opening yielding FAPY which is more stable than the parent compound. This stability is demonstrated by experiments in the absence of base at elevated temperature (100°C), depurination approaches completion within three minutes whereas NaOH treated samples under identical conditions experience little decomposition.[30] The piperidine treatment method relies on the strong nucleophilicity of the base to form an iminium ion (Figure 1.4). This results in elimination of the 2'-deoxyribose followed by the formation of a Schiff base which then causes strand breakage via β-elimination of the phosphates which is not observed with NaOH to any significant degree.[30] Neither method produces addition to, or decomposition of, non-alkylated 2'-deoxyguanosines.[24] Furthermore, base treatment with temperature elevation can serve to act as confirmation of N^7-addition as the labile glycosidic bond of N^3-adducts of purines have not been demonstrated to be stabilized by base and as such decompose at high temperature.[31,32,33] Similarly the relatively stable glycosidic bonds of C^8 adducts of purines have not displayed any increased stability or chemical modification by

11

base treatment versus neutral pH when heated.[32,34,35] As such base stabilization of the glycosidic bond would not be expected for alkylations at positions other than the N^7-atom.
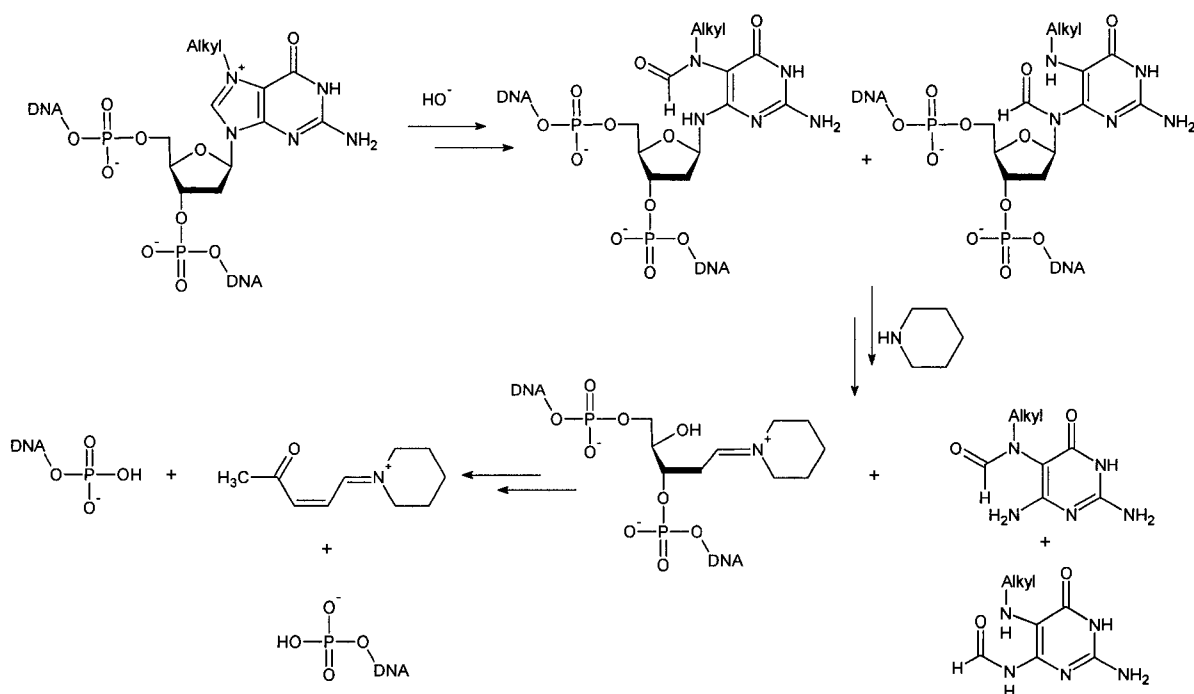
Figure 1.4 Sequential addition of NaOH to N^7-alkylated 2'-deoxyguanosine to form both the N^7 or N^9 FAPY derivatives of DNA. Subsequent displacement by piperidine catalyzes β-elimination of both phosphates from the sugar.[31,30] NaOH will only add to N^7-alkylated 2'-deoxyguanosine whereas piperidine results in strand breakage at N^7-alkylated positions aiding in the determination of the site of alkylation in short oligomers. A small population of adenines are also alkylated at N^3 but they are unreactive towards piperidine and do not result in strand breakage.

Elution patterns demonstrate the formidable challenges in separating ICL containing duplexes from parent duplexes.[24] The potential to form linkages between RNA is possible but mention is largely absent from literature due to short lifetime of mRNA and tRNA within the cell with DNA being the major target of BFA to form an ICL. To determine the position of alkylation for HM, guanosine was treated with excess alkylating agent followed by RP-HPLC analysis.[24] Isolation of ribonucleic acid derivatives was possible due to the increased stability to depuration as a result of the electron withdrawing nature of the

2'OH which destabilizes carbocation formation (Figure 1.5). Derivatization of N^7-guanosine adducts to FAPY analogues can be clearly discerned based on spectrometric experiments which consist of observing the reversibility of the UV signature as a function of increasing pH conditions and mass-spectral analysis.[24] Under high pH conditions it is observed that the UV spectral signature of N^7-alkylated guanosine undergoes an irreversible base-initiated transition in UV signature whereas guanosine UV spectra is completely reversible once brought back to neutral pH.[36] This is accounted for by addition of a hydroxide ion to the C^8-position resulting in ring opening to yield the FAPY derivative which processes a prominent shoulder centered around 280nm and a bathochromic shift of the absorption peak of 2'-deoxyguanosine.[24] The molecular weights of the derivatives can then be confirmed by mass spectral analysis.
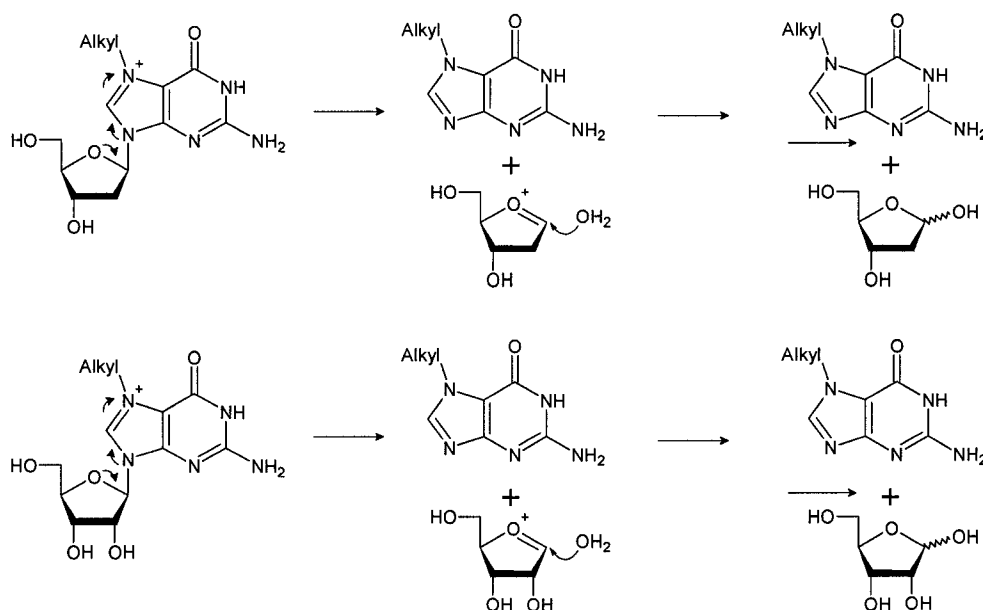
Figure 1.5 Depurination of an alkylated 2'-deoxyguanosine and guanosine at the N^7-position via a carbocation intermediate resulting in abasic sites in DNA and RNA. The presence of the 2'OH alcohol in ribose sugars greatly impedes the formation of the carbocation resulting a greatly diminished rate for the depurination reaction in RNA in contrast to DNA.[30,31]

1.4 Cross-link Position and Identification

The connectivity and three dimensional shape of a segment of DNA containing an ICL is of considerable importance with respect to enzymatic recognition and repair. Determination of the site of adduction, the orientation of a cross-link, degree of bending and helical unwinding have as of yet unknown effects on recognition of and repair of an ICL. Confirmation of the aforementioned parameters can be performed in a variety of ways including chemical derivatization, BFA treatment of comparative sequences, and biophysical studies.

More traditionally, identification of cross-link location has been performed via a variant of the Maxam-Gilbert reaction which relies on hydrolysis of a terminally ^{32}P radiolabeled duplex (Figure 1.4). N^7-FAPY-2'-deoxyguanosine adducts are generated using piperidine inducing strand breakage at N^7-adduct sites.[30,31] Derivatized samples are then purified using polyacrylamide gel electrophoresis (PAGE) and compared to standards to elucidate major cleavage sites which are indicative of adduct formation at specific sequences.[30]

Support for the regioselectivity of the HM was evidenced from reaction of oligonucleotides of identical molecular weight and identical net base composition containing exclusively 5'-GC-3' or 5'-GXC-3' (where X is any nucleotide) sequences separated by AT tracts.[24] It was observed after RP-HPLC purification that individual fractions analyzed using denaturing PAGE of the 5'-GC-3' eluted as a single band. Different 5'-GXC-3' HPLC fractions eluted as two bands leading to the conclusion that HM reacts with 5'-GXC-3' sequences to produce ICL in the synthetic DNA studied (Figure 1.6).[24] Mass spectral analysis also displayed masses corresponding to a depurinated dimer

14

in the 5'-GXC-3' sequence. In contrast, it has been observed that BU tends to induce primarily DNA intrastrand cross-links whereas HM induces predominantly ICL.[18,37] The exact orientation of lesion formed with BU remains unknown however it is thought that adducts between both strands remains the primary mechanism of cytotoxicity as in HM.[24]

5'-GXC-3'   **5'-GC-3'**

| 36 | 35 | 34 | 33 | 32 | 31 | | 36 | 35 | 34 | 33 | 32 | 31 |

HPLC Fraction Number
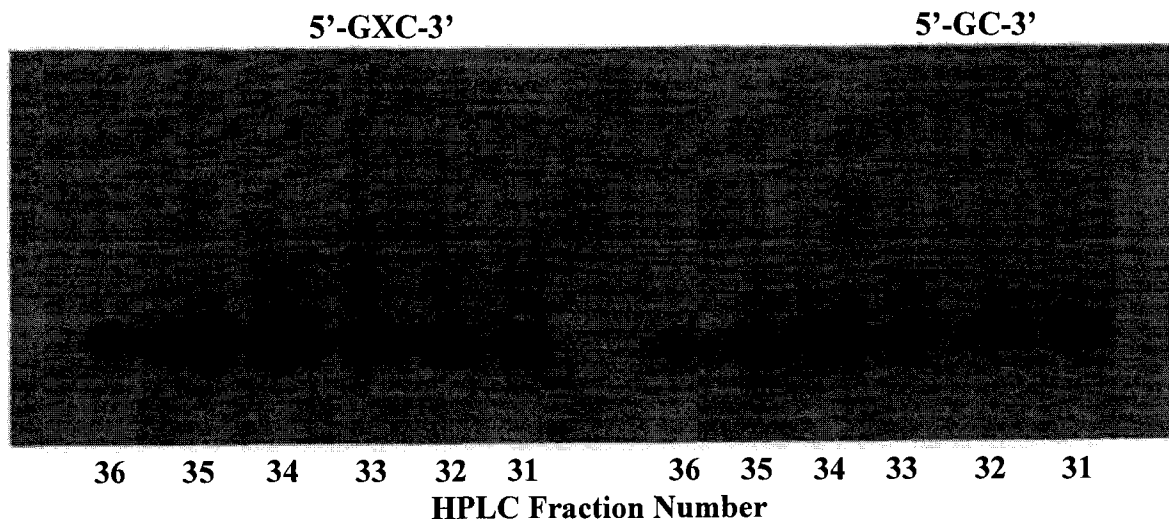
Figure 1.6 Denaturing PAGE comparing HPLC fractions from HM treated oligonucleotides.[24]

The extent of bending and helical unwinding induced by an ICL can be determined in a variety of ways. Conventionally, ligation experiments employing one-helical turn tracts have been employed to measure the shift in mobility via nondenaturing PAGE by an additive effect exaggerating the helical perturbation induced by a single ICL.[38] The amplified distortion impedes mobility within the polyacrylamide gel against a standard acting as a measure of the bending angle. The change in mobility can then be used to calculate the bending angle induced by a single lesion in a one-helical turn segment using the empirical relationship developed by Koo and Crothers.[39,40] This approach can be laborious as ligation conditions must be optimized often for each DNA duplex studied to prevent over or under ligation.[38] Under ligation can result in chain lengths too short to observe a significant change in the bulk three-dimensional properties producing little

15

change in mobility. Over ligation can result in DNA that requires excessive run times with little resolution.

More recently as higher field NMR instruments have been developed, 2D-experiments can be performed on short DNA duplexes of defined sequence. Most studies conducted on short one helical turn oligomers (11 to 13-mers) can be conducted at 500-600 MHz (^{1}H basis), combining ^{1}H , ^{13}C, and sometimes ^{31}P experiments in conjunction with NOESY.[41] Many studies are performed via a comparative analysis of a control duplex which is compared to an adduct containing duplex.[42-45] Backbone torsion angles can be discerned which can then used to develop a model consistent with the observed structural information. From NMR data the regioselectivity of a cross-link, the amount of helical unwinding and the degree of bending induced in the strand by a cross-link can be elucidated. Valuable insight is gleaned from bending and helical unwinding characteristics which have been deemed some of the most probable features allowing for enzymatic recognition of ICL and other DNA damage.

X-ray crystallographic diffraction studies have proved their value from the time that Watson and Crick elucidated the double helical structure of DNA more than half of a century ago.[46] Despite the long history of X-ray studies in nucleic acids NMR remains a more accessible and commonly used technique. Even with a pure sample, there remains a significant challenge to obtain a crystal of high quality allowing for meaningful resolution. Whether or not a sample will crystallize is highly sequence and environment dependent.[47] Additionally to obtain meaningful information, the DNA duplex containing an ICL must be of a single orientation (e.g. 1,3). Crystals of insufficient quality tend to result in low resolution diffraction leading to ambiguities in structural elucidation. The importance of

obtaining both NMR and X-ray data in establishing hydrogen bond structure remains as O^6-MedG·dT mismatch pairs have been shown to give conflicting alignments from the liquid to crystalline state (Figure 1.7).[21,43,48] The benefit remains that the high resolution of x-ray structures provides information to aid in the establishment of the 3-dimensional structure of a modified oligonucleotide. This information is inherently useful for more sophisticated studies involving repair enzymes.

Solution   Crystalline
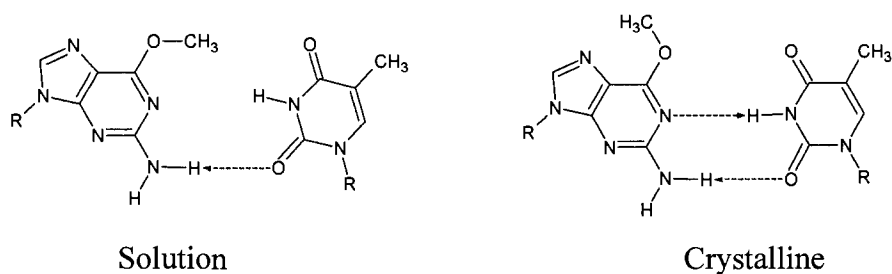
Figure 1.7 Differing orientations of an O^6-MedG·dT mismatch pair dependant upon the method of analysis in both solution (NMR) and crystalline (x-ray) states.[21]

Conformational variability in FAPY derivatives from the clinically relevant N^7-adduct results in unknown variables which potentially affect molecular repair studies. Insights derived from repair studies may not translate to the clinically relevant lesion formed by BFAs. There are significant differences in stability and structure of rigid purine ring systems to FAPY derivatives which exist as a mixture of structural and rotational isomers.[30] Despite a significant increase in stability with FAPY derivatives over the parent N^7-adducts, FAPY are still unstable at room temperature.[24] It has been demonstrated that N^7-FAPY-2'-deoxyguanosine adducts exist as mixture of structural forms with the formyl group located at the N^7 or N^9 position (Figure 1.4).[49-51] Further variability exists in the form of rotational isomers where the formyl group is either *anti* or *syn*.[49-51] The altered ring system produces notably different structural properties from the parent adduct resulting in ambiguities as to the relevance of the recognition and repair studies.

17

1.5 The Effect of Tautomeric and Resonance Forms on Base-Pairing Geometry

All four natural nucleobases possess tautomeric forms which are typically depicted in the more stable lactam form which is the basis for Watson-Crick (WC) bonding. Each structure can also be drawn in its respective enol tautomer for which non-standard base pairings can be shown.[47] The pK_a of the N^1-hydrogen atom of 2'-deoxyguanosine is ~9, as such a modest degree of ionization could exist at physiological pH.[47] The N^1 deprotonated form of 2'-deoxyguanosine has an altered hydrogen bonding pattern which corresponds to the pattern of an O^6-alkylated dG (Figure 1.8). It is known that O^6-alkylated 2'-deoxyguanosine can lead to non-standard base pairing with dT which can result in dGdC → dAdT transitions.
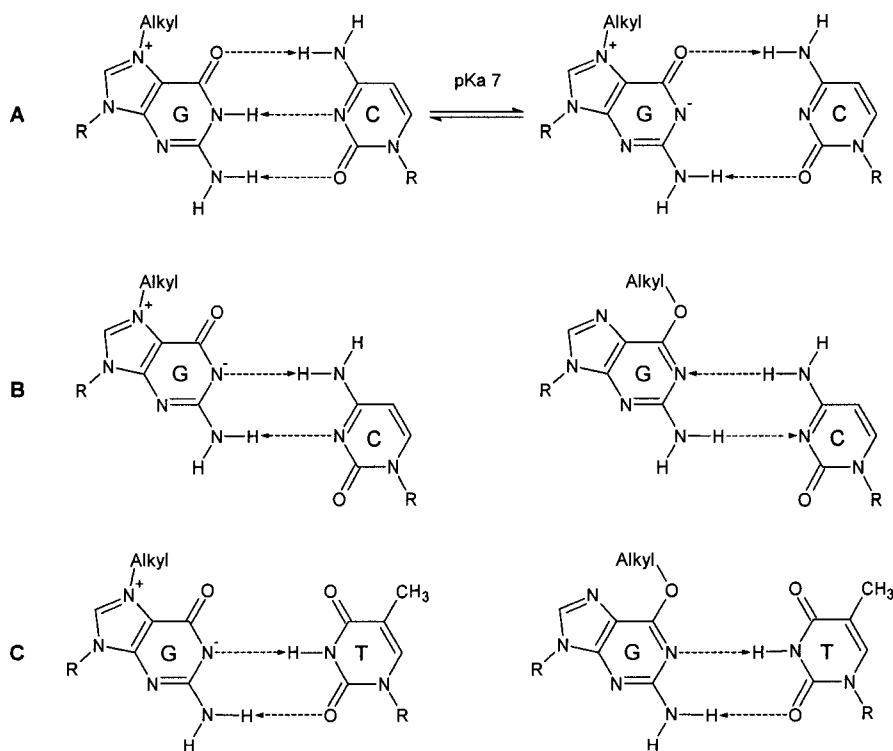
Figure 1.8 (A) Ionization of an N^7-alkyl-2'-deoxyguanosine and possible pairing of with cytosine[52] (B) Base pairing of N^7 and O^6-alkylguanosine and wobble base pairing of with cytosine[42] (C) Base pairing of N^7 and O^6-alkylguanosine mismatch with thymine having B-form geometry.[43]

Evidence for the number of minor O^6-alkyldG and O^4-alkyldT lesions and not the total number of alkylations being the source of carcinogenicity can be derived from the ability of AGT to lower the number of mutations produced by an alkylating agent.[28,29] Studies with methylating agents such as procarbazine form a variety of adducts with O^6-MedG being among the most cytotoxic.[53] These studies do not necessarily prove that an N^7-MedG adduct cannot be mutagenic however they do provide the impetus for investigation of the mutagenic or cytotoxic potential of O^6-alkyldG cross-links. NMR and X-ray studies also provide insight as to the potential geometry of the bonding structure around a 5'-GXC-3' ICL if base pairing at each site in a 1,3 orientation is treated as a mono-alkylation. Structural data also presents the opportunity to augment susceptibility to repair and the toxicity studies of O^6-MedG·dT base pairing.

The possibility of the enol tautomer playing a significant role in dG·dT mismatch base pairing is still under debate. Both X-ray and NMR studies have been unable to find evidence for the existence of tautomerism.[54] X-ray structures with dG·dT mismatches support wobble base pairing while the possibility exists for only one form crystallizing.[55] ^{1}H NMR experiments can detect downfield protons but these resonances cannot confidently be used to determine hydrogen bonding patterns.[47] Other studies using ^{31}P NMR and similarities in NOESY spectra indicate dG·dC, O^6-MedG·dT,[43] and O^6-EtdG·dT[44] containing duplexes have B-form conformations of normal WC geometry.[21] In contrast, O^6-MedG·dC[42] and O^6-EtdG·dC[45] containing duplexes have spectra similar to dG·dT having wobble geometry at the lesion site.[21] The wobble geometry of O^6-MedG·dC[42] and B-form conformation of O^6-MedG·dT[43] potentially has an impact on recognition and repair of both lesions.

The existence of a shift in the hydrogen bonding pattern of N^7-MedG has been suggested as a potential mechanism for mutagenesis with thymine.[46,47,56,57] The presence of an N^7-adduct, although addition at a position not directly involved in hydrogen bonding, can disrupt normal WC base pairing motifs. It has been reported that N^7-MedG has only one exchangeable proton (N^1-H) that has a $pK_a \sim 7$ which can result in the loss of a proton under physiological pH.[56] A loss of the N^1-proton would result in a change to the hydrogen bonding pattern with the N^1-position becoming a hydrogen bond acceptor (Figure 1.8). NMR characterization of duplexes containing a single N^7-MedG adduct have shown relatively little distortion from the parent conformation.[58] Alkylation of O^6-dG has similarly been shown to cause only a small conformational change to the helical conformation, as evidenced from torsion angles of the phosphate backbone.[42] The small amount of distortion induced in both lesions suggests that the deprotonated N^7-MedG may make a significant contribution to the hydrogen bonding character for the N^1-position from a donor to an acceptor. Prevention of the traditional WC base pair alignment inducing similar minor distortions in the respective helices leave the potential for similar repair pathways due to a comparable amount and type of helical distortion induced.

It is unlikely that a duplex containing a BU or HM ICL at either the O^6 or N^7-position would undergo an O^6-alkyldG·dC transition to form an O^6-alkyldG·dT mismatched base pairs. Mismatch base pair transitions (O^6-alkyldG·dC $\rightarrow$ O^6-alkyldG·dT) are typically the result of faulty polymerase incorporation in mono alkylations. BU and HM treated cells have a peak ICL formation after 12 hours[21] as such mono-adducts would be present for times which could allow a polymerase induced mismatch transition to occur however this can be excepted to be a minor occurrence. A more likely impact on repair would occur

from a false recognition and attempted repair of the O^6-alkyldG·dC lesion by mismatch repair (MMR) as the geometry more closely resembles a dG·dT mismatch.

1.6 Resistance to Therapeutic Alkylating Agents – DNA Repair

The mechanisms involved in drug resistance for chemotherapeutic agents such as BU and HM are a multifactorial combination of altered drug delivery and uptake, decreased drug activation, increased metabolic inactivation, active drug efflux, enhanced repair of DNA, and overlapping repair systems.[59] The interplay between overlapping repair mechanisms can be seen in the therapeutic effect of methylating drugs and MMR activity which adds a clinically relevant dimension to AGT repair of O^6-MedG. It has been shown in MMR deficient cell lines depleted of AGT with O^6-benzylguanine (O^6-BG) results in a failure to sensitize cells to methylating agents. Clinically achievable concentrations of the methylating agent temozolomide produce cytotoxic effects only in MMR proficient cells with low AGT activity.[60] The O^6-alkylated dG adducts formed do not elicit a cytotoxic effect in the absence of MMR even if they are mutagenic.[53] Cell lines that are AGT deficient or depleted and have normal MMR activity display heightened sensitivity to O^6-alkylating agents.[53,60] It is thought that this heightened sensitivity is the result of strand breakage induced by MMR recognition.[60] Cell lines with high levels of both MMR and AGT exhibit drug resistance mainly, but not exclusively, through up-regulation of AGT expression.[60] Exposure to methylating agents and O^6-BG has been shown to favour development of AGT-independent mechanisms of resistance.[60] As such, MMR mutations seem to override the AGT mechanism of resistance as a cell's failure to detect lesions may attenuate their toxicity. If AGT is incapable of recognizing a BU or HM O^6-dG-alkyl-O^6-dG ICL the potential exists to identify alternate mechanisms of drug resistance involving lesion recognition by MMR. Minor O^6-alkylation products of BU and HM, if present, may be critical in the development of drug resistance to BU and/or HM.[29,53]

Classically it is assumed that a cell with a gene that has been silenced which is involved in genetic repair would be less fit as the vigour of an organism is dependant on the maintenance of its genome. Several genes and enzymes that play a critical role in MMR have been identified for which many reviews are available.[61-63] It is clear that loss of MMR results in destabilization of the genome and causes high mutation rates in coding and non-coding regions.[64,65] In a variety of human cancers, the loss of DNA MMR has been observed which leads to drug resistance.[66-70] By impairing the ability of a cell to detect DNA damage, the ability to regulate cellular metabolism and initiate apoptosis is diminished.[53] Apoptosis can also be attenuated indirectly by increasing the mutation rate throughout the genome.[53] MMR-deficient cells have been reported to be resistant to a variety of chemotherapeutic agents that form ICLs[3] including BU which has been demonstrated in clinical trials to induce a reduced response *in vivo* in tumour cells.[53]

The importance of MMR in resistance to BU is in part due to MMR enzyme's ability to recognize and bind to various types of adducts in DNA as well as mismatches.[53] As opposed to being a primary repair protein, MMR appears to be more involved in recognition of specific types of DNA damage.[53] Although MMR does not appear to have the capacity to recognize O^6-MedG[71], it can however detect O^6-MedG·dT mispairing after incorporation from a replication cycle.[72] The current model of MMR recognition for a O^6-MedG·dT mismatch involves excision of a small DNA segment around the mismatched thymine. Thymine is then reincorporated triggering another round of repair resulting in an infinite loop of repair and excision which eventually has the potential to result in a double strand break triggering apoptosis.[53] This predicts that loss of MMR confers resistance to

O^6-alkylators as the cell does not attempt repair. This is further supported by the fact that chromosomal transfer reinstituting MMR function eliminates methylation tolerance.[73]

The minor structural alteration of HM from BU results in differences in the mechanisms of decomposition which may account for some of the divergence in therapeutic effect once in the body.[18] BU undergoes hydrolysis via nucleophilic attack by water to give 4-methanesulfonyl-oxybutanol, which subsequently undergoes an intramolecular displacement to yield tetrahydrofuran and methanesulfonic acid.[74] In contrast, no evidence of such an intramolecular mechanism has been observed with HM.[75] Instead, HM formed the intermediate 1,7-heptanemonoylsulfamic acid ester prior to complete hydrolysis to 1,7-heptanediol.[75] The lack of an intermolecular reaction with HM may help explain the differing pharmacokinetics and blood half-life of HM over BU. HM displays a 16 hour blood half-life versus 2-3 hours for BU which may also play a role in the observed difference of HM to show efficacy against BU resistant tumours.[76,77]

Overexpression of Glutathione-S-transferase (GST) is regarded as a significant mechanism by which human tumour cells become resistant to chemotherapy with cytotoxic alkylating agents.[78] Evidence also exists for increased sensitivity, and thus a possible mechanism of resistance development, for cells exposed to BU and HM under L-buthionine sulfoximine (BSO) depleted glutathione (GSH) conditions.[19] Increased reactivity of BU over HM exists in the presence of GST with BU reacting with GSH to form a sulfonium ion metabolite.[77,79] Early studies demonstrated HM was unable to react with GSH in either the absence or the presence of GST in blood using exposure times of only 1 hour in buffer and rat liver cytosol.[19,80] Studies by Pacheco *et. al.* showed that peak HM induced ICL formation occurred after a 2 hour drug treatment followed with washing by centrifugation

and subsequent incubation for 12 hours followed by alkaline elution.[18] Later studies by Davidson and Colvin using human as opposed to rat liver cytosol proteins, demonstrated HM forms conjugates with GSH when analyzed by mass spectrometry after a 2 hour drug exposure and a 12 hour incubation time.[19] The multiple cell lines tested showed no correlation to HM with basal levels of GSH but did show increased sensitivity with BSO depletion of GST in a dose-dependant manner.[19] BU's susceptibility to the GSH detoxification mechanism contrasts to the slower kinetics of HM to react with GST may be one dimension in the observed lack of drug cross-resistance of BU tolerant tumours.[18,19] It remains that increased GST expression correlates positively with resistance to BU and to a lesser extent HM.

1.7 Direct Repair

Currently more than 100 alkyltransferases are now known with crystal structures available for three of the enzymes (Figure 1.9).[81,82] Human O^6-alkylguanine-DNA alkyltransferase (AGT) directly reverses endogenous alkylation at the O^6-position of 2'-deoxyguanosine and O^4-alkylthymine by transferring the O^6-alkyl group to an active site cysteine (Cys145) in an irreversible and stoichiometric reaction.[83,84] The presence of a stable O^6-2'-deoxyguanosine lesion is promutagenic, having the potential for mispairing with thymine giving rise to a transition mutation during replication (i.e. C→T transition resulting in GC→ AT after subsequent replication).[82] It is unknown if the altered WC binding pattern of an O^6-alkylated BU or HM ICL will result in mispairing if cellular attempts at repair employ lesion bypass. It has been demonstrated that both (GSH) and AGT confer resistance to alkylation chemotherapies making GSH and AGT active anticancer drug targets.[19,85]
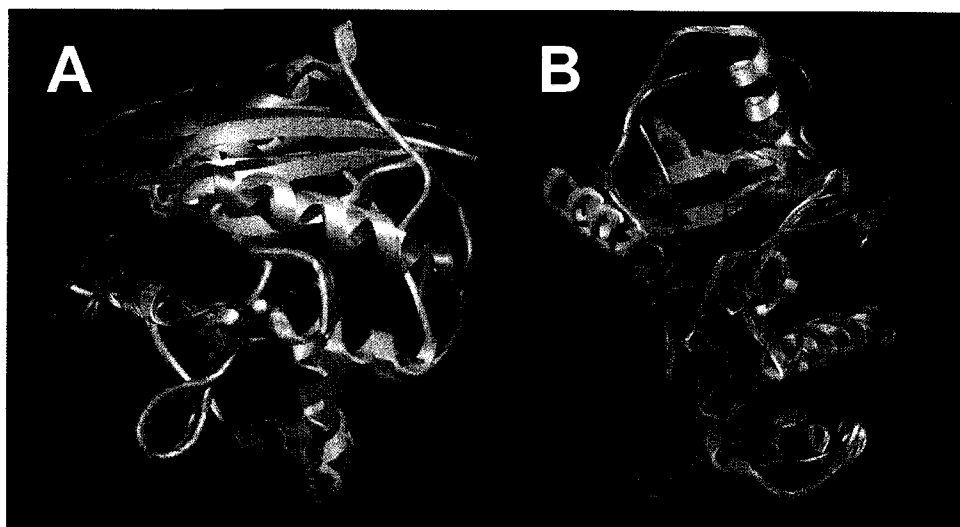
Figure 1.9. The three dimensional structure of AGT which is highly conserved in three kingdoms of life despite low sequence homology as revealed by a structural overlay of human (blue), *E. coli.* (yellow), and *P. kodakaraensis* (purple) AGT enzymes. The active site cysteine, Asn-hinge, and arginine finger sidechains are also shown (sulphur yellow, oxygens red, nitrogens blue, carbons green). (B) is rotated 90° about the vertical axis relative to (A).[82]

The conformational and structural similarity between numerous AGT class enzymes have been well documented representing examples from 28 species as diverse as Eubacteria, Archaea, and Eukarya.[82,84,85] Minor differences in structure have been reasoned to account for experimental results demonstrating varying rates of activity between different species.[84] Variability in activities has not lead to evidence supporting differing mechanisms of action divergent from use of amino acid residues at non-conserved positions.[86] Conformational studies using NMR data for the *Escherichia coli* Ada-C protein which aids in the repair of pre-mutagenic O^6-alkylguanine lesions in DNA revealed only a small conformational change is experienced upon DNA binding.[82] A similar DNA binding model also has been suggested for the human homologue from high resolution AGT crystal structures allowing for the use of Ada-C data for insight into human repair mechanisms (Figure 1.9).[82,85]

Previous work has reported that BU does not elicit toxicity via alkylation at the O^6-position in a way that can be influenced by AGT repair.[6] This contrasts to stem cell lines with overexpression of GST, specifically the MGSTII form of GST, having increased survival rates when exposed *in vivo* to BU and HM.[19,78] Studies employing O^6-BG served to potentiate chloroethylating agent 1,3-bis(2-chloroethyl)-1-nitrosourea (BCNU) toxicities in normal human bone marrow cells.[6,87,88] Similar treatment of cells with O^6-BG left BU activity unaffected by depletion of AGT.[6,87,88] This serves to further advance the claim that GST is the only known route for elimination of BU.[78] Efficacy enhancement of chloroethylating agents in the presence of O^6-BG but not BU leaves the potential for N^7 alkylation by BU or O^6-alkylation that cannot be repaired by AGT.[6,89] Given the homology in structure, it is likely that BU and HM share similar mechanisms of addition and repair.

27

The mechanism(s) of action for BU and repair mechanism(s) for BU or HM, to date remain unknown.

Crystal structures of Ada-C and human AGT serve to provide a comprehensive framework for structure activity relationship studies (SAR) (Figure 1.10). It has been determined that the molecular basis for recognition is extrahelical flipping of the alkylated nucleobase into the binding pocket.[82,85] Subsequent release of repaired DNA for methylated O^6-MedG lesions results when AGT undergoes a conformational change allowing dissociation from the repaired duplex.[82,85] It has generally been found that substituents of increasing steric bulk at the O^6-position react with a decreasing rate.[82] Adducts having decreased rates for repair by AGT due to increased steric bulk has suggested other repair systems such as nucleotide excision and MMR systems play an important role for the removal of derivatives other than O^6-MedG lesions.[62,63,90] To this end, studies have demonstrated that nucleotide excision repair (NER) seems to repair O^6-ethyl-2'-deoxyguanosine lesions *in vivo* in *Drosophila* germ cells.[90]
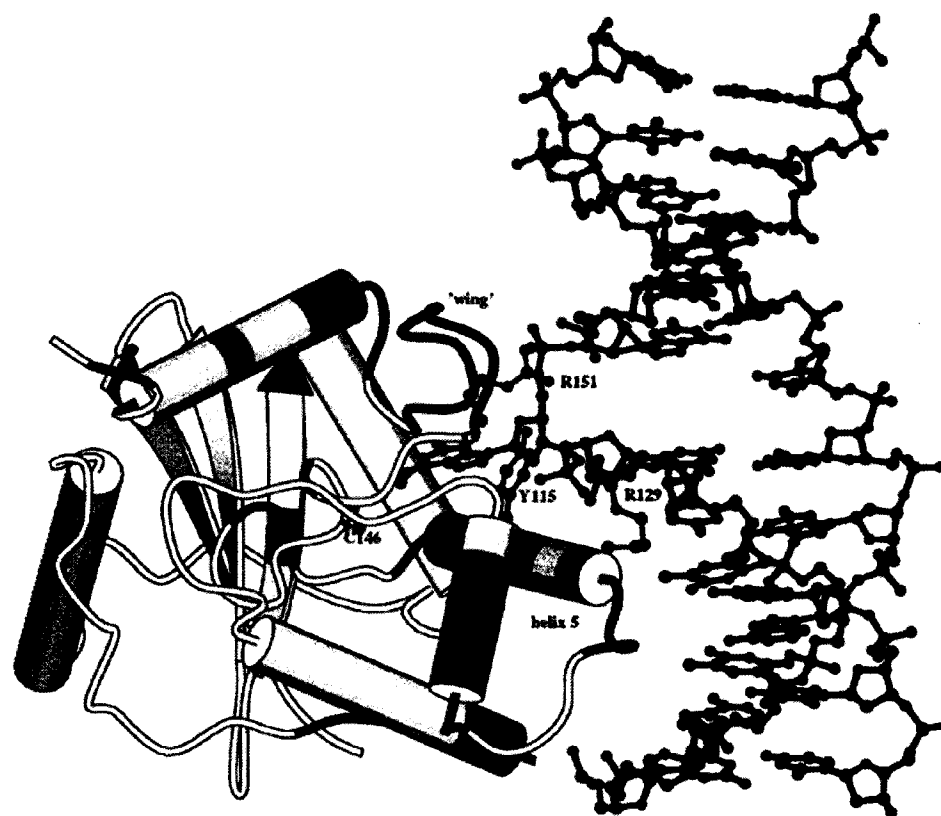
Figure 1.10. DNA binding model of Ada-C with O^6-MedG flipped into the active site through the opening between the recognition helix and wing. The orientation of the recognition helix of Ada-C with respect to the DNA is based on the CAP-DNA complex.[91] The colour scheme represents two categories of chemical shift movement observed for Ada-C backbone resonances upon DNA binding. Backbone regions of Ada-C most affected by DNA binding are coloured red, those less affected are coloured blue. Unaffected regions are coloured grey or white. Ada-C predominantly contacts a single strand of the duplex. Sidechains from residues showing potential interactions to the DNA backbone of the flipped strand are depicted. The exocyclic methyl group from O^6-MedG is shown as a blue ball in the active site, juxtaposed with the side chain Cys146 (blue).[82]

29

1.8 Synthesis of Modified Oligonucleotides for Repair Studies

Nucleic acids, their derivatives and corresponding polymers are among some of the most studied groups of compounds known to man. Synthesis of nucleic acids has resulted in a significant number of drugs currently in late phase III clinical trails or on the market.[91] Interest in nucleic acids has not been limited to the synthesis of monomeric analogues. Many oligonucleotides form an emerging class of potent anti-viral agents and treatments for of a variety of cancers and inflammatory disorders.[91] Automated oligonucleotide synthesis developed in the early 1980's by Ogilvie *et. al.* laid the ground work for the current ability to produce oligomers of DNA up to 150 nucleotides in length (Scheme 1.2). This access to specific sequences can then be amplified using PCR which has revolutionized the study of genetics. The study of cancer, being at its most fundamental level a genetic disease, has not been untouched by these developments. The study of molecular systems of repair is typically performed using segments of DNA which contain lesions from specific endogenous or exogenous damaging agents. To adequately study enzymatic DNA repair systems, sequences possessing modified bases having a single well defined modification at an individual site are required.

Many enzymes display a high specificity for the identity of the substrate that is to be bound making many assays using modified oligonucleotides highly sensitive to unmodified impurities. One common problem in both cell-based and non-cell based assays is the presence of an unmodified oligonucleotide mixed even at small percent composition with modified oligomers.[47] Enzymatic activity past a segment that does not have an ICL could lead to the ability of a cell to maintain cellular metabolism. This ability to bypass a site not containing a modification could lead to a poor correlation of an adduct to toxicity.[94] With

defined structures more meaningful repair mechanism studies can be conducted. These deviations from clinically relevant DNA adducts highlight the need for access to sequences of defined structure and native base composition accessible through synthetic methods. This has prompted the development of various methods to produce oligonucleotides containing ICLs by three major approaches.[47]

Scheme 1.2 Solid phase synthesis cycle employed for synthesis of DNA oligomers.[92] The solid support is functionalized with a long chain alkyl linker terminally attached to a nucleoside. Cycle entry is started by the introduction of TCA resulting in detritylation. The free 5'-alcohol is then rinsed with solvent to remove traces of acid. Coupling is initiated by the addition of the next phosphoramidite in the presence of excess tetrazole resulting in the formation of the phosphite triester backbone. Capping is performed using acetic anhydride to arrest further chain growth of any unreacted free alcohols. Oxidation is carried out with iodine in the presence of water to form the phosphate backbone. Detritylation is then performed in an identical manner to cycle entry at which point another phase of coupling can be undertaken. Deprotection of all functional groups and removal from the solid support occurs after treatment with 1:3 ethanol/ammonia hydroxide overnight.

31
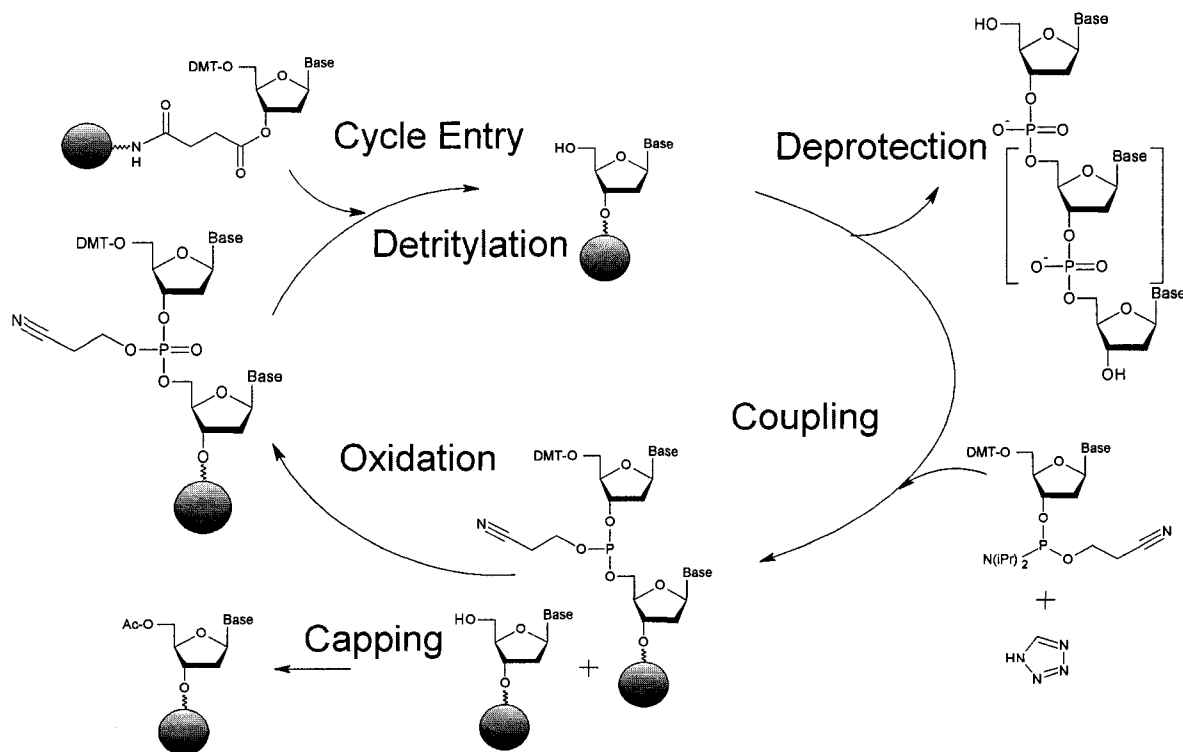
The classical method for alkylated oligonucleotide synthesis mirrors a cellular mutagenesis assay where small segments of native random sequence or chemically synthesized DNA of defined sequence are used.[47] This approach is commonly employed for chemistries that are incompatible with phosphoramidites or deprotection conditions such as N^7-guanyl adducts.[47] Access to many clinically relevant structures, particularly N^7-guanyl adducts, has been limited as the classical method suffers from multiple modification sites producing several types of damage. DNA is treated with excess alkylating agent resulting in a myriad of products of unknown structure and sequence comprised of monoalkylated, intrastrand and interstrand cross-linked sequences having low yields of the desired product (Figure 1.11).[24,47] Isolation of the desired adduct is at best problematic as the complex mixture generated containing a variety of different and/or multiple lesions may be inseparable (Figure 1.3, 1.6).[24] In fact, it has been observed in PAGE experiments/analysis of oligonucleotides having multiple alkylations elute as one band along with DNA control segments not treated with alkylating agent.[24] These results suggest that BFA in a different regiopositions are inseparable, making the presence of artefacts highly probable. Further complicating the interpretation of subsequent enzymatic studies, an alkaline treatment is commonly required for N^7-adducts previous to purification to form the FAPY derivative to prevent decomposition through depurination.
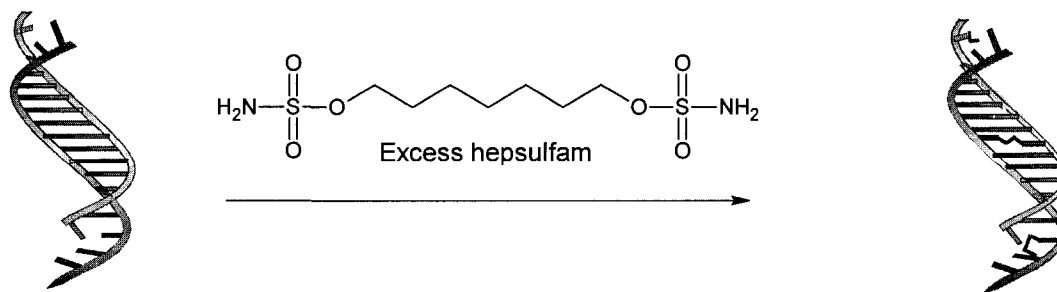
Figure 1.11. Post-synthetic strategy resulting in a mixture of mono-alkylation, intra and interstrand crosslink formation.[24]

The second method to prepare a DNA that is cross-linked occurs via a hybridization triggered reaction.[93] One requirement of the hybridization triggered method is the requirement for compatibility with standard phosphoramidite and subsequent deprotection chemistries. This has the advantage of a limited number of side reactions that are constrained to a specific area of a given oligonucleotide which can be separated by HPLC with a reduced chance for artefacts. This efficient and direct approach combines traditional nucleic acid synthesis techniques in solution to prepare dimers with solid-phase synthesis to produce DNA sequences of defined structure containing the cross-links in a known orientation. This approach still suffers from the potential for minor contamination given the difficulty in separating oligomers with and without modifications.

Covalent interstrand cross-linked segments of DNA that emulate lesions from BFA were first explored in the mid-1980's.[95,96,97] Since then several covalently linked structures have been developed containing mitomycin C,[98] psoralen,[99,100] trimethylene[101,102] or HM (Figure 1.1 & 1.12).[40] Unlike working with BU or HM on synthetically relevant scales using excess alkylating agent, synthesis of cross-links from non-toxic precursors offers the comparative advantage of being very safe to work with. A dimer containing an ICL is synthesized and incorporated into a DNA chain via solid-phase synthesis. This approach has the advantage of far fewer side reactions thus giving a much simpler purification and greater confidence in the structure of the final DNA duplex.

| Mitomycin C | Psoralen | Trimethylene |
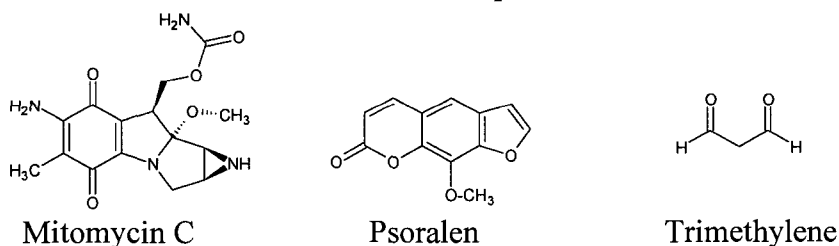
Figure 1.12. Therapeutics which have been mimicked through covalently linked duplex DNA structures which have been synthesized to emulate lesions formed by BFAs.[101,102]

33

1.9 Purification and Common Characterization Methods of Oligonucleotides

Oligonucleotide purification is typically performed with a combination of HPLC and PAGE. Oligonucleotides are routinely purified with separations possible by HPLC using ion exchange chromatography using a low composition of ACN (~10%, not exceeding 20%) and increasing NaCl gradient with longer oligomers eluting later than smaller ones.[103] Typically sufficient resolution can be achieved between oligomers up to 42 units in length differing by one nucleotide in length.[103] Once fractions are collected and combined the desired oligomer can then be desalted by loading onto a disposable C18 cartridge which is then eluted with water followed by methanol/water (1:1). The organic fractions containing the oligomer are collected and concentrated to yield the dried final compound. Individual nucleic acids and shorter oligonucleotides (up to 24-mers) can be separated using C18 reversed phase chromatography.[47] In order to improve peak shape and resolution gradients employing methanol are generally superior to those used with ACN. Volatile aqueous buffers (e.g. ammonium acetate) are typically used allowing lyophilization. With both methods it is desirable to use elevated temperatures (maximum <10°C of the lowest boiling point solvent) to reduce self association of the oligomer resulting in departures from ideal elution characteristics.

Purification via electrophoretic methods provides the highest level of sensitivity and resolution.[47] Visualization using UV detection or intercalating staining techniques (STAINS-ALL™) can be used on oligonucleotides to enable a more sensitive level of detection with an increasing length of the oligonucleotide. Preparative PAGE is an alternative to HPLC separation however it is not commonly used as it suffers from lower recovery, lower mass loading, significantly longer run times, and is more labour intensive.

As in preparative HPLC, resolution decreases when performing preparative PAGE highlighting the need for verification of the purity prior to further steps.

Nucleotide compositional analysis can be performed via a variety of techniques. Sequence composition can be performed via selective fragmentation creating a number collision induced products by ESI-TOF.[104] In addition, an enzyme digestion using snake venom phosphodiesterase (SVPDE) followed by hydrolysis of phosphate bonds using calf intestine phosphatase is performed. The digested nucleosides are then analyzed by RP-HPLC and relative integrations of the peaks eluting at known times are corrected for using the molar extinction coefficients. The corrected areas serve as a simple tool to ensure the observed ratio correlates with expected nucleic acid composition. This method suffers from a limited assessment of purity (~5-10% error) in the ratios that are present.[47] This limitation can be offset by the enzymatic digestion having the potential to provide confirmation that the cross-linked dimer has remained intact throughout the synthetic and characterization processes. The intact dimer is evidenced by the presence of a single late eluting peak, no mono-adducts eluting at times intermediate to the monomer and dimer, and accurate unmodified nucleotide composition ratios.

As is standard in the majority of modern chemistry, mass spectrometry (MS) is required to establish the molecular weight of a given oligonucleotide. Typically due to the large molecular weights of oligomers matrix assisted laser desorption/ionization – time of flight (MALDI-TOF) is frequently used for its ability to analyze molecules of high molecular weight. Properly calibrated MALDI-TOF instruments can deliver accuracy better than 1 part in 10^3 to 10^4 with the presence of a $[M + H]^+$ and $[M + Na]^+$ (or $[M - H]^-$) providing positive confirmation of oligonucleotide identity.[47]

1.10 Biophysical Analysis of Oligonucleotides

For approaching 40 years, circular dichroism (CD) has been used in the study of DNA conformation.[105] The pervasiveness of the technique is in part due to its minimal sample requirements, the non-destructive nature of the procedure, the ability to study DNA conformations in solution, and the ease and sensitivity of such measurements.[105] The value of CD spectra is apparent in their ability to distinguish between B-form or A-form conformations when compared to a standard. Deviations in the CD signature reveal qualitative information about possible alterations in conformation from native A or B-form structure of DNA. B-form structure can be characterized by the presence of a positive band at 280 nm and a negative band at 250 nm.[47] The presence of an ICL and the amount of bending or helical unwinding induced can be qualitatively measured as indicated by a large or small perturbation from the control signature. Resultantly CD is commonly used to detect changes in secondary structure.

The qualitative information derived from CD spectra can be used to ascertain some generalizations concerning specific regions of the UV spectrum as it relates to structure. The high variability in CD spectra is a result of the factorial nature of DNA sequence composition and the secondary and tertiary structures that result therein. Despite this variability, absorption profiles of a nucleic acid polymer are a function of nearest neighbour approximations at wavelengths above 210nm which can be used for compositional analysis.[106] Illustrating the sequence dependency of CD, the magnitudes of the major CD bands above 230 nm are not greatly changed upon denaturation.[105] Below 230 nm there can be a change in the magnitude around 220 nm of the band.[105] The lower wavelength regions display little information beyond a signature pattern due to the number of bonds that begin

36

absorbing below 200 nm. For meaningful analysis to be determined any adduct addition, dehydration or other alterations must always be compared against a standard.

Thermal melt (T_m) experiments are commonly used in DNA analysis to measure the change in stability of a given DNA duplex with respect to altered sequence or the presence of an adduct. The change in stability is measured as a plot of the normalized hyperchromicity at 260 nm against the temperature. As the temperature increases the annealed duplex begins to dissociate resulting in an increase in A_{260} referred to as hyperchromicity. Non-modified DNA duplexes produce sigmoidally shaped dissociation curves while less sigmoidal curves with large temperature range transitions generally indicate a disruption from normal WC base pairing in modified sequences. The 'melting point' of a given duplex is the point at which the duplex is annealed/dissociated or denatured in a 1:1 ratio as measured by a first derivative analysis of the T_m sigmoidal curve reaching a maximal rate of change at the inflection point. The observed hyperchromicity results in duplex DNA from the nucleobases being asymmetric with respect to electron density. When a nucleobase is excited with UV light a transition moment is created and the asymmetric bases couple with one another due to their van der Waals proximity. As the duplex dissociates under increasing temperature the base stacking becomes more disordered, diminishing the ability to transfer the induced transition dipole moment. This change in dipole transfer is observed as an increase in absorbance of ssDNA versus dsDNA. This change can be as great as 30% for some sequences.[107]

Chapter 2: Research Objectives

2.1 Design and synthetic strategy of O^6-alkyl linked dG dimers and ICL DNA duplexes

2'-Deoxyguanosine is the most common nucleobase to undergo modification under biological conditions.[5,26] This biological sensitivity translates to susceptibility to decomposition or side reactions under a variety of commonly employed synthetic conditions. Distinct reactivity exists at multiple positions on the purine ring making chemistry with the heterocyclic base a formidable challenge.[108] The presence of a tautomeric form in the molecule which shifts from a lactam ($pK_{a,N1} \sim 9$) to an enol form at high pH produces phenol type reactivity at the O^6-position. This phenol character enables a wide variety of aromatic nucleophilic substitution reactions for the purpose of protection or alteration of the reactivity of positions on the aromatic ring.[109] Replacement of the O^6-position oxygen typically requires extreme reaction conditions or that it be made a more effective leaving group. The N^1-position displays nucleophilicity being involved as a site of addition for multiple G-C cross-linking agents such as BCNU.[6] The N^2-exocyclic amine displays reactivity towards a variety drugs and is the basis for its need for protection, traditionally with isobutyryl chloride in pyridine.[3,26] The N^7 and N^9-positions are nucleophilic displaying unique electronic and steric properties resulting in preferential addition of ribose to the N^9-position. The C^8-position is an nucleophilic centre with a proton that is acidic enough to undergo exchange in deuterated solvents or be displaced by halogenating agents.[110] In addition, the presence of primary and secondary alcohols on the sugar moiety allow for orthogonal protection and removal of protecting groups. Careful selection of the conditions utilized is critical to avoid a multitude of unwanted side reactions.

Synthetically prepared 2'-deoxyguanosine cross-links can utilize deaza-heterocyclic bases to limit side reactions[97] or increase the stability of adducts formed at the C^7-position.[111] The formation of 7-iodo-7-deaza-2'-deoxyguanosine requires multiple synthetic steps in contrast to the readily available natural analogues. It has also been found that attempts to alkylate C^7-iodinated-7-deaza-2'-deoxyguanosine using palladium catalyzed reactions results in multiple side products.[111] Palladium reactions on halogenated guanosines have been shown to proceed in relatively high yield using mono-yne reagents. The requirement for a bifunctional alkyl linker to install an ICL is suspected to be the source of unwanted side-reactions. Difficulties encountered in this lab with palladium catalyzed reactions resulting in multiple side reactions, arduous purifications and exceptionally low yields led us to synthesize a cross-linked dimer via alternative methods.

Multiple alkylating agents display selectivity for either N^7 or O^6 with several alkylating at both positions.[90,112] For example, HM has been shown to form a number of N^7-alkylated products with 2'-deoxyguanosine including 1,7-bis(guanyl)heptane.[24] The synthesis of DNA duplexes containing chemically stable structural mimics presents the possibility for a method to elucidate the enzymes involved in repair of an ICL that links two 2'-deoxyguanosine residues between the strands. The orientation of the O^6 of 2'-deoxyguanosine in B-form DNA is such that the lone pair of electrons on the oxygen protrudes prominently into the major groove providing free access to small electrophilic alkylators such as HM (Figure 1.1).

Additionally it is known that stable O^6-2'-deoxyguanosine lesions are produced by a number of electrophilic alkylators and can lead to mispairing with thymine giving rise to a transcriptional mutation.[82] A structural mimic can provide a conduit for the investigation of

the role of O^6-alkylation for mismatch and direct repair studies. The design of an O^6-alkyl-O^6 analogue for BU and HM will allow for the construction of DNA segments of defined structure and sequence (Scheme 2.1). This in turn allows for future studies to confirm or disprove O^6-alkyl addition as a mechanism of action or the role of direct repair in removal of O^6 lesions. Further potential exists for our O^6 ICL to act as a structural mimic for addition to the N^7 position with BU and HM (Scheme 2.3).

Scheme 2.1: Retrosynthetic pathway from (**I**) the dimer integrated into the DNA duplex (**II**) the dimer that is amenable to solid phase synthesis and finally (**III**) a bifunctional linear alkyl chain and 2'-deoxyguanosine building blocks used to synthesize the dimer.
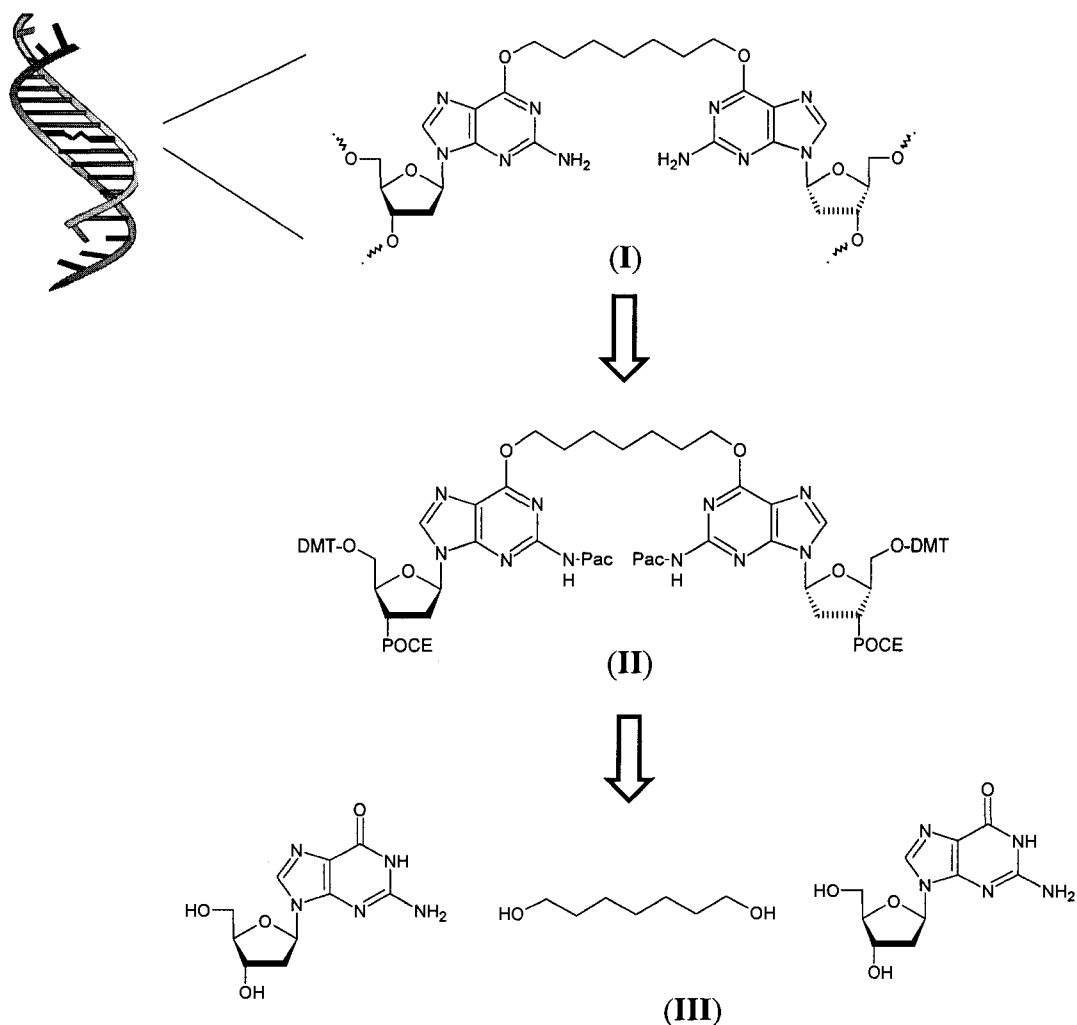
Multiple methods exist for the alkylation of the O^6 position of 2'-deoxyguanosine many of which require the replacement of the oxygen with a better leaving group, typically a halogen atom or a quaternary nitrogen (Scheme 2.2).[109] The most common of these reactions involves nucleophilic substitution with alkoxides having chlorine functionalization of the O^6 position allowing for incorporation of a number of heteroaromatic alcohols.[109] Reaction conditions require reflux conditions in the presence of POCl$_3$ followed by a reaction with an alkoxide.[109] The chlorinated methodology has the disadvantage of being unfeasible for syntheses in which the alcohol to be coupled is either expensive or difficult to prepare as the alcohol serves as both reactant and solvent.[109] Alternatively use of 2-aminopurin-6-yl-trimethylamonium chloride circumvents the need for the alkoxides as the solvent and provides a better leaving group. The disadvantage is that another synthetic step is required under harshly basic conditions using sodium hydride.[109] Both approaches are only compatible with the purine ring system requiring subsequent addition of the 2'-deoxyribose.[109] Syntheses that manage to overcome poor anomeric addition selectivity frequently require a sterically bulky adduct at the O^6-position.[113] This would preclude the ability for placement of an O^6-linkage prior to placement of the sugar moiety. Many O^6-alkylation routes employ conditions that would quickly remove protecting groups commonly present in solid-phase oligonucleotide synthesis. Use of the Mitsunobu reaction was explored due to its exceptionally mild conditions, compatibility with all protecting groups employed in solid-phase oligonucleotide synthesis, minimal side reactions, and direct route to the desired alkyl linker for the synthesis of **II** in Scheme 2.1.[115-117]

Scheme 2.2 Common synthetic methods for O^6-alkylation by installing a better leaving group.[109]
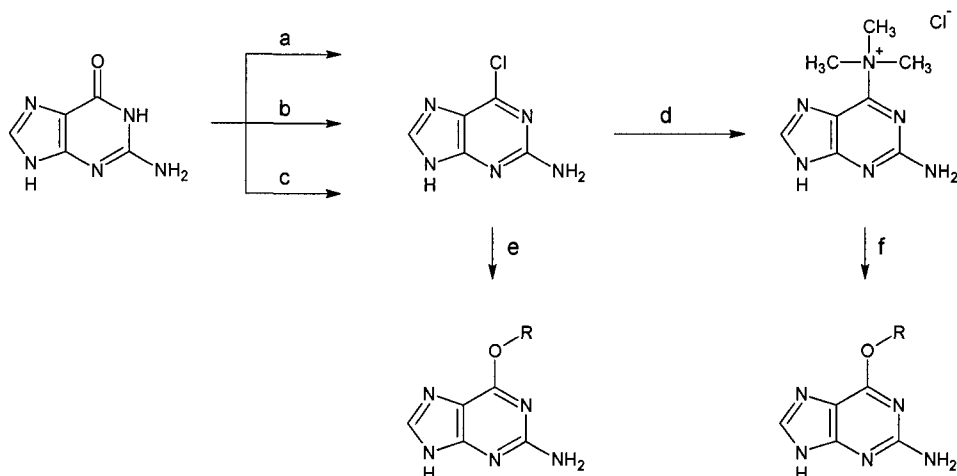
Conditions: a) POCl₃, *N,N*-diethylaniline (Sato); b) POCl₃, DMF (Igi and Hayashi); c) POCl₃, ACN, Et₃NMe⁺Cl⁻ (Hanson); d) NMe₃, r.t.; e) DMAP, NaH, R-OH, 60°C, DMSO; f) NaH, R-OH

Scheme 2.3. Structure of BU (n=2) and HM (n=5) with suggested clinically relevant N^7 alkylation, decomposition and O^6 alkylation products.
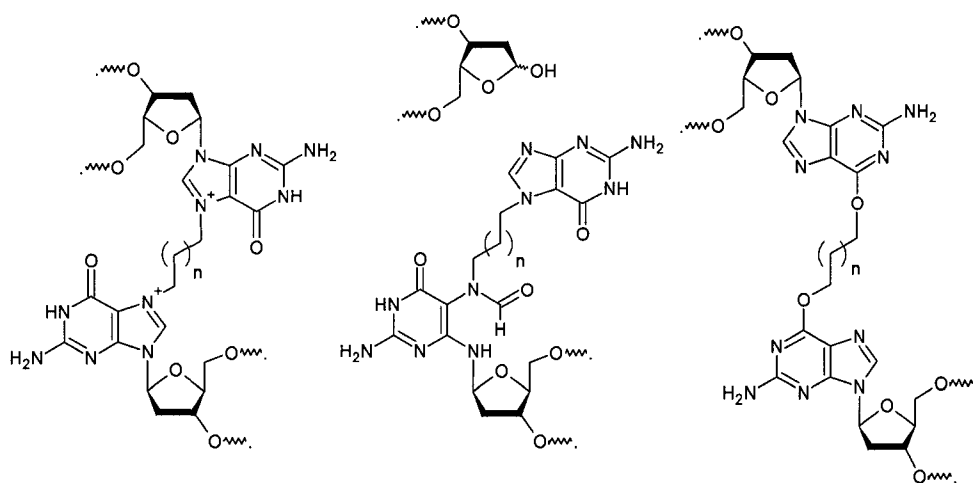
This thesis will deal specifically with studies on the synthesis of 2'-deoxyguanosine-O^6-alkyl-O^6-2'-deoxyguanosine ICLs and their incorporation into one helical turn DNA duplexes of defined sequence and cross-link orientation. Use of traditional solution phase synthesis and solid phase chemistries that allows for ordered removal of various hydroxyl

42

protecting groups providing units of asymmetry in the final DNA duplexes will be discussed (Scheme 2.2). The monopartate strategy has one unit of asymmetry enabling the synthesis of directly opposed 1,1 mismatched duplexes in addition to 1,2 ICL duplexes which can be later incorporated into a DNA plasmid for general repair studies. Introduction of an additional unit of asymmetry in the bipartate scheme allows synthesized duplexes to adopt the clinically relevant staggered 1,3-orientation also allowing incorporation of 'sticky ends' for future plasmid incorporation studies to examine repair pathways. Furthermore, the tripartate approach allows for complete strand asymmetry within a duplex but likewise requires complete orthogonal protection at all of the alcohols on the 2'-deoxyribose moieties of the dimer with protecting groups amenable to solid phase synthesis. Duplexes synthesized in this manner can add nucleoside phosphoramidites in a sequential manner to synthesize four unique DNA chains around the ICL dimer. These duplexes provide DNA sequences of defined structure and sequence that will allow more systematic investigation into repair mechanisms that reduce the efficacy of bifunctional alkylating antineoplastic agents.

Scheme 2.4 Progressively increasing control over the sequence composition of each DNA duplex is a function of increasing asymmetry. The ability for selective deprotection of individual alcohol functionalities imparts the ability to grow nucleotide chains individually. Nucleotide chains circled are symmetrical to another nucleotide chains in the duplex centred on the ICL.
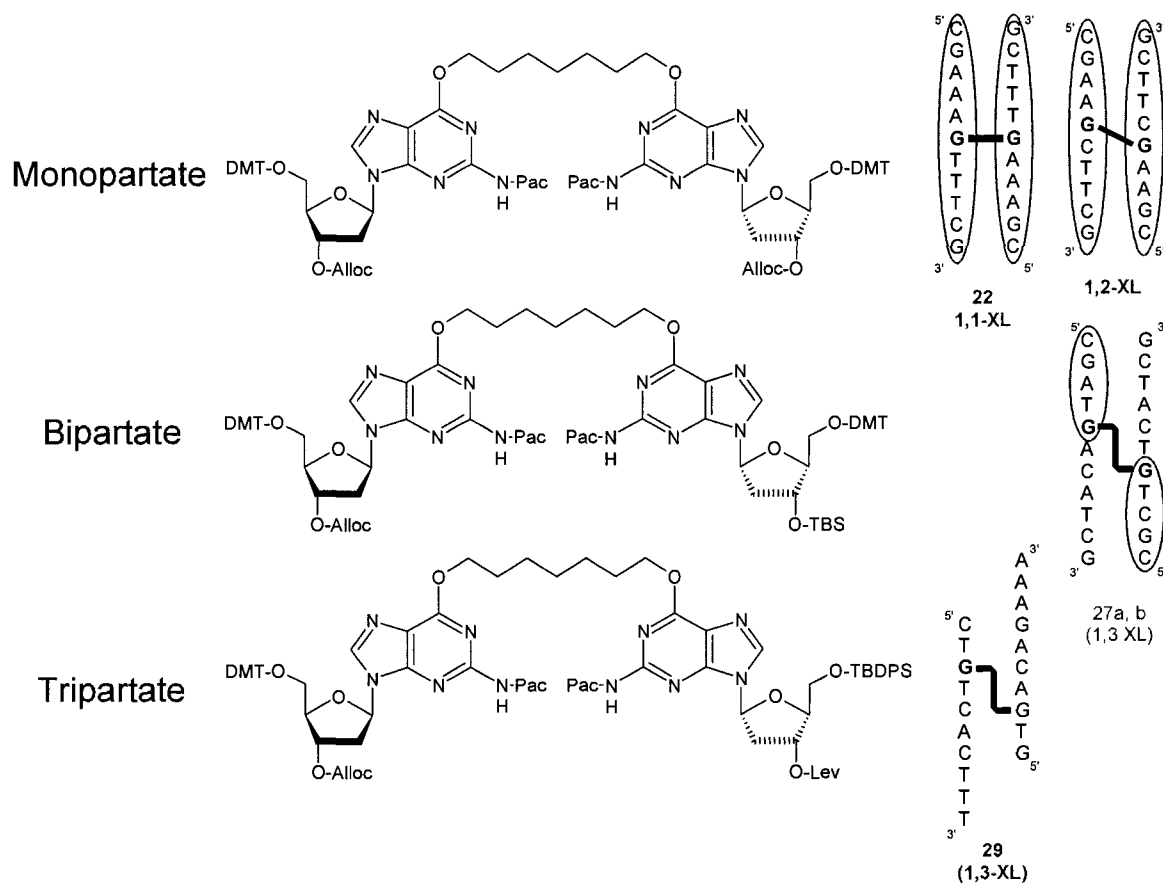
Chapter 3 : Results and Discussion

3.1. Synthesis and Properties of the Dimers.

Synthesis of the mono-, bi-, and tripartate dimers began with the introduction of the phenoxyacetyl group at the N^2-position. It should be noted that the most common group for N2 position of 2'-deoxyguanosine is the isobutyryl group. Initial attempts to synthesize our desired duplex using the traditional isobutyryl protection were successful for the O^6-cross-linked dimer. Unfortunately at the final deprotection stage of the duplex the two isobutyryl groups were retained as shown by MALDI-TOF analysis of the oligomers.[118] In an effort to use a more labile N^2 protecting group synthesis of 1 employed a modified Jones transient phenoxyacetyl (Pac) protection with a yield of 88% (Scheme 3.1).[119] The crude product consisted of a mixture of starting material and product. Reduced temperature (0°C) and timing (20 min) of NH_4OH/H_2O desilylation was critical to optimized yields. Removal of pyridine via high vacuum of the crude material greatly aided the successive liquid-liquid extractions (CH_2Cl_2/H_2O) which eliminated nearly all traces of starting material. Optimized yields of the more labile N^2-phenoxyacetyl group were comparable to that of the traditional isobutyryl method despite the increased sensitivity to base. Subsequent tritylation of 1 proceeded to form 2 in high yields (88%) after quenching of the reaction mixture with stoichiometric amounts of NH_4OH in methanol. Base treatment permitted neutralization of HCl and consequently removal of pyridine by rotary evaporation prior to aqueous work-up of the organic phase. Overnight removal of pyridine by high vacuum also greatly aided purification by flash chromatography.

Initially allyl 1–benzotriazolyl carbonate (Alloc-OBt) was commercially available however during the course of this work another route for allyl carbonate protection had to

be explored. Attempts to use allyl-chloroformate in the presence of triethylamine (TEA) for 3'-allylation produced exclusively decomposition products. Alloc-OBt was synthesized via a condensation reaction with allyl-chloroformate and 1–benzotriazole producing highly favourable results (85%).[122]   During Alloc-OBt recrystallization, removal of excess TEA and other side products proved unsatisfactory in our hands based on procedures described from literature methods.[120] Yields were low (>25-40%) and NMR revealed only a partial enrichment in purity of the crystals obtained versus the mother liquor with the resulting product in lower than expected yields.   A simple chromatography using CH_2Cl_2 on dehydrated silica gel was sufficient to produce Alloc-OBt in greater than 95% purity in comparable yields to that of the previously reported recrystallization method.[120,121] Subsequent reactions with chromatographed Alloc-OBt proceeded efficiently producing **3** in higher yields (>85%) in contrast to (<65%) the recrystallized material.

Scheme 3.1 Synthesis of **5a** (butyl, n =2) and **5b** (heptyl, n = 5) monoalkylated 2'-deoxyguanosine

/ **1** (R = H, R'=H)

ii { **2** (R = DMT, R' = H), 88%

iii \ **3** (R = DMT, R' = Alloc), 85%

v / **4a, 4b** (R = TBDPS), 78.9%, 61.9%

\ **5a, 5b** (R = H), 70.3%, 89.6%
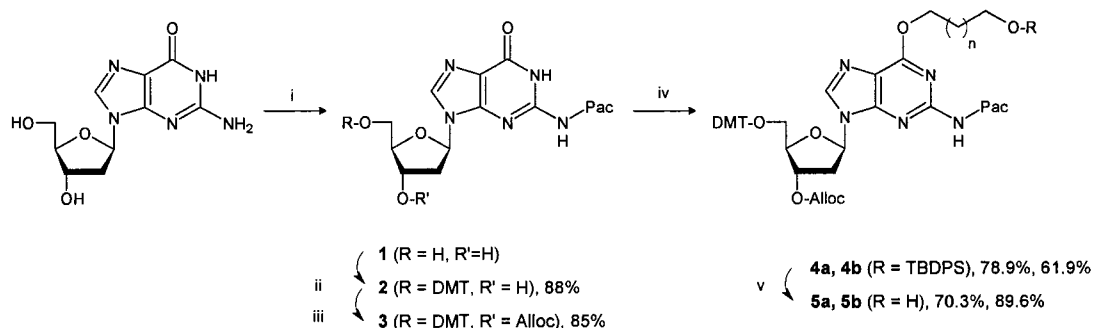
Conditions: (i) 1) TMS-Cl, pyridine, 0°C, 30 min, 2) Pac-Cl, Pyridine, 3-4 hours, 3) NH_4OH/H_2O (1:1), 0°C, 20 min, (ii) DMT-Cl, Pyridine, 12 hours, (iii) Alloc-OBt, CH_2Cl_2/pyridine (9:1), 12 hours, (iv) 4-(*tert*-butyldiphenylsiloxy) butanol or 7-(*tert*-butyldiphenylsiloxy) heptanol (1.1 eqv.), triphenylphosphine (1.25 eqv.), DIAD (1.2 eqv.) (v) TBAF (1M in THF).

3.1.1 Monopartate dimer synthesis

It is difficult to imagine a more efficient method of dimer formation than by a direct alkylation using two equivalents of the fully protected monomer and one equivalent of the

bifunctional linear alkyl chain. Alkylation using Mitsunobu methodology at the O^6-position was first reported by Pfleiderer.[122,123] Mitsunobu coupling using 1,4-butanediol and two equivalents of **3** was attempted yielding only **3** and spent Mitsunobu reagents. It is thought that the 1,4-butanediol underwent an internal cyclization to produce tetrahydrofuran due the same tendency of BU under biological conditions.[74] A mono-protected diol strategy was then employed for the synthesis of **4a** and **4b** which was optimal with slight (1.05-1.1 eqv.) excess of mono-silylated diol producing near quantitative addition. All reagents were added to the reaction mixture followed by drop-wise addition of diisopropylazidodicarboxylate (DIAD). Excess triphenylphosphine (PPh₃) (1.1 eqv.) and DIAD (1.05 eqv.) ensured that all of the monosilylated alcohol remained in the acyloxyphosphonium ion form minimizing any homocoupling. A slight excess of PPh₃ over DIAD was used to limit DIAD anion additions to the O^6-position.[124] Work-up of the organic phase followed by flash chromatography with a short column on the crude mixture (sample:silica, 1:3, *w/w*) loaded in dichloromethane (DCM) and eluted with hexanes results in nearly all of the product and non-polar reaction components being eluted in the first fraction leaving the majority of triphenylphosphine oxide on the column as confirmed by thin layer chromatography (TLC). It was also observed that solvation in Et₂O after purification resulted in selective precipitation of residual triphenylphosphine oxide contaminant for **4b** (61.9%) and to a lesser extent **4a** (78.9%), which was not successful on the crude mixture. Facile removal of the TBDPS group was performed using TBAF in THF to give **5a** (70.3%) and **5b** (89.6%) with flash chromatography allowing for the complete removal of any residual triphenylphosphine oxide contaminant.

Syntheses of monopartate dimers **6a** (53.1%) and **6b** (66.3%) (Scheme 3.2) were

accomplished via a second Mitsunobu reaction of the **5a** or **5b** with a slight excess of

3.[122,123] Separations using various solvent compositions were attempted with

ethylacetate/hexanes providing reasonable separation from all side products. The two alloc

protecting groups of **6a** and **6b** were then removed using 5 mol% of

tetrakis(triphenylphosphine) palladium (0) (Pd(PPh₃)₄) and 4 equivalents of a stock solution

of n-butylamine and formic acid (1:1) to produce **7a** (84.5%) and **7b** (77.1%) after work-up

and flash chromatography.[120,121] Sample **7a** was stored at -20°C for future use.

Scheme 3.2. Synthetic strategy for compound **7a** (n=1) and **8b** (n=4).
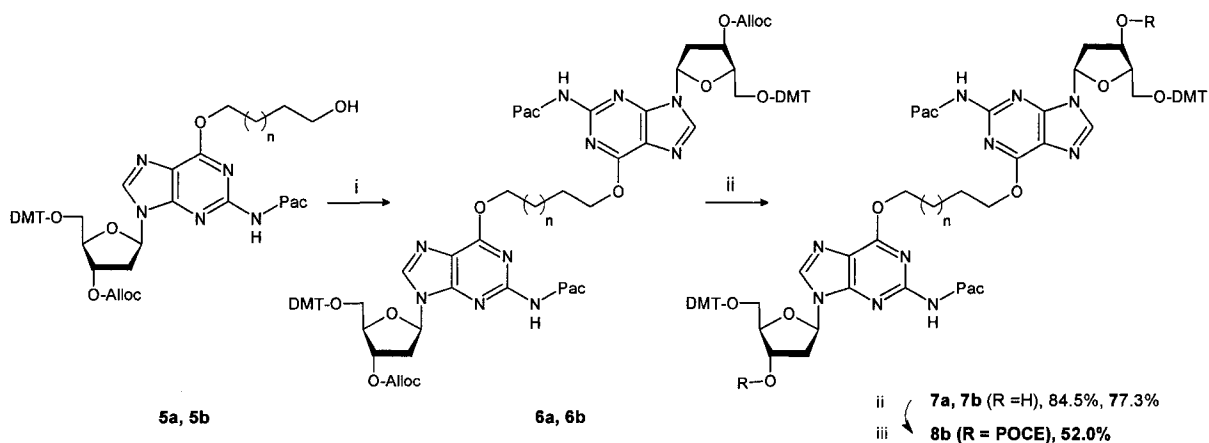
Conditions: (i) **3** (1.01 eqv.), triphenylphosphine (1.25 eqv.), DIAD (1.2 eqv.), 1 hour (ii)
Pd(PPh₃)₄ (5 mol%), PPh₃ (0.25 eqv.), *n*-butylamine/formic acid (1:1, 4.0 eqv.), THF, 20
min., (iii) diisopropylammonium tetrazolide (3.4 eqv.), *β*-2-cyanoethyl-*N*,*N*,*N'*,*N'*-
tetraisopropylphosphane (3.5 eqv), THF, 30 min.

NMR analysis of **7b** prior to reaction was useful for determination of the moisture

content of the sample as indicated by a singlet at 2.84 ppm in acetone-d_6. Once rigorously

dried **7b** was solvated in anhydrous DCM followed by the addition of

diisopropylammonium tetrazolide (3.4 eqv) then the atmosphere was exchanged with argon.

Using a 1 mL disposable syringe, β-2-cyanoethyl-*N*,*N*,*N'*,*N'*-tetraisopropylphosphane agent

(3.5 eqv.) was measured by volume and added over 30 seconds to a vigorously stirred

reaction mixture.[131] Reaction progress was monitored by the visual formation of a white precipitate and the appearance of three diastereomeric phosphoramidite spots on TLC signalling complete consumption of the starting material. The reaction was quenched using 5% NaHCO$_3$ (aq) followed by a brief work-up with the organic phase dried over Na$_2$SO$_4$ (MgSO$_4$ is slightly acidic) followed by immediate solvent rotary evaporation. The sample was purified via a hexanes precipitation from a minimum amount of DCM to yield an off-white solid **8b** (58.6%) which was sealed and stored at -20°C for use in solid phase synthesis. It should be noted that **8b** is extremely acid and heat sensitive. Typically compounds were stored as the free alcohols and the final solution phase phosphitylation reaction performed as close as possible to the date of solid phase synthesis.

3.1.2 Bipartate dimer synthesis

Progress to the bipartate synthesis of a more asymmetric dimer allows synthesis of the clinically relevant lesion but requires a labile 3'-protecting group that does not allow for removal of both the allylcarbonate and 5'-*O*-dimethoxytrityl (DMT) groups during deprotection. The *t*-butyldimethylsilyl (TBS) protecting group is well established in nucleoside chemistry. Silylation of **2** with imidazole (6.0 eqv.) followed by TBS-Cl (3.0 eqv.) in dimethylformamide (DMF) overnight afforded **9** (83.2%) after aqueous work-up and flash chromatography (Scheme 3.3). Compound **9** displays a lower polarity than **3** eluting at approximately 20% lower ethylacetate (EtOAc) composition.

Scheme 3.3 Synthesis of monomer **9** for the bipartate strategy.
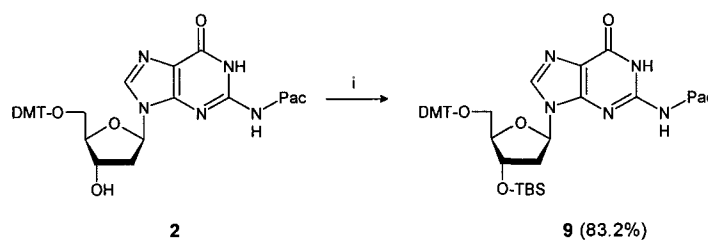
Conditions: (i) TBS-Cl, imidazole, DMF, 16 hours.

With both fully protected 2'-deoxyguanosine monomers in hand a second Mitsunobu reaction was performed to yield **10a** and **10b** (Scheme 3.4). Stoichiometric quantities of each monomer were critical to achieving maximal yields (**10a**, 62.9%; **10b**, 71.8%) and the most straightforward purification. Subsequent treatment of **10a** and **10b** with tetrabutylammonium fluoride (TBAF, 1M) in tetrahydrofuran (THF) followed by purification was performed to give the pre-monoamidite dimers **11a** and **11b** (68.5 and 71.1% respectively). Once again the samples were rigorously dried prior to phosphitylation. Using the more reactive N,N'-diisopropylamino cyanoethyl phosphoramidic chloride (1.5 eqv.) with DIPEA (2.0 eqv.) offered the advantage of using less equivalents of reactants with shorter reaction times.[131,132] The reaction was quenched when the starting material was consumed and two diastereoisomers were visible by TLC. The phosphoramidite work-up was brief with drying over Na_2SO_4 followed by immediate solvent rotary evaporation. The samples were purified by hexanes precipitation from a minimum amount of DCM to yield off-white solids **12a** and **12b** (88.3% and 97% yields) and the samples sealed and stored at -20°C. Later solid phase synthesis experiments with **12b** to generate test oligonucleotides produced lower yields after overnight deprotection of the allyl carbonate protecting group. This prompted removal of the allyl carbonate from **10b** via previously higher yielding 'solution phase' chemistry leaving the TBS group to generate **13** (76.0% yield) after aqueous work-up and flash chromatography (Scheme 3.5). Compound **13** was then phosphitylated using N,N'-diisopropylamino cyanoethyl phosphoramidic chloride (1.5 eqv.) with DIPEA (2.0 eqv.) yielding **14** (86.3% yield) from a hexanes precipitation.

Scheme 3.4. Strategy for the bipartate synthesis of **12a** (n=1) and **12b** (n=4).

5a, 5b   10a, 10b, 63.1%, 71.1%

ii ⎧ 11a, 11b (R =H), 75.0%, 86.6%
iii ⎩ 12a, 12b (R = POCE), 88.3%, 97%
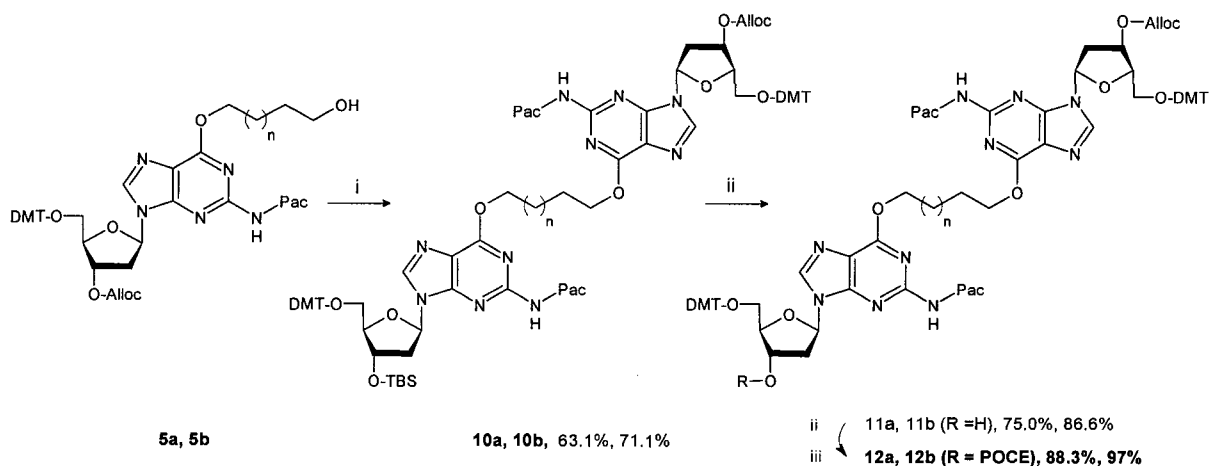
Conditions: (i) **5a/5b** (1.01 eqv.), PPh₃ (1.25 eqv.), DIAD (1.2 eqv.), 1 hour (ii) TBAF, THF, 30 min., (iii) DIPEA (2.0 eqv.), *N,N'*-diisopropylamino cyanoethyl phosphoramidic chloride (1.5 eqv), THF, 30 min.

Scheme 3.5. Synthetic strategy for the bipartate dimer **14** (n = 4).

10b   ii ⎧ 13 (R = H), 76.0%
             ⎩ 14 (R = POCE), 86.3%
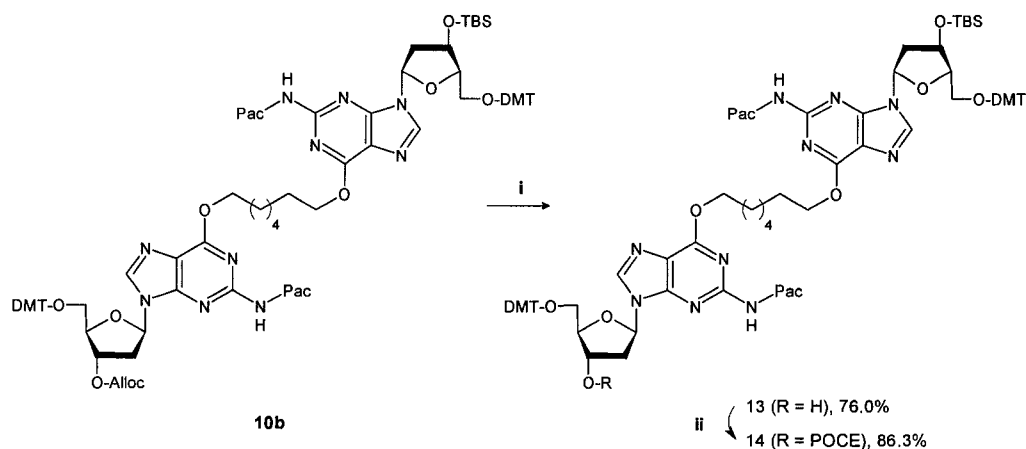
Conditions: (i) **10b**, Pd(PPh₃)₄ (5mol%), PPh₃ (0.25 eqv.), *n*-butylamine/formic acid (1:1, 2.5 eqv.), THF, 20 min., (ii) DIPEA (1.7 eqv.), *N,N'*-diisopropylamino cyanoethyl phosphoramidic chloride (1.5 eqv), THF, 30 min.

3.1.3 Tripartate dimer synthesis

To develop a completely asymmetric DNA duplex the use of orthogonal protection is required at all 2'-deoxyribose alcohol positions. Building on previous chemistry monomer **5** was used and two additional protecting groups were chosen. The requirement for a sterically demanding group capable of direction to the primary 5'-alcohol led to use of *t-*

butyldiphenylsilyl chloride (TBDPS-Cl) in dilute reaction conditions to further promote regioselectivity. Compound **1** was treated with imidazole (4.0 eqv.) in DMF followed by addition of TBDPS-Cl (1.1 eqv.) in five increments over 30 minutes to generate **15** (72% yield) after aqueous work-up and flash chromatography (Scheme 3.6). The observed polarity of the TBDPS group was greatly diminished, as observed by flash chromatography, over the DMT group resulting in elution with less polar solvents. Additionally, it was observed that TBDPS-Cl showed a similar reactivity to DMT-Cl but a slightly lower selectivity towards primary versus secondary alcohols. Longer reaction times which allowed the reaction to progress to completion had to be balanced with an increase in the amount of depurinated product observed.

Scheme 3.6. Synthesis of tripartate monomer **16**.

HO—[structure]—OH **1** →(i)→ TBDPS-O—[structure]—OH **15, 72.0%** →(ii)→ TBDPS-O—[structure]—O-Lev **16, 73.8%**
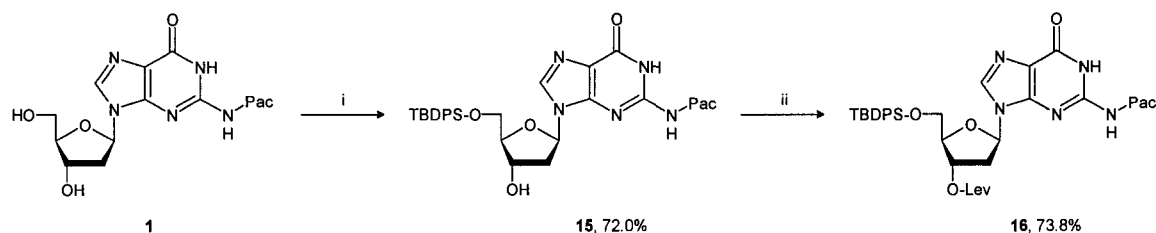
Conditions: (i) **1** (1.0 eqv.), TBDPS-Cl (1.2 eqv.), imidazole (4.0 eqv.), DMF, 16 hours (ii) levulinic acid (6.0 eqv.), DCC (3.5 eqv.), dioxane, 3 hours.

Small-scale 'litmus tests' were performed with hydrazine, employed in the deprotection of the levulinyl group showing no decomposition of **11b**. This suggested no decomposition would be observed with dimers tripartate protecting groups. Synthesis of the fully protected monomer **16** was performed using **15** which was solvated in dioxane followed by addition of a dicyclohexylcarbodimide (DCC, 3.5 eqv.) and a catalytic amount of dimethylaminopyridine (DMAP, 0.1 eqv.).[125] Levulinic acid (6.0 eqv.) was added to the vigorously stirred reaction. The reaction mixture changed to dark-orange brown in colour

with a white precipitate suspended in solution which was filtered off and rinsed with a minimal amount of dioxane. Sample work-up was then performed on the organic phase followed by flash chromatography. The product **16** displayed very little difference in polarity from that of **15** however purification was nonetheless straightforward as very little starting material remained.

Syntheses of the tripartate dimer **17** (Scheme 3.7) was accomplished via a second Mitsunobu reaction of **5** with a slight excess of **16**. Separations using various solvent compositions were attempted with hexanes/ethylacetate providing reasonable separation from all side products. The alloc protecting group of **17** was then removed using 5 mol% of Pd(PPh$_3$)$_4$ and a stock solution of n-butylamine and formic acid (1:1) (4 eqv.) to produce **18** (84.5%) after aqueous work-up and flash chromatography.[120,121] The bulk amount of sample **18** was stored at -20°C for future use. A portion of **18** was rigorously dried and treated with N,N'-diisopropylamino cyanoethyl phosphoramidic chloride (3.0 eqv.) with DIPEA (3.4 eqv.) in THF. The reaction was complete when starting material was consumed and two diastereoisomers were visible by TLC. The aqueous phosphoramidite work-up was brief and dried over Na$_2$SO$_4$ followed by immediate solvent rotary evaporation. The decision was then made to perform flash chromatography in light of the number of equivalents of reagents that were required to bring the reaction to completion and the orange colour of the reaction mixture. The sample was solvated in dichloromethane/hexanes (1:1) and loaded on top of the column. All solvents were doped with five drops of TEA for every 100 mL of eluent. In the first fraction a white precipitate was noticed to emerge, as such, the collection vial was immediately switched to fraction two. TLC revealed the solid in fraction two to be **19** and fraction one contained excess

phosphitylating agent. After elution of precipitated **19** normal chromatography followed to

yield an off-white solid (58.3%) and the samples sealed and stored at -20°C.

Scheme 3.7. Strategy for the synthesis of **19**.
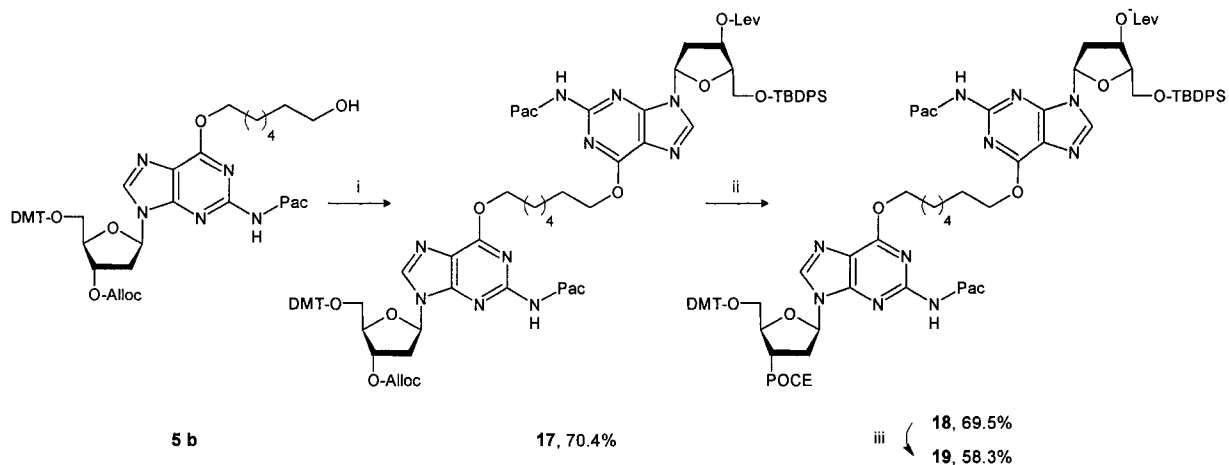
Conditions: (i) **5b** (1.01 eqv.), **16** (1.0 eqv.), triphenylphosphine (1.25 eqv.), DIAD (1.2 eqv.), 1 hour (ii) Pd(PPh₃)₄ (5mol%), PPh₃ (0.25 eqv.), *n*-butylamine/formic acid (1:1, 4.0 eqv.), THF, 20 min., (iii) DIPEA (1.7 eqv.), *N,N'*-diisopropylamino cyanoethyl phosphoramidic chloride (1.5 eqv), THF, 30 min.

3.2 Synthesis, Characterization, and Properties of Oligonucleotides.

Solid-phase synthesis was performed on an Applied Biosystems 3400 DNA Synthesizer via standard methods using commercially available 3'- and 5'-O-2'-deoxyphosphoramidites on a 1-μmol scale for each duplex synthesized (Scheme 1.2).[92,107,126] The use of 'fast-deprotecting' phenoxyacetyl protection on all dimers made it necessary to use phenoxyacetic anhydride in place of acetic anhydride as the capping reagent to prevent an undesired N-acetylation displacement reaction at the N^2-position observed during solid-phase oligonucleotide synthesis.[127] Base sequence compositions were chosen to allow structural comparisons with previous studies on synthesized duplexes containing a variety of cross-links of defined structure and sequence.[129-132] Selection of two alternating dG·dC pairs at each end allows for higher stability in the duplex form favoured for NMR studies.[132]

Synthesizer parameters were programmed for coupling in the 3'→5' direction for each respective phosphoramidite (15-20 eqv.) and tetrazole (45 eqv.). Standard coupling times (120 s) were used for 2'-deoxyphosphoramidites and extended couplings for dimer incorporations (600 s) and the nucleotide proceeding after dimer couplings (180 s) to reduce the likelihood of diminished yield due to steric influences. Sequences were typed into the DNA synthesizer method in a 5'→3' direction. It was verified that the sequence selected on the computer software was correctly transferred to the DNA synthesizer. The number of times a phosphoramidite was to be incorporated into a DNA chain was used to calculate the quantity of phosphoramidite to be placed on each respective line. ACN was dried over CaH_2 and transferred via syringe to each 5 mL test tube which was then capped and lightly vortexed to dissolve each respective compound. After installation on the

synthesizer of all reagents bottles and sample lines were purged with solvent under argon atmosphere to ensure anhydrous conditions. A visual check of all solvents levels, argon gas level, and monitoring for leaks was then performed.

Linear sequences were coupled to each respective dimer were followed by a capping step, then visual monitored for the presence of the orange trityl cation eluting from the column after treatment with trichloroacetic acid (TCA). With the trityl protecting group removed, normal 3'→5' direction chain growth is resumed. For sequences requiring deprotection of the 3'-alcohol after the dimer has been coupled, the use of 5'-*O*-2'-deoxyphosphoramidites were required to continue chain growth for the final segment. After addition of the last nucleotide, oligomers were removed from the solid support and deprotected via treatment with 1:3 ethanol/ammonium hydroxide (28%) at 55°C for 4 hours. The crude oligonucleotides were then purified by strong anion exchange (SAX) HPLC (Table 3.1). Purified samples were desalted by absorbing each sample onto a SEPPAK (C18) column followed by washing with deionized water (30 mL) to remove salts and desorbed using ACN (3 mL x 2). The desalted samples were then evaporated using vacuum centrifugation to yield a dried solid.

Table 3.1. Standard SAX-HPLC method for separation of oligonucleotides using a Waters 1525 pump and 2487 detector with a Dionex DNAPac PA-100 (4x250 mm) column.

	Time	Flow	%A	%B	Curve
1	0.01	1.00	100.0	0.0	6
2	30.00	1.00	50.0	50.0	6
3	32.00	1.00	50.0	50.0	6
4	35.00	1.00	0.0	100.0	6
5	37.00	1.00	0.0	100.0	6
6	40.00	1.00	100.0	0.0	6
7	45.00	1.00	100.0	0.0	6
8	46.00	0.00	100.0	0.0	6

Solvent A: 1 M Tris-HCl (pH 7.8) / 10% ACN / H_2O (1:1:8)
Solvent B: 1 M Tris-HCl (pH 7.8) / 10% ACN / 1.25 M NaCl (1:1:8)

Coupling of 3'-phosphitylated dimers to a deprotected 5'-alcohol of a linear chain were performed in an 'on-column' and 'off-column' manner. The on-column method had the advantage of all dimer couplings to the linear chains being automated under anhydrous and inert argon atmosphere. The on-column method however suffers an unavoidable loss by sending the unused phosphoramidite dimer, used to prime sample delivery lines, to waste. Employing the off-column method involves removal of the column containing the solid support from the DNA synthesizer followed by transfer to a screw cap microfuge tube. The solid support is then suspended in an ACN solution containing phosphoramidite dimer and tetrazole. The off-column method offers the advantage of using only the required equivalents of dimer but suffers the disadvantage of manual manipulations exposing the reaction to non-anhydrous conditions via atmospheric moisture. After reacting off-column samples for an hour, the solid support is then transferred back into DNA synthesizer columns and rinsed using a vacuum apparatus with ACN (30mL) followed by THF (30mL). After removing all reagents, supports bearing oligomers were placed in columns then reinstalled on the DNA synthesizer and all sample lines purged with anhydrous solvent under argon then re-primed with reagents.

The most symmetrical oligonucleotide to be synthesized involved an on-column synthesis of a 1,1-mismatch one helical turn ICL duplex (Schemes 2.2 & 3.8). The incorporation of the bis-phosphoramidite **8b** (**8a** was not prepared) using the monopartate strategy contained the same type of functional and protecting groups as nucleotide monomers used in oligonucleotide synthesis making it highly amenable to on-column procedures. The quantity of each phosphoramidite nucleotide was calculated and placed on each respective line in a test-tube (Table 3.2). Allowance was made for 250 μL of priming

volume with an additional 50 µL for each commercially available nucleotide to ensure adequate volume in the base of the sample test-tube. Dimer coupling parameters were optimized with respect to concentration and coupling times. Analytical SAX HPLC runs revealed the optimal concentration range observed for coupling was at 0.10 M (coupling time 600 s) to balance insufficient amounts of dimer with over concentration resulting in mono-addition. A significant amount of **20**, the pentamer L-sequence product, was observed when 0.05 M concentrations of **8b** were employed (Figure 3.1). Concentrations of 0.15 M resulted in a diminished amount of pentamer **20** and two major products. The first is **21** resulting from mono-coupling of **8b** to the linear oligonucleotide chain to yield a Y-like 17-mer. The second is the fully cross-linked duplex **22** which was easily separated by SAX HPLC from **20** and **21**. The final yield of the cross-linked duplex **22** after purification and desalting was 12.6% (27.2 od units from a 1 µmol scale synthesis). PAGE studies were used as a confirmation of purity of the final purified duplex **22** (Figure 3.2).

Table 3.2. Calculated quantities required for solid-phase synthesis of ICL duplex **22** using 3'-O-phosphoramidites using on-column coupling of **8b**.

Phosphoramidite	Number of Incorporations	Mass (mg)	Concentration (mol•L^{-1})	Volume (µL)
dA	6	131.1	0.100	1500
dG	2	58.8	0.100	700
dC	4	93.5	0.100	1100
dT	6	111.7	0.100	1500
8b	1	66.6	0.050	700

Scheme 3.8 Synthesis (on-column) of 1,1-mismatch one helical turn ICL duplex **22**.
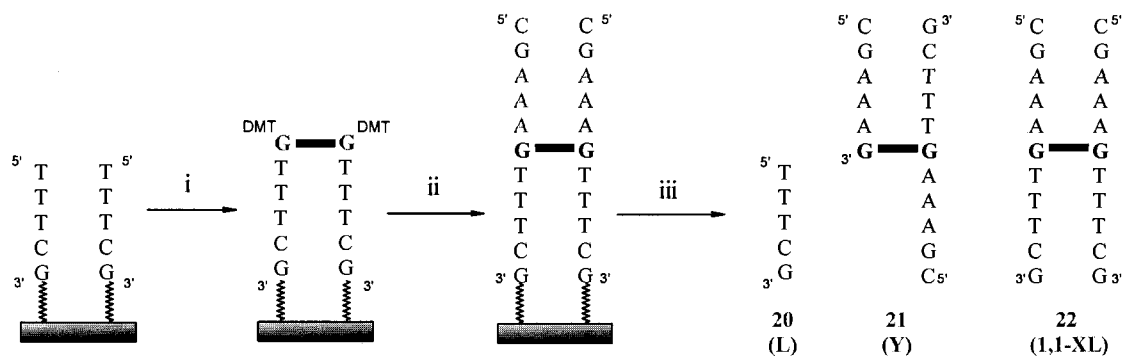
Conditions: (i) The addition of **8b**, (ii) continued chain growth with 3'-*O*-2'-deoxyphosphoramidites, (iii) final detritylation using TCA followed by removal from the DNA synthesizer with subsequent cleavage from the solid-support and deprotection using 1:3 ethanol/ammonium hydroxide (28%). Nucleoside loading on CPG was 65 μmol/g.
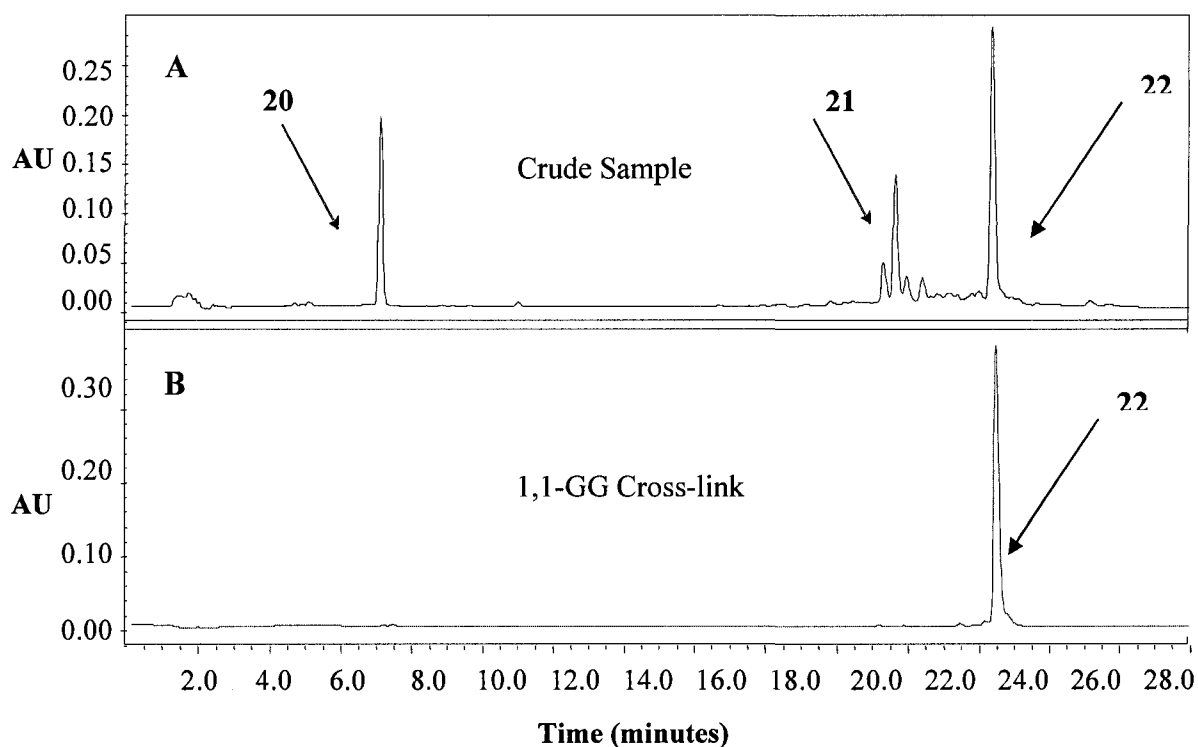
Figure 3.1 Analytical chromatograms using SAX HPLC of (A) the crude synthesis products of duplex **22** and (B) final purified product.

Figure 3.2 Analysis of purified oligonucleotide duplex **22** using denaturing PAGE (20 %) at room temperature.

Compositional analysis of the cross-linked duplex **22** was performed by enzymatic digestion with SVPDE and calf intestine phosphatase followed by analysis of the nucleosides via RP-HPLC (C18) as described previously[14] as well as MALDI-TOF MS. An HPLC chromatogram shows the first four peaks elute using a gradient with retention times corresponding to dC, dG, dT, and dA standards (Figure 3.3, Table 3.3). The final peak at 25.4 min coelutes with a completely deprotected standard of dimer **8b**. The ratio of the peak areas agrees with the nucleotide composition of the cross-linked oligonucleotide **22** (Table 3.4). MALDI-TOF analysis of **22** indicated a molecular weight of 6806.9 (expected 6805.4) (Apx A).
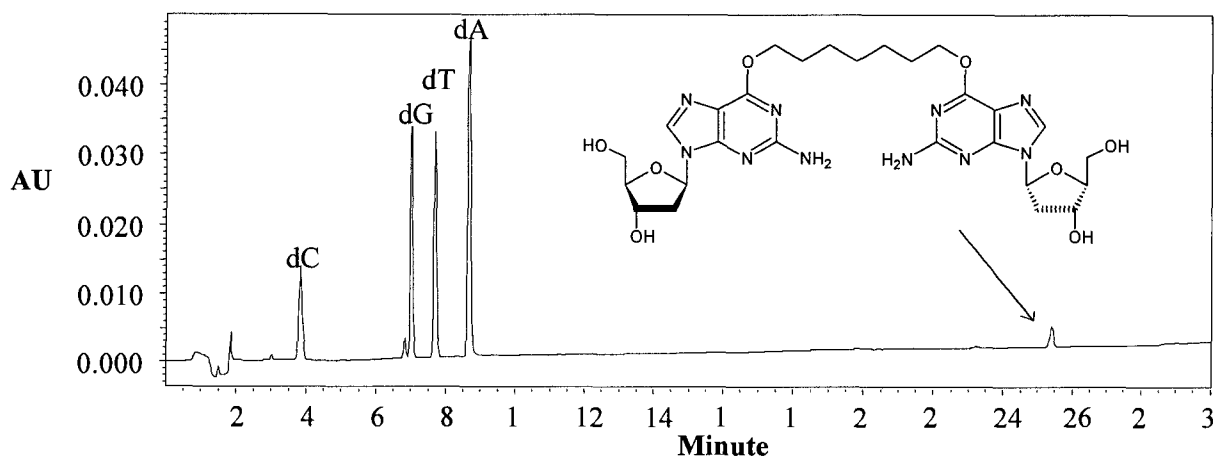
Figure 3.3: Analysis of a snake venom phosphodiesterase and calf intestinal phosphatase digest using a Waters 1525 pump and 2487 detector using a Symmetry C18 (4.6 x 150mm, 5μm) column.

60

Table 3.3. Standard HPLC method for separation of oligonucleotides using a Waters 1525 pump and 2487 detector with a Waters Symmetry C18 (4.6 x 150mm, 5μm) column.

	Time	Flow	%A	%B	Curve
1	0.00	1.00	100.0	0.0	6
2	20.00	1.00	40.0	60.0	6
3	22.00	1.00	40.0	60.0	6
4	24.00	1.00	0.0	100.0	6
5	26.00	1.00	0.0	100.0	6
6	28.00	1.00	100.0	0.0	6
7	30.00	1.00	100.0	0.0	6

Solvent A: 50mM Na_2HPO_4 (pH 5.8) / 2% ACN / H_2O
Solvent B: 50mM Na_2HPO_4 (pH 5.8) / 50% ACN / H_2O

Table 3.4 Nucleoside composition ratios corrected for molar extinction coefficients and normalized to one for the cross-linked duplex **22**. MALDI-TOF: observed 6806.9 (expected 6805.4)

Nucleoside	Peak area	Nucleoside ratios	
		Expected	Observed
dC	101140	1.00	1.00
dG	165006	1.00	1.03
dT	178764	1.50	1.44
dA	272597	1.50	1.39
dG-dG	23911	0.25	0.20

A variety of methods including ultraviolet transition T_m, CD, and gel electrophoretic mobility studies were employed to characterize synthesized duplexes. Deviations from the WC control duplex (**24**) dissociation temperatures are the result of primarily two effects giving an indication as to the effect of modifications on duplex stability. Entropic stabilization in **22** is observed as the two strands have less rotational degrees of freedom remaining unimolecular when fully denatured. Destabilization from the alteration to normal WC base pairing induced by the presence of an ICL or mismatched 2'-deoxyguanosine bases (**23**, 1,1-mismatched (5'-dCGAAAGTTTCG)₂) also serves to decrease dissociation temperatures. A sigmoidal denaturation curve qualitatively similar in shape to the **24** was exhibited for **22** (Figure 3.4). First derivative analysis of T_m plots

indicates dissociation temperatures of 35°C for duplex **23**, 46°C for duplex **24**, and 58°C for

duplex **22**. The presence of a mismatch in **23** resulted in a broadening of the sigmoidal

curve as the duplex dissociated over a larger temperature range as evidenced by a

diminished slope.
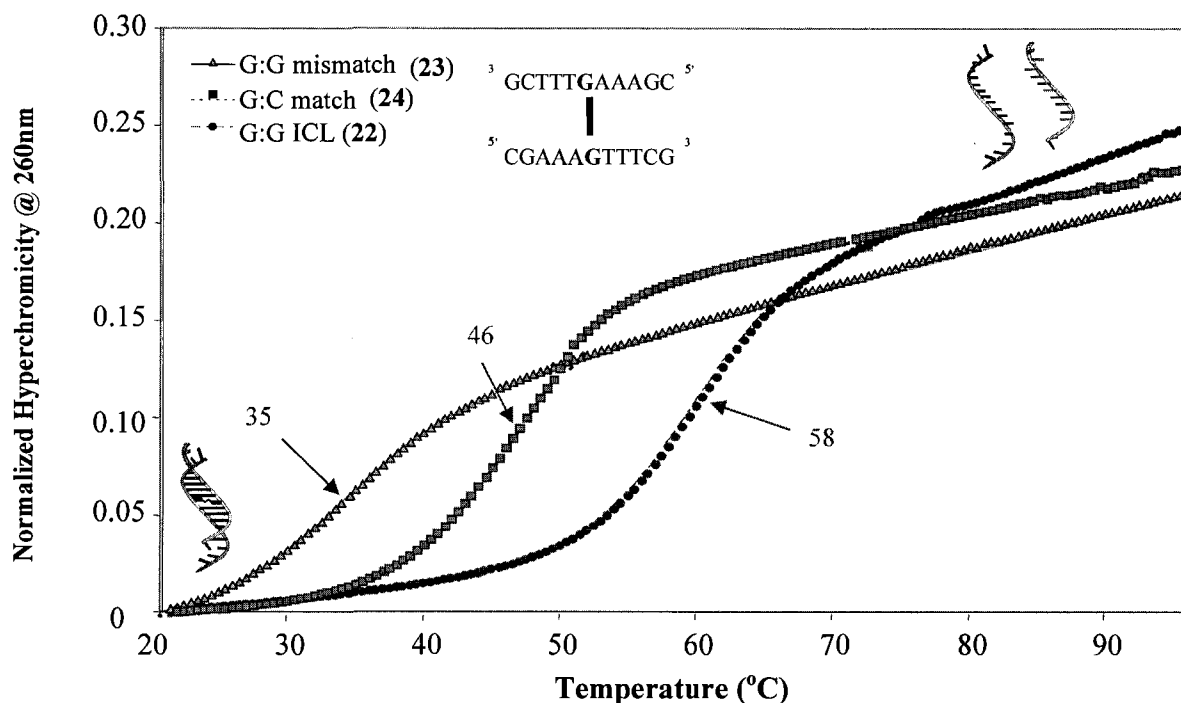
Figure 3.4. UV thermal denaturation profile (Δ0.5°C/min) of the cross-linked (**22**), non-cross-linked mismatch control duplex (**23**) (5'-dCGAAAGTTTCG)₂, and non-cross-linked control (**24**) duplexes.

The degree of hyperchromicity is similar at equivalent concentrations for **22** and **24**.

Duplex **23** possesses minimally diminished hyperchromicity possibly due to dG·dG

mismatches being well tolerated resulting in little disruption in base-stacking interactions.

The similarity in the magnitude of hyperchromicity for **22** and **24** suggests the disorder

induced by the dG·dG mismatching may be balanced by the order conferred by the ICL.

The forced proximity due to the unimolecular nature of **22** potentially maintains a degree of

dipole transfer between bases at the site of the ICL accounting for the similarity in

hyperchromicity and higher dissociation temperatures.

The CD spectra of **22, 23,** and **24** duplexes are displayed in Figure 3.5. These spectra were recorded at 5°C where only the annealed form exists as evidenced from T_m data. The control duplex **24** exhibits a spectrum consistent with the signature of natural B-form DNA having a maximum at 284 nm, a minimum at 253 nm, and a crossover at 270 nm. The CD spectrum of duplexes **22** and **23** display a high degree of similarity with the general shape of **24**, each having a relatively sharp positive band with maxima at 283 nm and 284 nm respectively. Crossover points for all three duplexes are identical at 270 nm. The molar ellipticity of **22** observed in the minima area centered on 255 nm has an identical shape however it displays 1.6 times greater molar ellipticity than that of **24**. This is contrasted in the maxima area for **22** centered on 283 nm being 0.7 times less than the **24**. This could result from subtle changes in stacking interactions of the bases at or near the site of the ICL.[129] At lower wavelength duplex **23** displays a deviation from the signature pattern exhibited by the **24** having slight shoulder at 250 nm and a shift of the minima to 259 nm. There is overall lower amplitude of the **23** sequence in all regions compared to the **24**. The observed relative difference of **23** from the **24** is 0.42 times in the maxima region and 0.57 times in the minima region. This may be a result of the slight distortion of the duplex expected from the presumed wobble dG·dG base pair. The overall signature similarity of both **23** and **22** to that of **24** suggest no major structural change is induced. Observed patterns also indicate a minimal amount of helical distortion induced from the 7 carbon linker, closely approximating native form DNA.
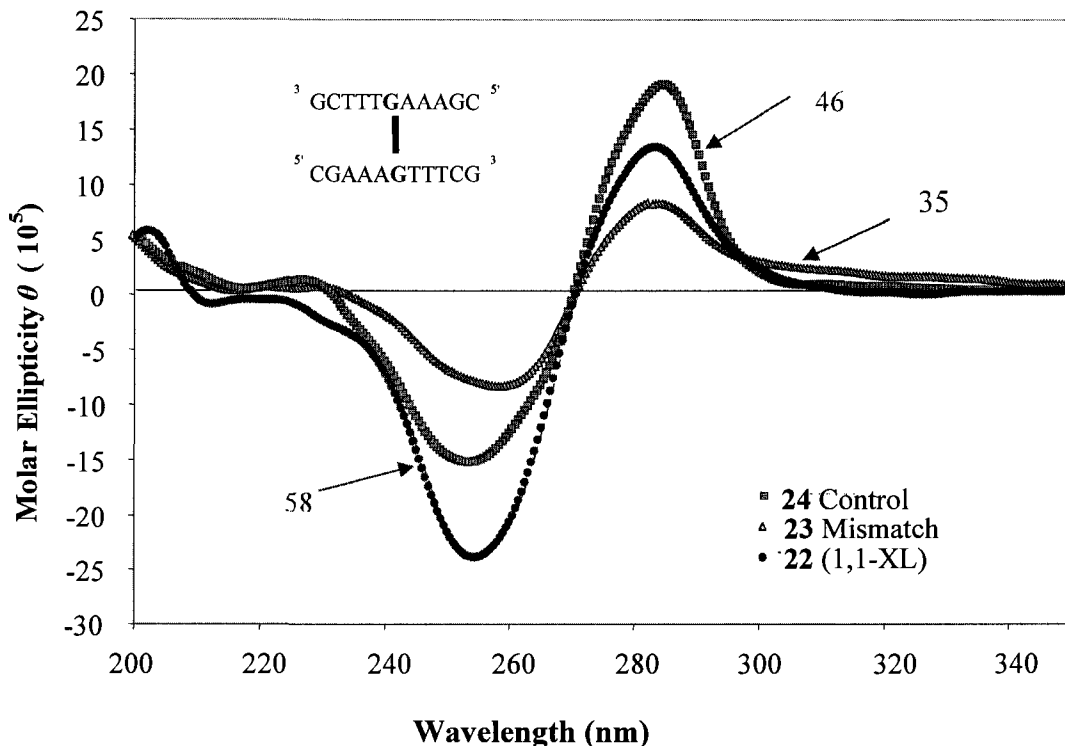
Figure 3.5. CD profile of the cross-linked oligonucleotide **22**, **23**, and **24** duplexes with dissociation temperatures labelled.

3.2.1. Bipartate synthesis

Progressing to the synthesis of the clinically relevant orientation of the ICL requires the design of a DNA duplex that is composed of a sequence with two units of asymmetry (Scheme 2.1, 3.9). Asymmetry is a prerequisite for the capability of forming a complimentary 5'-GXC-3' motif around the site of the ICL capable of WC hydrogen bonding, where X is any nucleotide. Coupling of a monophosphoramidite dimer containing a protected 3'-alcohol which is robust to the conditions of solid-phase synthesis is essential. Further constraints imposed are the need for facile removal with no degradation of the intermediate DNA duplex during deprotection. The bipartate strategy involves the formation of only one phosphorous linkage to a linear strand during coupling of the dimer. This offers the advantage over the monopartate approach of not requiring optimization of

64

the dimer concentration also allowing a higher concentration to further push the reaction to completion.

Scheme 3.9 Synthesis of interstrand cross-link duplexes **27a** (butyl) and **27b** (heptyl).
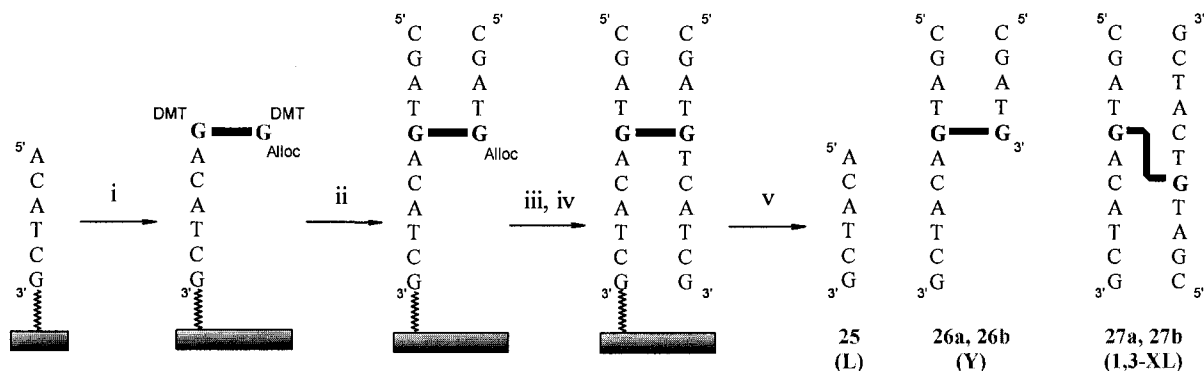
Conditions: (i) The addition of **12a** or **12b**, (ii) followed by 3'-O-2'-deoxyphosphoramidites, (iii) Pd(PPh$_3$)$_4$, PPh$_3$, butylamine/formic acid (1:1) (iv) continued chain growth with 5'-O-2'-deoxyphosphoramidites, (v) cleavage from the solid-support and deprotection using 1:3 ethanol/ammonium hydroxide (28%). Nucleoside loading on CPG was 65 μmol/g.

The bipartate synthesis is begun in the same manner as the monopartate with the generation of a linear sequence. This is followed by coupling of a 3-*O*-monophosphoramidite **12a** or **12b** in an off-column fashion to use only the required quantities of each dimer (Table 3.5). Dried dimers were added to separate screw-cap microfuge tubes with the L-sequence stationary phase. A tetrazole solution (0.45 M, 100 μL) was added to the microfuge tubes to dissolve the dimers upon which the samples were left to react for 1 hour. In addition to NMR analysis providing an indication of the samples' purity (**12a** and **12b**), the two solutions were clear in ACN at 0.200 M indicating minimal phosphoramidite side-products. After coupling the respective solid-supports were returned to DNA synthesis columns followed by a rinsing step using THF (30 mL) followed by ACN (30 mL). Samples were then placed under high vacuum for ~10 min to remove to residual solvents and then reinstalled on the DNA synthesizer to resume 3'→ 5' chain

extension to form the Y-sequence 17-mer **26a** and **26b**. A capping step is performed, aqueous iodine oxidation, then a detritylation step with TCA followed by continued chain growth using 3'-*O*-phosphoramidite nucleotides.

Table 3.5 Quantities of 3' and 5'-*O*-phosphoramidites required to generate **26a** and **26b** using off-column coupling of **12a** and **12b**.

Phosphoramidite	Number of Incorporations	Mass (mg)	Concentration (mol•L^{-1})	Volume (μL)
3'-*O*-dA	8	105.6	0.100	1900
3'-*O*-dG	4	66.2	0.100	1100
3'-*O*-dC	8	101.4	0.100	1900
3'-*O*-dT	6	66.4	0.100	1500
5'-*O*-dA	2	38.9	0.100	700
5'-*O*-dG	2	42.1	0.100	700
5'-*O*-dC	4	58.7	0.100	1100
5'-*O*-dT	4	48.7	0.100	1100
12a	1	36.2	0.200	100
12b	1	37.1	0.200	100

A variety of challenges were encountered during the syntheses of **26a** and **26b** with the failure to observe an orange coloured trityl cation when the reinstalled samples were treated with TCA. It was thought that failure to couple was the result of moisture or a contaminant in both syntheses using **12a** and **12b**. The presence of water can be expected to compete with the 5'OH of the support bound **25** as a nucleophile to displace the isopropyl amine of a phosphoramidite which is readily protonated in the presence of tetrazole. Water would be expected to have a greater rate of reaction than **25** resulting in greatly diminished yields of the desired duplexes. In an effort to dry the monophosphoramidites while also increasing sample purity, column chromatography was performed on **12a** and **12b**. All solvents contained TEA (1%) to prevent decomposition with the slightly acidic silanols present in silica gel. To further remove water from the system, all solvents and the final products were dried over sodium sulphate. Co-

evaporations with anhydrous THF were problematic as both **12a** and **12b** tend to bump excessively leading to loss of material and unwanted manipulations of the sample. The purified sample was sealed in a flask in solution over sodium sulphate for more than 24 hours (4°C) to remove any traces of water. NMR and TLC analysis showed an improvement in the purity however the presence of water (0.5 eqv.) at 2.86 ppm in acetone-d_6 (from an ampoule) was unavoidable with drying conditions used. A second set of solid-phase reactions were performed following the aforementioned procedure with the purified samples. The trityl cation was observed to elute from the DNA synthesizer columns upon reinstallation and treatment with TCA which was followed by continued 3'→5' extension of the free 5'OH groups.

Subsequent chain growth in the 5'→3' direction necessitates the removal of the alloc group[120,121] via an off-column method due to incompatibility of the DNA synthesizer to Pd(PPh$_3$)$_4$. After the Y-sequence synthesis was completed the solid-support was removed from the DNA synthesizer and placed into 4 mL vials with Teflon septa. A solution (1 mL) consisting Pd(PPh$_3$)$_4$ in the presence of a 1:1 butylamine/formic acid from a stock solution emulsified in THF was prepared (Table 3.6). Equal portions (500 μL) of the parent solution were added to each 4 mL vial. The vials were then sealed with parafilm, wrapped in tin foil, and placed on a shaker at the lowest setting (30 rpm) overnight. It was observed that the solution which was originally pale yellow turned to black overnight, staining the solid support even after successive washings. Using a Pasteur pipette, CPG portions (~10 beads) were taken for NH$_4$OH/EtOH deprotection at 55°C overnight. The oligonucleotides were then analyzed by SAX HPLC to confirm successful dimer couplings as indicated by a large ratio of Y-sequence to n-1 failure and pre-dimer linear sequences.

Table 3.6 Quantity of reagents used for the step (iii) off-column removal of the alloc protecting group.

Compound	FW	Mass (mg)	mmoles	Equivalents
Pd(PPh$_3$)$_4$	1155.58	20.7	0.0087	8
PPh$_3$	262.29	22.2	0.0846	84
Butylamine/Formic acid	119.17	15.9	0.133	133

Significant challenges were also encountered with the stability of the solid support bound Y-sequence cross-link. Initially it was thought that the deprotection conditions (NH$_4$OH/EtOH 3:1 overnight at 55°C) were responsible. Variation in time and temperature with additional solid-phase samples employed a 2 hour deprotection at room temperature and continued treatment for an additional 24 hours at 55°C of the previously analysed samples. No difference between samples deprotected for 2 hours or 48 hours were observed ruling out ammonia degradation.

To confirm the point of decomposition another 1 μmol scale solid-phase synthesis was conducted using **12a** and **12b**. The off-column coupling of **12a** and **12b** were successful as indicated by the trityl cation removal during TCA treatment. Analysis was performed before and after alloc deprotection demonstrating a dramatic change in sample composition with decomposition occurring during off-column Pd(PPh$_3$)$_4$ treatment (Figure 3.6). This contrast from yields in solution phase deprotection can likely be explained by the decreased time of exposure to the Pd(PPh$_3$)$_4$ catalyst in addition to a dramatic increase in the number of equivalents called for in the solid phase methodology.[120,121]
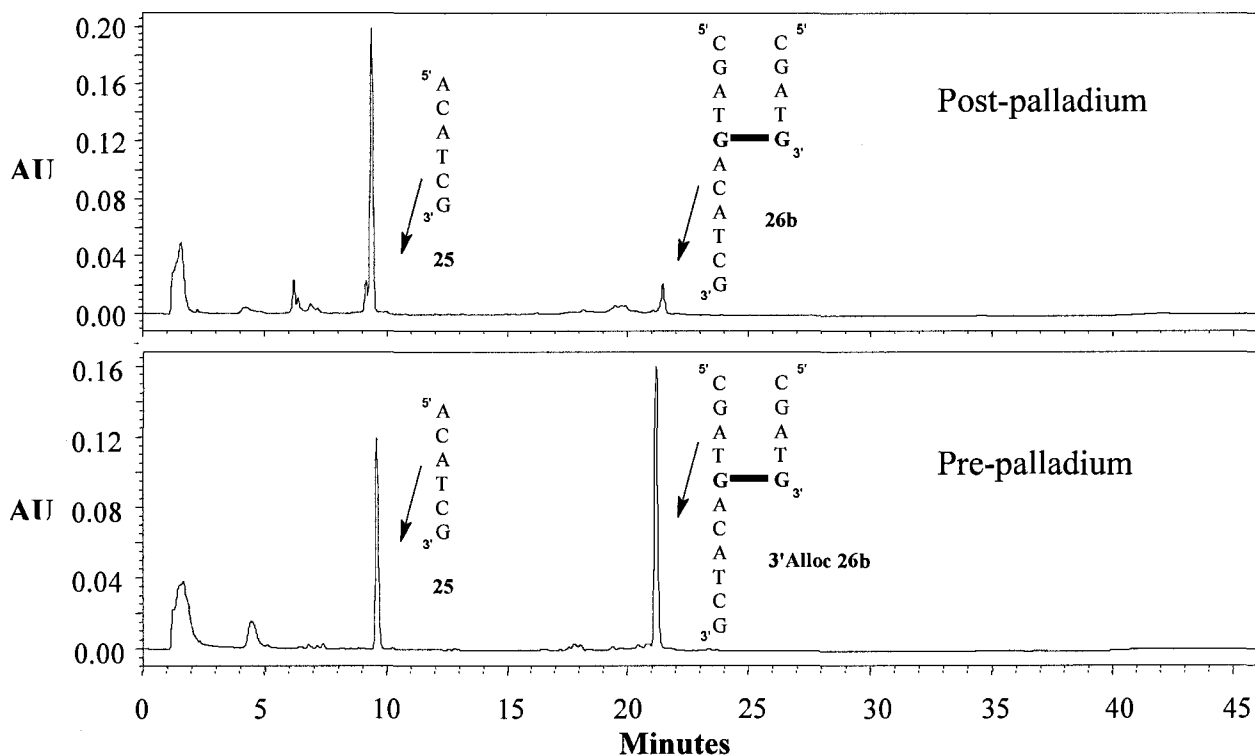
Figure 3.6. Pre and post Pd(PPh₃)₄ alloc deprotection group of Y-sequences **26** using SAX

In a desire to synthesize duplex **27b**, the blackened solid-support was reinstalled on the DNA synthesizer after washing. Deprotected Y-sequence extension with normal 3'-*O*-phosphoramidites demonstrated it was possible to generate the desired full-length product. The successful coupling of the more sterically congested 3'OH to a 3'-*O*-phosphoramidite also displayed the coupling times were sufficient for less strained 5'-*O*-phosphoramidite coupling at the site of the cross-link. Continued 5'→3' extension on the intact Y-sequence using 5'-*O*-phosphoramidite nucleotides was performed followed by a final detritylation step. After the completed synthesis the black stained appearance of the solid-support had been removed. The solid-support was then deprotected overnight with NH₄OH/EtOH (3:1) and analysed revealing the major component being **25**, a minor population of n-1 failure sequences and **27b** (Figure 3.7). Final analysis of the purified **27b** showed a highly pure

sample suitable for further structural studies. Purified **27b** was analyzed by SVPDE and calf intestine phosphatase digestion as described previously.[14] MALDI-TOF indicated a molecular weight of 6765.28 (expected 6768.29) (Figure 3.8, Table 3.7).
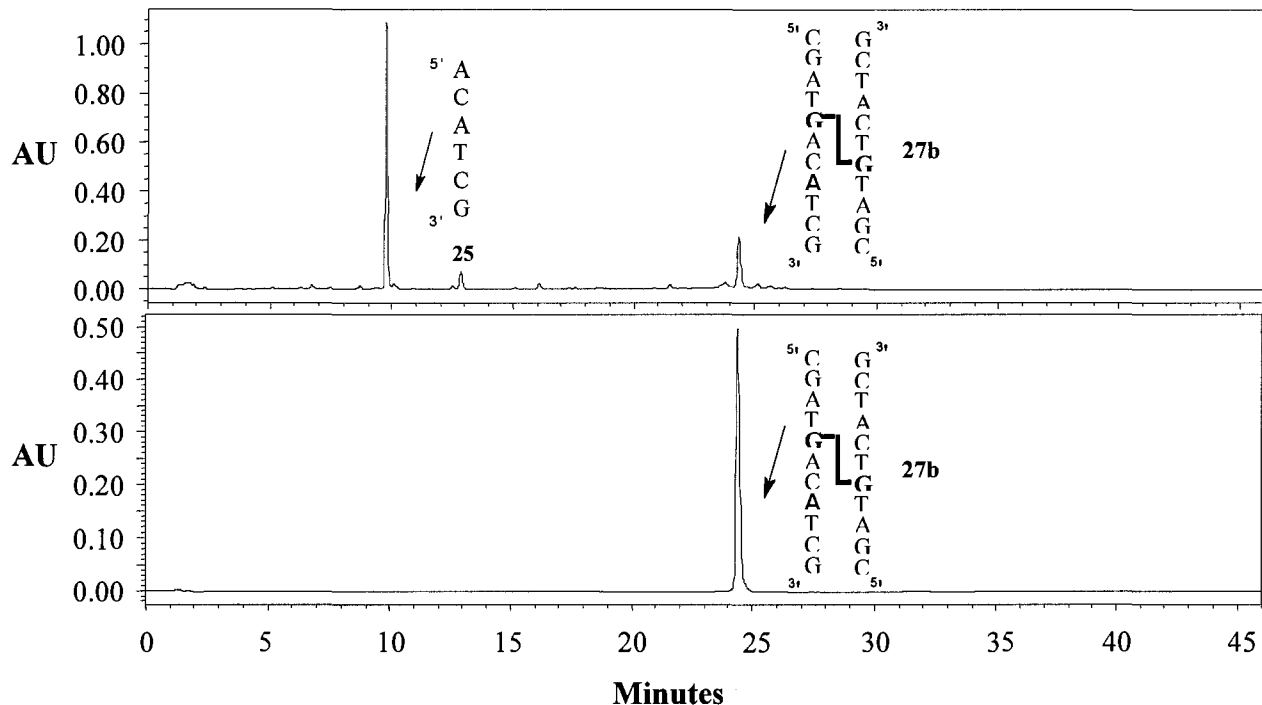
Figure 3.7. Analytical chromatograms using SAX HPLC of the crude synthesis of duplex **27b** and final purified product.
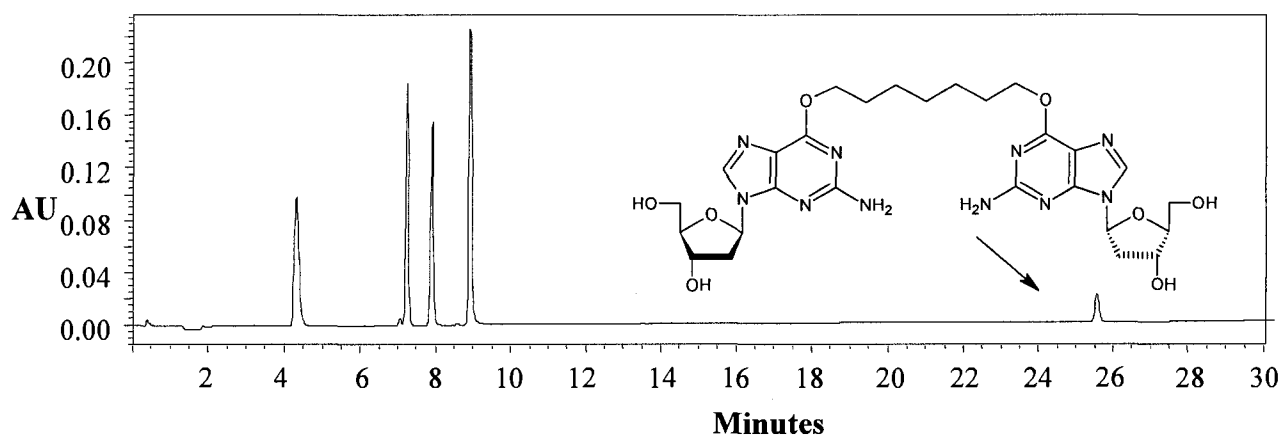
Figure 3.8 HPLC analysis of a snake venom phosphodiesterase and calf intestinal phosphatase digest of **27b** using a Waters 1525 pump and 2487 detector using a Symmetry C18 (4.6 x 150mm, 5μm) column.

Table 3.7 Nucleoside composition ratios corrected for molar extinction coefficients and normalized to one for the cross-linked duplex **27b**. MALDI-TOF: observed 6765.28 (expected 6768.29)

Nucleoside	Peak area	Nucleoside ratios	
		Expected	Observed
dC	828094	1.50	1.53
dG	882950	1.00	1.03
dT	824288	1.25	1.24
dA	1314029	1.25	1.25
dG-dG	159606	0.25	0.25

In light of the degradation observed for the synthesis of **27a** and **27b** using the alloc protecting group, consideration was given to its removal using the higher yielding solution phase reaction on dimer **10b** to generate **13** which was then phosphitylated to yield **14**. Slight modifications to solid phase synthesis protocols were explored as TBS deprotection with TEA·3HF is incompatible with controlled pore glass (CPG) and glass components on the DNA synthesizer. A polystyrene solid support was utilized with off-column deprotection of the Y-sequence and step iii deprotection conditions in Scheme 3.9 were replaced with TEA•3HF. Chromatograms of the post-deprotection yields using TEA•3HF on TBS and Pd(PPh$_3$)$_4$ on alloc were compared with respect to overall yield. Pre-deprotection coupling yields on polystyrene were comparable to CPG yields. TEA•3HF removal of the 3'protecting group generated the most favourable results with no observable decomposition during TEA•3HF treatment (Figure 3.9).
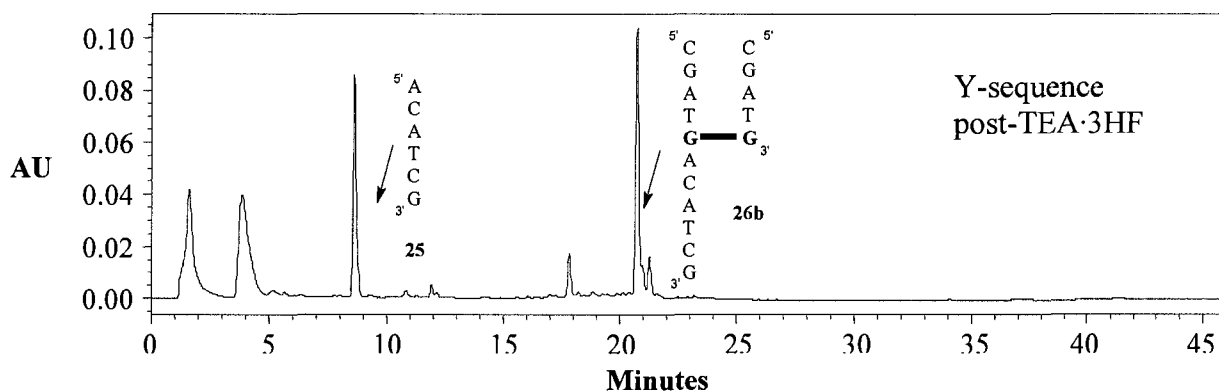
Figure 3.9. SAX-HPLC chromatogram demonstrating improved yield from deprotection of 3'TBS using TEA•3HF on a polystyrene solid-support.

Previously published data on **24** serves to demonstrate the destabilizing effect that the mismatched 2'-deoxyguanosine base pairs exert as indicated by a 23°C shift in dissociation temperature for **22**.[40] It should be noted that **22** only has one set of base pairs' hydrogen bonding disrupted in contrast to the two disrupted in **27b**. Sufficient quantities of **27b** were obtained from the initial alloc deprotection strategy to permit T_m and CD studies. Comparison of **27b** (51.0°C) to its respective non-cross-linked control duplex **28** (50.0°C) indicates only a minimal increase in the stability of the duplex which is inside the range of experimental error (±1.0°C) (Figure 3.10). A sigmoidal denaturation curve qualitatively similar in shape to that of **28** was exhibited for **27b**. Unlike **22** where control duplexes **24** and to a lesser extent **23** have similar hyperchromicity profiles, the degree of hyperchromicity for **27b** is 48% of **28** at each respective melting temperature. The lower net change in hyperchromicity suggests a significant amount of perturbation to the base stacking in **27b** induced by the ICL and its disruption of two WC base pairings to a degree not present in **22**. The near identical dissociation temperatures appears to be the result of the entropic stabilization of the two strands remaining unimolecular being balanced by destabilization from the disruption of WC base pairing induced by the presence of the 1,3-

ICL. Evidence for the duplex destabilization and entropic effects balancing each other is displayed in the magnitude of hyperchromicity for **27b** along side the dissociation temperature for **22** which supports this hypothesis.
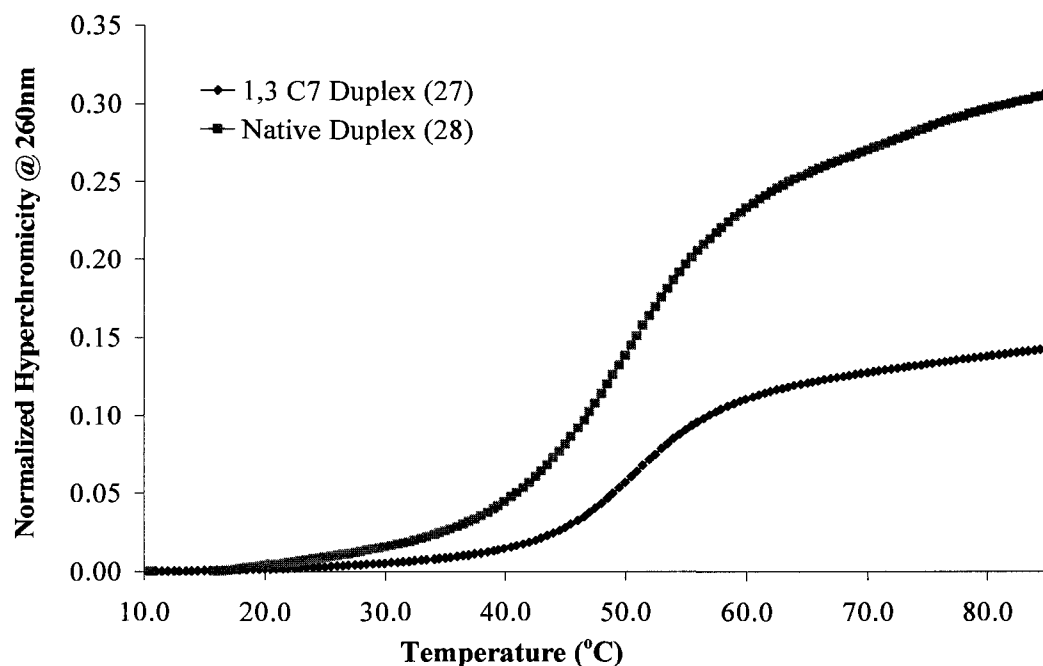
Figure 3.10. UV thermal denaturation profile ($\Delta 0.5°C/min$) of the cross-linked (**27b**), non-cross-linked control duplex (**28**) (5'-dCGATGACATCG)$_2$.

The CD spectra of **27b** and **28** duplexes are displayed in Figure 3.11. Duplex **28** exhibits a spectrum consistent with the signature of natural B-form DNA having its maximum at 281 nm, a minimum at 251 nm, and a crossover at 266 nm. The CD spectrum of **27b** displays a degree of similarity with the general shape of the control duplex while having shift to lower frequencies for spectral features. The minimum of **27b** is at 247 nm, the crossover at 261 nm, and the maxima at 278 nm. At low and high wavelength (<211 and >290 nm) the molar ellipticities of **27b** and **28** become convergent. The minimum of **27b** has a deviation from the signature pattern exhibited by the **28** having 0.48 times less molar ellipticity than **28** in this region. The contrasts to **22** showing 1.6 times greater molar ellipticity than its respective control **23** in its minimum area. The magnitude of the maxima

for **27b** is 0.86 times the magnitude of **28** more closely approximating the relationship

observed between **22** having 0.57 times less ellipticity than **23**. These differences could

result from changes in base stacking interactions induced from mismatched $dG^*{\cdot}dG^*$ in **22**

when compared to two $dG^*{\cdot}dC$ base pairs (where dG^* is O^6-alkylated). Modified duplexes'

ability to transfer induced dipoles between bases when compared to their respective

controls is a likely source of the observed divergences. The overall signature similarity of

both **22** and **27b** to that of their respective controls suggests the seven carbon ICL are well

tolerated in both ICL duplexes.
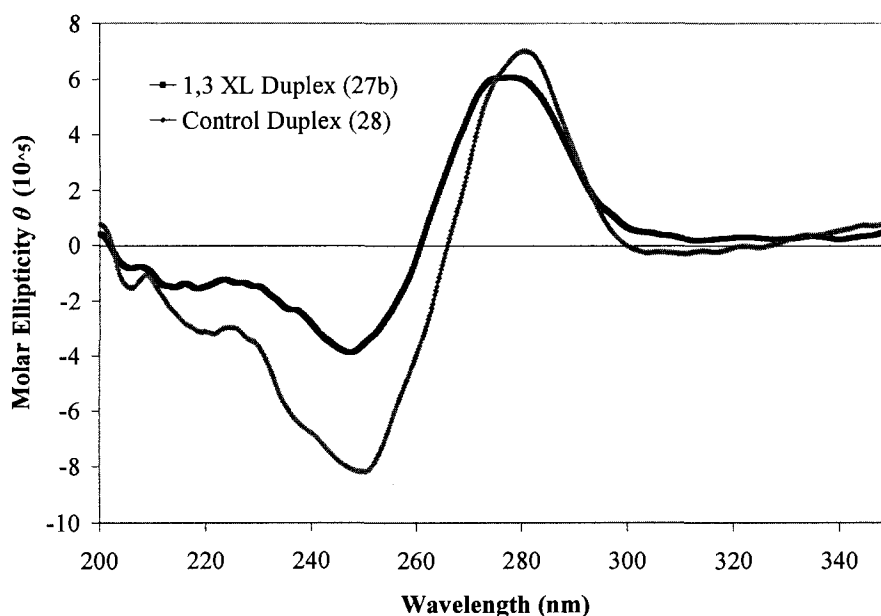
Figure 3.11 CD profile of the cross-linked oligonucleotide **27b** and control duplex **28**.

3.2.2. Tripartate synthesis

Advancement to the tripartate synthesis of **30** permits complete control over sequence

composition around the cross-link (Scheme 3.10). It represents the highest degree of

asymmetry in a DNA segment being reflected in many native DNA sequences. An

additional complexity is introduced in that the tripartate dimer **19** has the 3' and both 5'-

alcohols orthogonally deprotected while still being compatible with solid-phase synthesis.

The potential exists for optimized drying and silyl deprotection methodologies from the bipartate strategy to be directly transferred to the tripartate synthesis. The tripartate strategy involves three off-column steps for the coupling of **19** to the linear sequence in step i, the 5'-silyl deprotection in step iii, and the final removal of the 3'-levulinyl group in step v (Scheme 3.10).

Scheme 3.10 Proposed stepwise synthesis of ICL duplex **30** using the tripartate strategy permitting complete asymmetry induction.
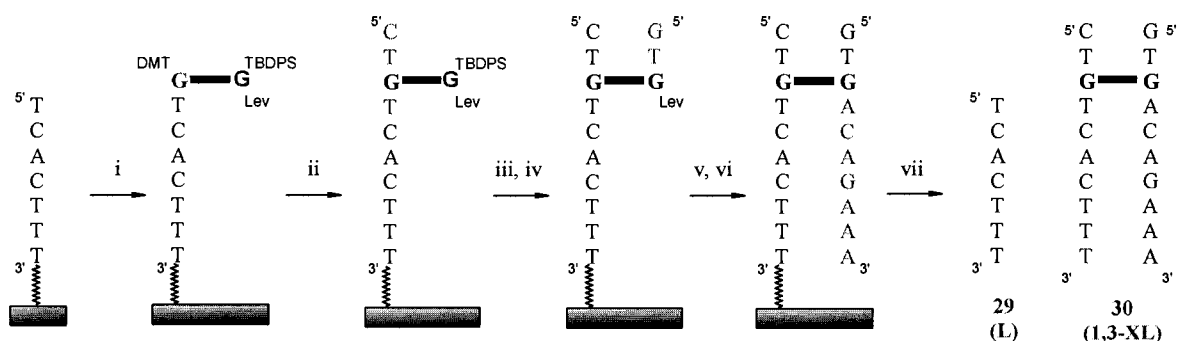
Conditions: (i) Off-column coupling of monophosphoramidite **19**, (ii) chain growth with 3'-O-2'-deoxyphosphoramidites, (iii) off-column TEA•3HF silyl deprotection (iv) continued chain growth with 3'-O-2'-deoxyphosphoramidites, (v) off-column hydrazine deprotection of the levulinyl group (vi) continued chain growth with 5'-O-2'-deoxyphosphoramidites, (vii) cleavage from the solid-support and deprotection using 1:3 ethanol/ammonium hydroxide (28%). A pre-packed 1 μmol polystyrene column was used.

The tripartate synthesis is begun in the same manner as the mono and bipartate strategies with the generation of a linear sequence. As in the bipartate methodology, this is followed by coupling of a 3-*O*-monophosphoramidite **19** in an off-column fashion (Table 3.8). Only one attempt was made at synthesizing duplex **30** which was performed concurrently with **27a** and **27b**. Once the L-sequence cycle had been completed the columns were removed from the DNA synthesizer. Vacuum dried **19** was added to separate screw-cap microfuge tube with the L-sequence stationary phase. A tetrazole solution (0.45 M, 100 μL) was added to the microfuge tube to dissolve amidite **19** upon which the sample

was left to react for 1 hour. After reacting the solid-support was placed into DNA synthesis columns followed a rinsing step using a vacuum apparatus with THF (30 mL) followed by ACN (30 mL). The column was then placed under high vacuum for ~10 min too remove to residual solvents and then placed on the DNA synthesizer to resume 3'→ 5' chain extension to form the Y-sequence product. Unfortunately during treatment with TCA no trityl cation was observed to elute from the column. As such the tripartate synthesis was not pursued until the optimized bipartate conditions could be applied.

Table 3.8 Quantities required for solid-phase synthesis of ICL duplex **30** using 3'-O-phosphoramidites to generate the Y-sequence 17-mer using off-column coupling of **19**.

Phosphoramidite	Number of Incorporations	Mass (mg)	Concentration (mol·L^{-1})	Volume* (µL)
dA	6	183.5	0.100	1500
dG	2	142.8	0.100	700
dC	4	195.5	0.100	1100
dT	6	141.5	0.100	1500
19	1	34.8	0.200	100

* Solvent is ACN.

3.3 Molecular Modeling

3.3.1 Comparison of Previous Modeling to Hyperchem Model for 1,3-ICL Orientations

Initial studies by Colvin $et.$ $al.$ on the connectivity of the N^7-N^7-ICL formed by HM were complemented by molecular modeling to discern if the seven carbon chain would span the N^7-N^7 distance within a typical DNA duplex.[24] Modeling studies in Macromodel (v.4.0) using AMBER force field with the GB/DA solvation demonstrated the heptyl alkyl linkage fitting into the major groove of DNA resulting in the generation of a slight wobble in the dG*·dC base pairs.[24] The length of a 1,7-diaminoheptane chain (9.8 Å) was calculated to be sufficiently long to bridge the N^7-N^7 distance for a typical x-ray structure of a 5'-GXC-3' sequence (8.9 Å) with minimal distortion in the DNA (Figure 3.12).[24] Using a dodecamer (ACTAGGCCTAGT) ICL duplex, Colvin calculated the optimized N^7-N^7 distance to be 8.05 Å and the same unmodified duplex to be 9.25 Å.[24] These distances diverge from values obtained using Hyperchem modeling software (geometry optimized using the AMBER forcefield in $vacuo$) for an ICL sequence producing a shorter N^7-N^7 distance of 6.41 Å and the control duplex to be 8.64 Å (Table 3.9). The Hyperchem model displays stronger correlation for N^7-N^7 distances for non-modified duplexes to typical of x-ray structures (Hyperchem: < 0.25 Å, Colvin: > 0.35 Å). The observed divergence in models is greater when the degree of distortion for respective ICL duplexes are compared for N^7-N^7 distances with Colvin's model (1.2 Å closer) against the Hyperchem model (2.23 Å closer) (Figure 3.13, 3.14, 3.15). It is possible that a Hyperchem model factoring in hydration effects could have a dramatic influence on the geometry of the phosphate backbone producing a stronger correlation between the two models, particularly in the ICL structures.

Overall the divergences in the geometry prediction between the two calculated models make direct comparison suspect.
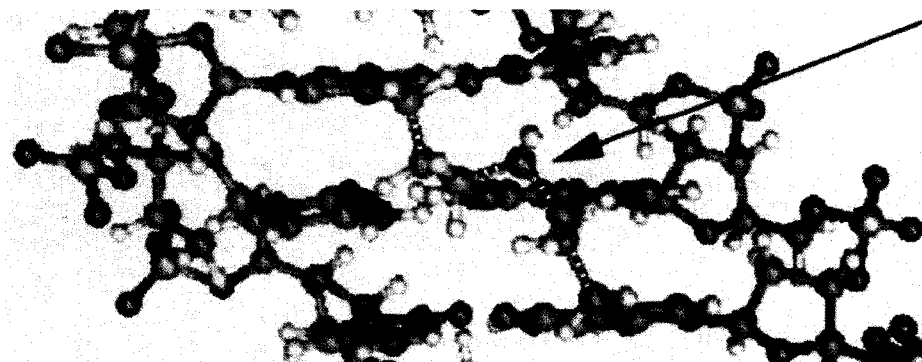
Figure 3.12 Theoretical model of the clinically relevant 1,3 N^7dG-heptyl-N^7dG cross-link lesion. The heptyl chain is outlined with a dashed black and white bond.[24]
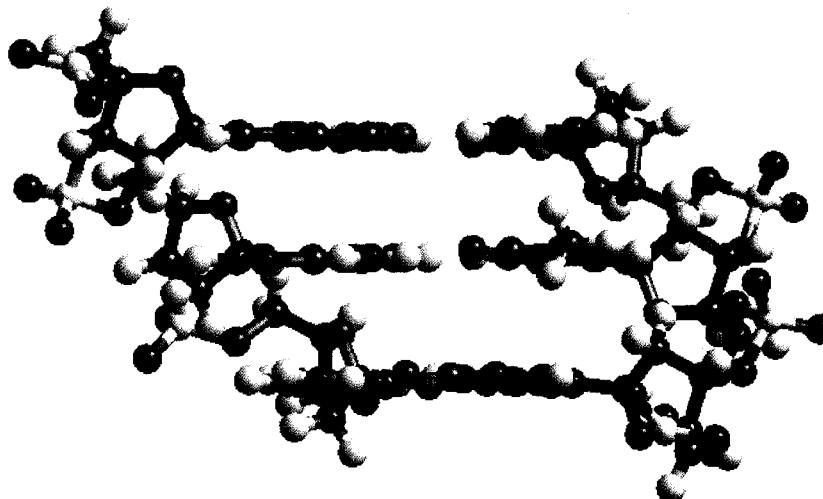
Figure 3.13 Control duplex **28**. Only three base-pairs of the eleven base-pair duplex are shown for clarity. The calculated O^6-O^6 distance is 6.41 Å and the N^7-N^7 distance is 8.64 Å
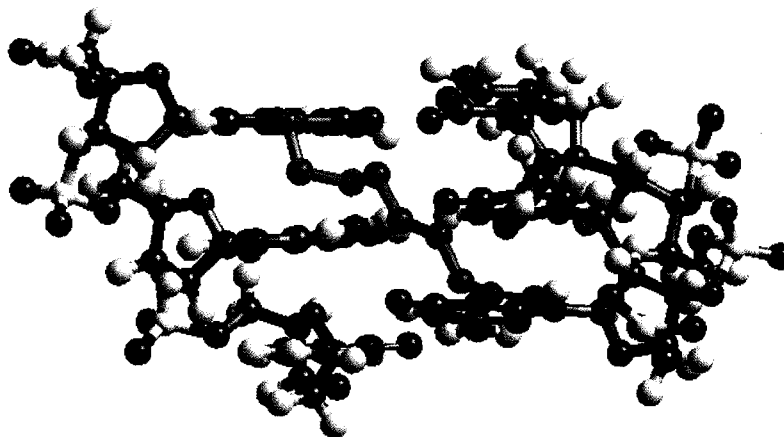
Figure 3.14. The 1,3-N^7dG-heptyl-N^7dG duplex. Only three base-pairs of the eleven base-pair duplex are shown for clarity. The heptyl chain is coloured green. The calculated O^6-O^6 distance is 5.47 Å and the N^7-N^7 distance is 6.41 Å
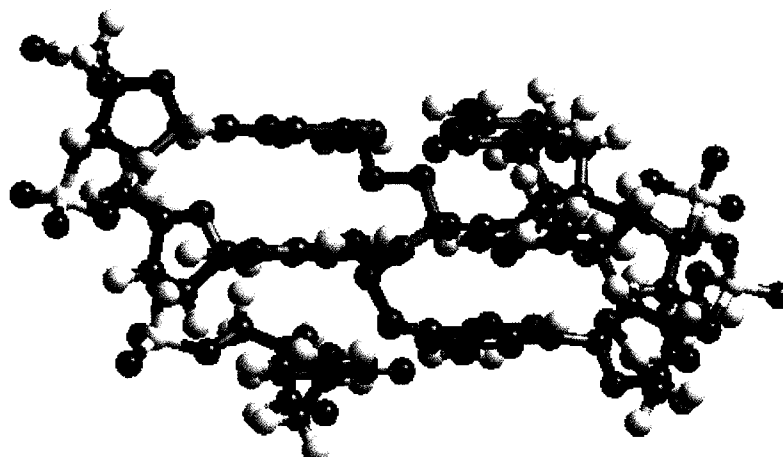
Figure 3.15. $1,1\text{-}O^6\text{dG-heptyl-}O^6\text{dG}$ duplex **27b**. Only three base-pairs of the eleven base-pair duplex are shown for clarity. The heptyl chain is coloured green. The calculated O^6-O^6 distance is 5.76 Å and the $N^7\text{-}N^7$ distance is 8.90 Å.

Table 3.9 Distances between shared reference point atoms for duplexes **27a**, **27b**, **28**, and their respective $N^7\text{-dG-alkyl-}N^7\text{-dG}$ ICL duplexes.

Nucleotide Duplex	$O^6\text{-dG-}O^6\text{-dG}$ Distance	$N^7\text{-dG-}N^7\text{-dG}$ Distance
28	6.41 Å	8.64 Å
$1,3\text{-}N^7\text{-dG-butyl-}N^7\text{-dG}$	5.32 Å	5.95 Å
27a	5.03 Å	8.71Å
$1,3\text{-}N^7\text{-dG-heptyl-}N^7\text{-dG}$	5.47 Å	6.41 Å
27b	5.76 Å	8.90 Å

The exact connectivity of the lesion formed by BU has to date not yet been discerned but it is theorized to form more intrastrand cross-links and a 1,3-lesion similar to HM.[3] Molecular models were calculated for BU using the clinically relevant $1,3\text{-}N^7\text{-}N^7$ connectivity for HM and the $O^6\text{-}O^6$ connectivity for duplex **27a**. The shorter butyl linker in a 1,3-lesion would not span the $N^7\text{-}N^7$ distance inducing a larger amount of distortion from the parent conformation for both $O^6\text{-}O^6$ and to a much greater extent for $N^7\text{-}N^7$ lesion models (Figure 3.13, 3.16, 3.17). As expected the $1,3\text{-}N^7\text{-dG-butyl-}N^7\text{-dG}$ lesion displays the greatest amount of distortion from **28** as interbase distance displays the shortest $N^7\text{-}N^7$ (5.95 Å) while **27a** displays the shortest $O^6\text{-}O^6$ distance (5.03 Å). This distortion in **27a** would be expected to produce a larger departure from a B-form signature and would alter

79

the T_m profile to a when compared with data obtained for **27b**. The larger distortion induced by a BU N^7-N^7 linkage could result in increased recognition and repair by a cell accounting for higher efficacy demonstrated by HM.
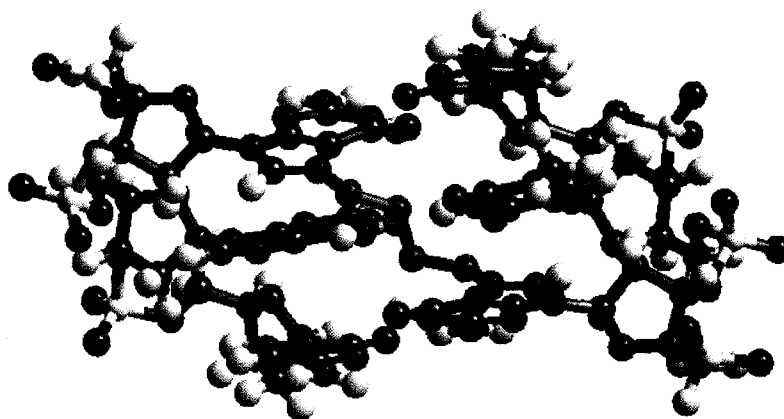
Figure 3.16. 1,3-N^7dG-butyl-N^7dG duplex. Only three base-pairs of the eleven base-pair duplex are shown for clarity. The butyl chain is coloured green. The calculated O^6-O^6 distance is 5.32 Å and the N^7-N^7 distance is 5.95 Å
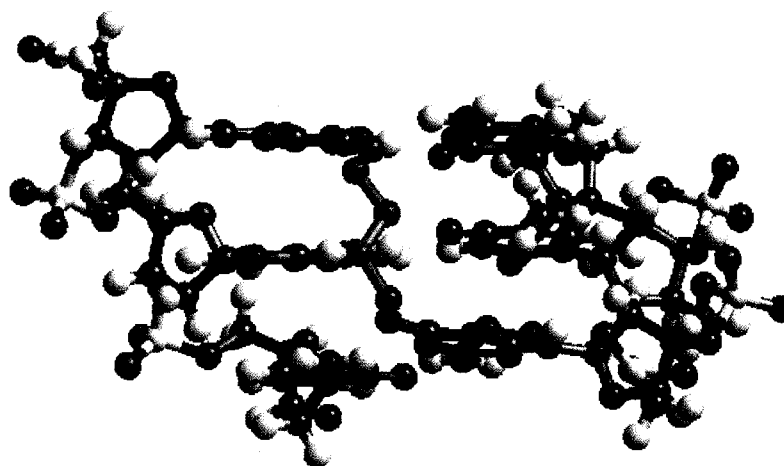
Figure 3.17. 1,3-O^6dG-butyl-O^6dG duplex **27a**. Only three base-pairs of the eleven base-pair duplex are shown for clarity. The butyl chain is coloured green. The calculated O^6-O^6 distance is 5.03 Å and the N^7-N^7 distance is 8.71Å

3.3.2 Hyperchem Model for 1,1 Duplexes

Molecular modeling was used to provide a comparative analysis for the possible structural effects produced by the presence of mismatched base pairs and chemically modified duplexes (Table 3.10, Figure 3.18, 3.19, 3.20). It can be seen that both the mismatched duplex **23** and ICL duplex **22** are expected to see a distortion from the parent

control duplex **24**. Calculations predict an increase in the 1,3-N^7-dA-N^7-dA distance in opposing strands for both non-WC duplexes. The mismatch modification in **22** and **23** are seen to push the phosphate backbone apart and displace both 2'-deoxyguanosines in opposite directions out of planarity from normal base stacking geometry. It is unknown if the change in hydrogen bonding pattern from **22** to the non-modified bases of **23** has a similar affect on duplex T_m values. Both duplexes display a similar shift from planarity in the heterocyclic bases greatly distorting normal hydrogen bond geometry. Lastly it can be seen that the presence of the ICL of **22** induces a greater distance for both O^6-dG-O^6-dG and N^7-dA-N^7-dA measurements.

Table 3.10 Distances between shared reference point atoms for duplexes **22**, **23**, and **24**.

Nucleotide Duplex	O^6-dG-N^4-dC Distance	O^6-dG-O^6-dG Distance	N^7-dA-N^7-dA Distance
22		4.74 Å	9.02 Å
23		4.09 Å	8.75 Å
24	4.74 Å		8.68 Å
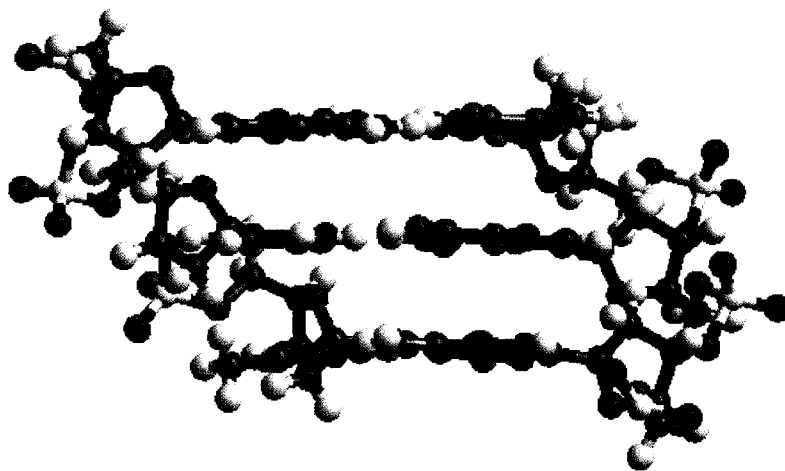
Figure 3.18 dG·dC native control duplex **24**. Only three base-pairs of the eleven base-pair duplex are shown for clarity. O^6dG-N^4dC distance is 4.74 Å N^7dA-N^7dA distance is 8.68 Å
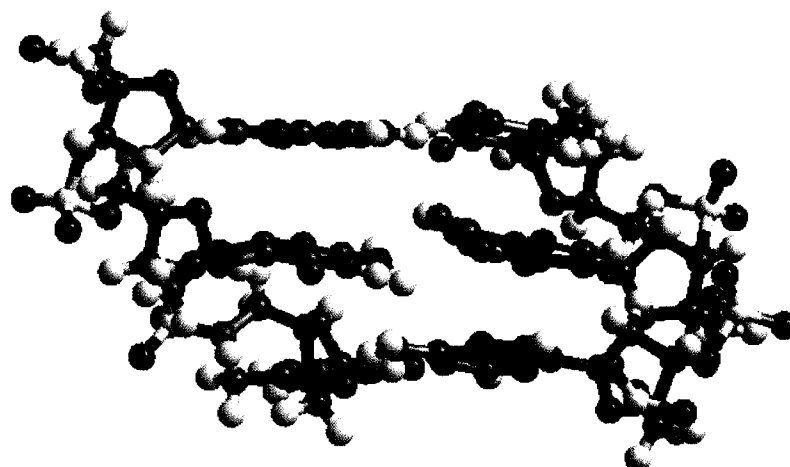
Figure 3.19 dG·dG mismatch control duplex **23**. Only three base-pairs of the eleven base-pair duplex are shown for clarity. O^6dG-O^6dG distance is 4.09 Å N^7dA -N^7dA distance is 8.75 Å
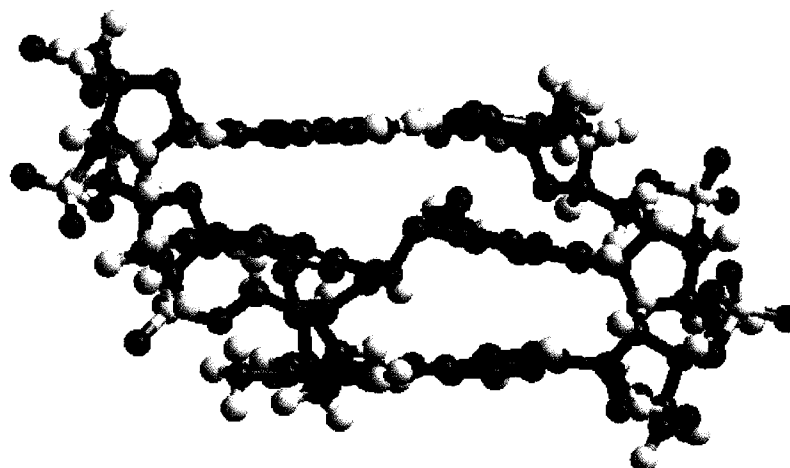
Figure 3.20 1,1-O^6dG-heptyl-O^6dG- mismatch duplex **22**. Only three base-pairs of the eleven base-pair duplex are shown for clarity. O^6dG-O^6dG distance is 4.74 Å N^7dA-N^7dA distance is 9.02 Å

3.3.3. Ada-C Direct Repair SAR from NMR Model

A closer examination of SAR results suggests that an O^6dG-alkyl-O^6dG or HM ICL might be inaccessible to the active site nucleophile Cys145 if the modified base remained stacked within the DNA duplex (Figure 1.10).[82] The potential inability of an O^6dG-alkyl-O^6dG BU or HM interstrand cross-linked 2'-deoxyguanosine to enter the active site of the enzyme could result in AGT's failure to recognize the lesion. Synthesis of a DNA segment containing a BU O^6dG-alkyl-O^6dG ICL of defined position and orientation could

offer insight into the ability of O^6-BG to potentate BCNU via depletion of AGT while leaving BU toxicity unaffected. Length of the alkyl chains in BU and HM could also affect the pharmacokinetics of repair due to an increased ability of a longer HM chain to fit into the binding pocket of AGT. A BU or HM ICL could also act as an absolute block to AGT if the nucleobase was prevented from becoming fully extrahelical and thus not allowed to enter the active site of the enzyme resulting in transition state inhibition. Alternatively if alkyl transfer occurred the AGT protein would be covalently tethered to the DNA duplex. It is uncertain if alkylation of this type would allow for the conformation changes within AGT that typically facilitate a mechanism of dissociation from DNA thus allowing cellular recognition prompting proteolytic degradation of alkylated AGT. This inability to recognize the covalently linked enzyme coupled with an inability to degrade the tethered enzyme could trigger apoptosis. Resultantly synthesis of structures of defined structure, composition and sequence would allow for the systematic investigation of the clinical relevance these repair pathways to determine the role of O^6-alkylation in the toxicity of BU and HM.

Chapter 4: Conclusions and Future Directions

In summary, a synthetic methodology to prepare DNA duplexes containing an O^6dG-alkyl-O^6dG dimer amenable to the conditions of solid phase synthesis has been developed. The first monopartate dimer was incorporated into a one helical turn duplex processing an ICL of defined structure and sequence. Biophysical studies consisting of CD measurements which confirm **22** exists as B-form DNA and T_m studies illustrating the presence of the 1,1-ICL has a stabilizing effect of 12°C over **23** and 23°C over **24**. Observed changes in hyperchromicity were minimal indicating the 1,1-lesion was well tolerated inducing little deformation from control duplex **23**. Sufficient quantities were obtained to permit biological and NMR structural studies which are currently ongoing with collaborators.

Synthesis of the bipartate dimer allowed for the incorporation of an additional unit of asymmetry around the site of the ICL in **27b** permitting the synthesis of the clinically relevant orientation. Yields of **26b** were low due to decomposition encountered with overnight Pd(PPh$_3$)$_4$ deprotection. It is likely that the order of magnitude increase in the amount of Pd(PPh$_3$)$_4$ along with longer reactions times were responsible for this observation. Reaction conditions using lower amounts of catalyst loading for shorter periods of time in argon atmosphere could be explored to remedy this problem. Progressing with 5'-O-phosphoramidites after deprotection on **26b** yielded sufficient quantities of **27b** which was isolated for biophysical characterization. Enzymatic digests demonstrated good agreement with observed versus expected composition. CD studies revealed **27b** to be in B-form DNA conformation. T_m dissociation temperatures for **28** were

1°C lower than **27b** which also showed reduced hyperchromicity (48%) reflective of a probable disruption to base stacking.

Subsequent experiments resulted in a convenient route to circumvent decomposition observed with the removal of the alloc group during solid-phase synthesis. Previous alloc removal using solution phase chemistry resulted in high yields (90%) as such the alloc of **10** was removed in place of TBS to yield **13** which was phosphitylated to yield **14**. Deprotection conditions with TEA•3HF necessitated a change from a CPG to a polystyrene solid support. Analysis of samples before and after silyl deprotection demonstrated no decomposition of the oligonucleotide after TEA•3HF treatment. No further synthesis was performed on the TEA•3HF deprotected **26b** as it is believed that the quantity of **25** present was due to moisture present in **14**. Drying conditions require improvement for the bipartate phosphoramidite dimer as traces of water after overnight drying in THF with excess Na_2SO_4 were observed by NMR before another synthesis will be attempted. Optimization of the drying of **14** could result in near quantitative yields common to coupling of commercially available phosphoramidites.

After the synthesis of **26b** is optimized using TEA•3HF deprotection, sufficient quantities of **27a** and **27b** can be synthesized for NMR and X-ray crystallographic structural studies. This will enable the possibility to discern the solution phase and solid state torsion angles about each DNA segment containing the 1,3-orientation of the ICL. Comparisons can be made between the amounts of distortion induced by the 4 versus 7 carbon linkers. Enzymatic repair rates of ICL duplexes which have been incorporated into plasmids can then be correlated in SAR studies in terms of bending and helical unwinding versus enzymatic recognition and repair rates.

The final synthesis of the tripartate pre-amidite dimer was obtained in high yield permitting long term storage of the compound. Optimized bipartate silyl removal procedures offer the chance of method transfer to the tripartate synthesis of duplex **30**. Preliminary deprotection tests display compatibility of all protecting groups with overnight levulinyl deprotection using hydrazine. The differential protection at all alcohol functionalities will permit stepwise addition of each nucleotide allowing for each chain to be unique. Enzymatic digest and MALDI-TOF analysis will be performed to confirm duplex composition. The completely asymmetric tripartate 1,3-duplex having 'sticky ends' will then be integrated into a DNA plasmid permitting repair studies. Studies are ongoing to determine the repair pathways involved in these lesions and their function as BU and HM mimics.

Chapter 5: Experimental

General.

Protected 2'-deoxyribonucleoside-CPG supports, 3'-O-(β-2-cyanoethyl-N,N'-diisopropyl)-5'-O-dimethoxytrityl-N^2-isobutyryl-2'-nucleotide and 3'-O-dimethoxytrityl-5'-O-(β-2-cyanoethyl-N,N'-diisopropyl)-N^2-isobutyryl-2'-nucleotide phosphoramidites were purchased from Proligo, Inc. (Boulder, Colorado). All anhydrous solvents were bought from EMD Chemicals (Gibbstown, New Jersey). All deuterated solvents were acquired from Cambridge Isotope Laboratories, Inc. (Andover, MA). Solid phase oligonucleotide syntheses were carried out with 500 Å pore size CPG solid support with a nucleoside loading of 65 μmol/g or prepacked 1 μmol polystyrene solid support for the synthesis of duplex **27b**. Silica gel and aluminum backed TLC plates were purchased from Silicycle (Ville de Québec, Québec). All other chemicals were obtained from Sigma-Aldrich Chemical Company (Milwaukee, Wisconsin) and used as is. Strong anion exchange (SAX) HPLC was carried out using a Dionex DNA PAC PA-100 column (0.4 cm x 25 cm) purchased from Dionex Corp. (Sunnyvale, California) on a Waters Breeze HPLC system. Reversed phase HPLC was carried out using a Symmetry C-18 5 μm column (0.46 cm x 15 cm) purchased from Waters Inc. (Milford, Massachusetts). Injected samples were monitored at 260 nm for analytical runs or 290 nm for preparative runs. Denaturing polyacrylamide gel electrophoresis was carried out on a 20 cm x 20 cm x 0.75 cm gel containing 20% acrylamide and 7 mol/L urea in TBE (89 mmol/L Tris. 89 mmol/L boric acid, and 0.2 mmol/L ethylenediaminetetraacetate buffered at pH 8.0). The running buffer was TBE. Oligonucleotides were detected by UV shadowing and developed with STAINS-ALL™. MALDI-TOF mass spectra for oligonucleotides were conducted at the John

Hopkins University Mass Spectrometry Facility using a Voyager DE-STR MALDI-TOF instrument. ESI mass spectra for all small molecules were collected at the Department of Chemistry at McGill University using a Kratos MS25RFA mass spectrometer. NMR spectra (^{1}H, ^{13}C, and ^{31}P) were recorded using a Varian INOVA 300 MHz spectrometer. Chemical shifts are relative to that of CDCl$_3$, DMSO-d_6, or Acetone -d_6 (^{1}H and ^{13}C) as internal standard or H$_3$PO$_4$ (^{31}P) as external standard.

Oligonucleotide synthesis and purification.

The cross-linked structures **22**, **27a**, **27b** were assembled using an Applied Biosystems Model 3400 synthesizer according to standard cyanoethylphosphoramidite chemistry protocols.[133] Respective control duplexes were assembled as individual linear chain sequences and annealed. Long chain alkylamine (LCAA) CPG was used as the solid support. The nucleoside phosphoramidites were diluted in anhydrous acetonitrile at a concentration of 0.10 mol/L for the 3'- and 5'-O-phosphoramidites. Assembly of the sequences was carried out as follows: (*a*) detritylation with 3% TCA in DCM; (*b*) nucleoside phosphoramidites coupling time of 120 sec for commercial phosphoramidites, 180 sec coupling time for nucleoside phosphoramidites couplings after the installation of a dimeric nucleoside, and 600 sec for the monopartate dimer **8b** using on-column methods; (*c*) capping with 1:1 (*v/v*) of phenoxyacetyl anhydride-2,6,lutidine-THF 1:1:8 (*v/v/v*. solution A) and 1-methyl-imidazole-THF 16:84 (*w/v*, solution B); (d) oxidation with 0.02 mol/L iodine in THF-water-pyridine, 2.5:2:1. Prior to final analysis the 5'-terminal trityl groups were removed on the synthesizer. The oligomer-derivatized CPG was transferred from the reaction column to screw-capped microfuge tubes fitted with teflon lined caps and the protecting groups were removed by treatment with 1 mL of concentrated ammonium

hydroxide-ethanol (3:1) for 4h at 55°C. The cross-linked duplexes were purified by SAX HPLC with a 30mL linear gradient of sodium chloride in a buffer that contained 100 mmol/L Tris (pH 7.8) in 10% ACN at a flow rate of 1.0 mL/min. The purified oligomers were then desalted using C-18 SEP PAK cartridges (Waters Inc.) as previously described.[131]

Nuclease digestion and MALDI-TOF mass spectrometry

The cross-linked oligomers (0.2 A_{260} unit) were characterized by digestion with a combination of SVPDE (0.28 unit) and calf intestinal phosphatase (5 units) in a buffer containing Tris (pH 8.1, 10 mmol/L) and magnesium chloride (2 mmol/L) as previously described.[131] The resulting mixture of nucleosides was analyzed by reversed phase HPLC and the ratio of nucleosides determined.

Thermal denaturation experiments

Solutions of each crosslinked duplex, mismatched duplex, and control non-cross-linked duplexes were prepared in a buffer containing 90 mmol/L sodium chloride, 10 mmol/L sodium phosphate, 1mmol/L EDTA, pH 7.0 to a final concentration of 2 μM. The molar extinction coefficients of the oligonucleotides and cross-linked duplexes were calculated using nearest-neighbor approximations.[134] Samples were degassed in a vacuum centrifuge prior to thermal denaturation experiments run in a Cary 300 Bio UV-Vis spectrophotometer fitted with a thermostated sample holder and temperature controller. The absorbance at 260 nm was monitored as the samples were heated from 5 to 90 °C at a rate of 0.5°C/min. Denaturation curves were generated by plotting the relative hyperchromicity, defined as (a_t - $a_{5°C}$)/($a_{90°C}$), where a_t is the absorbance at temperature t, $a_{5°C}$ is the absorbance at the initial

temperature of 5°C, and $a_{90°C}$ is the absorbance at the final temperature of 90°C, versus temperature.

Circular dichroism (CD) spectroscopy

Solutions of each crosslinked duplex, mismatched duplex, and control non-cross-linked duplexes were prepared in a buffer containing 90 mmol/L sodium chloride, 10 mmol/L sodium phosphate, 1 mmol/L EDTA, pH 7.0 to a final concentration of 2 μM. The samples were equilibrated for 5-10 min at 5°C, and their CD spectra were recorded on a Jasco J-700 spectropolarimeter equipped with a NESLAB RTE-111 circulating bath. Spectra were collected at a rate of 100 nm/min with a bandwidth of 1 nm and sampling wavelength of 0.2 nm using fused quartz cells (Hellma, 165-QS). The molar ellipticity was calculated from the equation $[\theta] = \theta / Cl$, where θ is the relative ellipticity (mdeg), C is the concentration of the oligonucleotide (mol/L) and l is the path length of the cell (cm). The data was processed on a PC computer using software supplied by the manufacturer and transferred into Microsoft Excel for presentation. Each spectrum shown is the average of 5 scans.

Molecular modeling studies

Molecular models of 11bp duplexes (5'-dCGAAAGTTTCG-3')$_2$ and (5'-dCGATGACATCG-3')$_2$ were performed as controls. Corresponding mismatched 2'-deoxyguanosine at the underlined site (5'-dCGAAA<u>G</u>TTTCG-3')$_2$ as well as O^6dG-heptyl-O^6dG interstrand cross-links linking 2'-deoxyguanosine at underlined sites (5'-dCGAAA<u>G</u>TTTCG-3')$_2$ (**XL 1-1**), N^7dG-butyl-N^7dG, N^7dG-heptyl-N^7dG, O^6dG-butyl-O^6dG and O^6dG-heptyl-O^6dG (5'-dCGAT<u>G</u>ACATCG-3')$_2$ (**XL 1,3**) were built using HyperChem molecular modeling software. All duplexes were geometry-optimized using the AMBER forcefield.

Chemical synthesis

Allyl 1-Benzotriazolyl Carbonate (AllocOBt) : To a solution of 1-hydroxybenzotriazole (6.06 g, 39.6 mmol) and DIPEA (5.76 g, 44.6 mmol) in THF (30 mL) was added with vigorous stirring Alloc-Cl (5.12 g, 41.2 mmol) at room temperature. After 30 min, the resulting precipitates were removed by filtration and washed with THF (20 mL x 2). The filtrate and washings were collected and evaporated to give a solid that was loaded onto silica gel in CH_2Cl_2 and eluted furnishing the pure AllocOBt; mp 107-111°C (7.09g, 81% yield) as a white powder; ^{1}H NMR ($CDCl_3$, $CHCl_3$ as reference) 7.5-8.3 (m, 4H, Ar), 6.14 (m, 1H , $CH_2CH=CH_2$), 5.4-5.7 (m, 2H, $CH_2CH=CH_2$), 5.07 (dt, 2H, J = 5.4 and 1.8 Hz, $CH_2CH=CH_2$). NMR spectrum was identical to the original Aldrich product.

1-O-*tert*-Butyldimethylsilylbutanol : To a solution of 1,4-butanediol (31.53 g, 349.9 mmol) and imidazole (2.40 g, 35.3 mmol) was added DCM (20 mL) and THF (10 mL) to form a homogeneous solution. *tert*-Butyldimethylsilyl chloride (1.11 g, 7.35 mmol) was solvated in THF (10 mL) and added dropwise over 5 min and the reaction left overnight. The sample was rotary evaporated and an extraction performed using DCM and 5% sodium bicarbonate (x3) and dried over magnesium sulphate. The sample was purified via silica gel chromatography using hexanes:ethylacetate (20:0 → 17:3) to yield a viscous liquid; 0.734 g (48.9% yield); R_f (SiO_2): 0.35, hexanes:ethylacetate 4:1; ^{1}H NMR (DMSO-d_6, DMSO as reference): 4.36 (t, 1H, OH); 3.62 (t, 2H, CH_2); 3.34 (t, 2H, CH_2); 1.39-1.59 (m, 4H, CH_2); 0.98 (s, 9H, $SiC(CH_3)_3$); 0.03 (s, 3H, $SiCH_3$); -0.04 (s, 3H, $SiCH_3$).

1-O-*tert*-Butyldiphenylsilylbutanol : To a solution of 1,4-butanediol (22.23 g, 246.7 mmol) and imidazole (0.696 g, 10.2 mmol) was added THF (10 mL) to form a homogeneous solution. *tert*-Butyldiphenylsilyl chloride (0.519 g, 1.89 mmol) was

solvated in THF (10 mL) and added dropwise over 5 min and the reaction left overnight. The sample was rotary evaporated and an extraction performed using DCM and 5% sodium bicarbonate (x3) and dried over magnesium sulphate. The sample was purified via silica gel chromatography using hexanes:ethylacetate to yield a viscous liquid; 0.562 g (90.6% yield); R_f (SiO$_2$): 0.35, hexanes:ethylacetate 4:1; ^{1}H NMR (DMSO-d_6, DMSO as reference): 7.57-7.63 (m, 4H, Ar); 7.38-7.49 (m, 6H, Ar); 4.37 (t, 1H, OH); 3.64 (t, 2H, CH$_2$); 3.36 (q, 2H, CH$_2$); 1.43-1.61 (m, 4H, CH$_2$); 0.98 (s, 9H, SiC(CH$_3$)$_3$).

1-O-*tert*Butyldiphenylsilylheptanol : 1,7-heptanediol (1.75 g, 13.3 mmol) and imidazole (2.80 g, 41.1 mmol) was added to THF (10 mL) to form a homogeneous solution. *tert*-Butyldiphenylsilyl chloride (3.46 g, 12.5 mmol) was added dropwise over 20 min to a vigorously stirred solution. A white precipitate was visible near the end of addition of *tert*-butyldiphenylsilyl chloride after which the reaction left overnight. The sample was rotary evaporated to dryness and an extraction performed using DCM and 5% sodium bicarbonate (x3) and dried over magnesium sulphate. The sample was purified via silica gel chromatography using hexanes:ethylacetate to yield a viscous liquid; 2.57 g (52.2% yield); R_f (SiO$_2$): 0.25, hexanes:ethylacetate 4:1; ^{1}H NMR (DMSO-d_6, DMSO as reference): 7.57-7.63 (m, 4H, Ar); 7.38-7.49 (m, 6H, Ar); 4.37 (t, 1H, OH); 3.61 (t, 2H, CH$_2$); 3.36 (t, 2H, CH$_2$); 1.18-1.58 (m, 10H, CH$_2$); 0.98 (s, 9H, SiC(CH$_3$)$_3$).

N^2-phenoxyacetyl-2'-deoxyguanosine (1) : 2'-deoxyguanosine monohydrate (11.73g, 18.7 mmol) was coevaporated with anhydrous pyridine (2 x 50 mL) and then suspended in pyridine (90 mL) and placed in an icebath (20 min). Trimethylsilyl chloride (28.0 mL, 101 mmol) was added via a syringe and the reaction stirred (1 h). Phenoxyacetyl chloride (6.6 mL, 21.7 mmol) was added and stirred at room temperature (4 h). The sample was cooled

to 0°C and concentrated ammonium hydroxide (28%, 100 mL) was diluted with water (100 mL) then added and allowed to stir (20 min). The sample was rotary evaporated and placed under high vacuum overnight. The sample was suspended in water (500 mL) and CH_2Cl_2 (100 mL), allowed to precipitate, and filtered (x2) to yield 4.03g (81.3% yield) of an light beige powder; TLC (10% MeOH/ CH_2Cl_2), R_f = 0.25, ^{1}H NMR (DMSO-d_6, DMSO as reference): 11.93 (s, 1H, NH1), 11.87 (s, 1H, NH), 8.24 (s, 1H, H8); 7.29-7.37 (t, 2H, Ar); 7.00-7.03 (m, 3H, Ar); 6.20 (dd, 1H, H1', J = 6.7); 5.34-5.42 (d, 1H, H3'), 5.00 (t, 1H, H4'); 4.95 (s, 2H, PhOCH$_2$CO); 4.37-4.43 (m, 1H, OH3'); 3.87-3.94 (m, 1H, OH5'); 3.43-3.57 (dd, 2H, H5', H5"); 2.52-2.59 (m, 1H, H2'); 2.27-2.34 (m, 1H, H2"); ^{13}C NMR δ 61.50, 66.23, 70.51, 83.10, 87.80, 114.58, 120.48, 121.37, 129.58, 137.75, 147.22, 148.19, 154.93, 157.61, 170.98.

5'-O-dimethoxytrityl-N^2-phenoxyacetyl-2'-deoxyguanosine (2) : A solution of **1** (5.82 g, 14.5 mmol) was coevaporated with anhydrous pyridine (3 x 30 mL) and then suspended in pyridine (50 mL), followed by the addition of dimethoxytrityl chloride (5.95 g, 17.4 mmol) at room temperature (16 h). Methanol (10 mL) and NH$_4$OH (28%, 4 mL) were combined and added whereupon the solution became cloudy. The solution was then rotary evaporated to dryness and the crude taken up in DCM (50 mL) and washed twice with 5% sodium bicarbonate (50 mL). The organic layer was then dried over sodium sulphate, evaporated to dryness and then placed on high vacuum overnight followed by silica gel chromatography using DCM: methanol (100:0.8 → 100:2.0) to yield a 8.88 g (87.1% yield) of an light yellow foam; R_f (SiO$_2$): 0.43, CH$_2$Cl$_2$:MeOH 10:1; ^{1}H NMR (DMSO-d_6, DMSO as reference): 11.82 (s, 1H, NH1), 11.78 (s, 1H, NH), 8.13 (s, 1H, H8); 7.14-7.34 (m, 12H, Ar); 6.94-7.04 (m, 3H, Ar); 6.74-6.84 (t, 3H, Ar); 6.26 (dd, 1H, H1', J = 6.3); 5.31-5.40 (d,

93

1H, H3'); 4.85 (d, 2H, PhOCH$_2$CO); 4.36-4.42 (m, 1H, OH3'); 3.93-4.00 (m, 1H, H4');

3.71 (s, 3H, OCH$_3$); 3.69 (s, 3H, OCH$_3$); 3.07-3.23 (dd, 2H, H5', H5"); 2.67-2.75 (m, 1H,

H2'); 2.31-2.38 (m, 1H, H2"); ^{13}C NMR δ 40.56, 55.28, 64.08, 66.98, 72.02, 83.96, 86.48,

86.56, 113.21, 114.87, 121.62, 122.80, 126.98, 127.92, 128.18, 129.95, 130.09, 135.67,

135.71, 137.59, 144.58, 146.50, 147.96, 155.62, 156.63, 158.57, 170.24.

3'-O-alloxycarbonyl-5'-O-dimethoxytrityl-N^2-phenoxyacetyl-2'-deoxyguanosine (3) :

To a solution of compound **2** (2.15 g, 3.06 mmol) in anhydrous THF/pyridine (9:1) was

added dimethylaminopyridine (catalyst) followed by allyl 1-benzotriazoyl carbonate (0.888

g, 4.05 mmol) and the mixture was allowed to stir at room temperature for 24 h. The

solvent was removed, and the crude taken up in DCM (50 mL) and washed twice with 5%

sodium bicarbonate (50 mL). The organic layer was then dried over sodium sulphate and

evaporated to near dryness, which was purified by silica gel column chromatography using

DCM: methanol (100:0.5 → 100:1.4) to yield an off-white foam 1.84 g (76.2% yield); R$_f$

(SiO$_2$): 0.37, CH$_2$Cl$_2$:MeOH 20:1; ^{1}H NMR (CDCl$_3$, CH$_2$Cl$_2$ as reference): 11.79 (s, 1H,

NH1), 8.99 (s, 1H, NH), 7.85 (s, 1H, H8); 7.35-7.44 (m, 4H, Ar); 7.18-7.35 (m, 9H, Ar);

7.12 (t, 1H, Ar); 7.02 (dd, 2H, Ar); 6.82 (dd, 4H, Ar); 6.32 (dd, 1H, H1', J = 6.0); 5.90-6.05

(m, 1H, allyl); 5.31-5.40 (m, 2H, allyl); 4.44-4.50 (m, 1H, H3'); 4.69 (d, 2H, PhOCH$_2$CO);

4.67 (d, 2H, vinylCH$_2$CO); 4.31-4.36 (m, 1H, H4'); 3.79 (s, 6H, OCH$_3$); 3.35-3.51 (dd, 2H,

H5', H5"); 2.83-2.88 (m, 1H, H2'); 2.64-2.70 (m, 1H, H2"); ^{13}C NMR δ: 38.50, 55.51,

63.30, 66.86, 69.24, 79.66, 81.68, 86.73, 86.83, 113.42, 115.10, 119.92, 123.29, 125.21,

127.32, 128.03, 128.10, 129.39, 130.27, 131.31, 138.99, 139.68, 139.74, 146.54, 146.96,

147.60, 154.48, 155.23, 156.56, 158.88, 169.89; MS (ESI), m/z (M + Na): calcd for

$C_{43}H_{41}N_5O_{10}Na^+$, 810.275, found 810.25; m/z (M - H): calcd for $C_{43}H_{40}N_5O_{10}^-$, 786.278, found 786.24.

3'-*O*-tert-butyldimethylsilyl-5'-*O*-dimethoxytrityl-*N²*-phenoxyacetyl-2'-deoxy-guanosine (9) : To a solution of compound **2**, (1.73 g, 2.46 mmol) in anhydrous *N',N'*-dimethylformamide was added imidazole (1.133 g, 16.64 mmol) followed by *tert*-butyldimethylsilyl chloride (1.24 g, 8.23 mmol) and the mixture was allowed to stir at room temperature for 16 h. The solvent was removed, and the crude taken up in DCM (25 mL) and washed twice with 5% sodium bicarbonate (25 mL). The organic layer was then dried over sodium sulphate and evaporated to dryness, which was purified by silica gel column chromatography using hexanes:ethylacetate (5:1 → 1:3) to yield 0.58 g (83.2% yield) of an off-white foam; R_f (SiO$_2$): 0.55, CH$_2$Cl$_2$:MeOH 10:0.8; ^{1}H NMR (DMSO-d_6, DMSO as reference): 11.85 (s, 1H, NH1), 11.78 (s, 1H, NH), 8.19 (s, 1H, H8); 7.10-7.36 (m, 12H, Ar); 6.92-7.02 (m, 3H, Ar); 6.74-6.84 (t, 3H, Ar); 6.21 (dd, 1H, H1', J = 6.9); 4.85 (d, 2H, PhOCH$_2$CO); 4.49-4.54 (m, 1H, H3'); 3.81-3.86 (m, 1H, H4'); 3.70 (s, 6H, OCH$_3$); 3.05-3.23 (dd, 2H, H5', H5"); 2.72-2.77 (m, 1H, H2'); 2.32-2.37 (m, 1H, H2"); 0.79 (s, 9H, SiC(CH$_3$)$_3$); 0.03 (s, 3H, SiCH$_3$); -0.04 (s, 3H, SiCH$_3$); ^{13}C NMR δ -4.60, -4.50, 18.15, 25.91, 41.15, 55.42, 63.48, 67.19, 72.68, 83.83, 86.74, 87.14, 113.31, 115.10, 122.38, 123.28, 127.14, 128.09, 128.14, 130.24, 135.74, 135.78, 137.16, 144.61, 146.28, 147.85, 155.52, 156.55, 158.75, 169.67; MS (ESI), m/z (M + Na): calcd for $C_{45}H_{51}N_5O_8SiNa^+$, 840.340, found 840.20; m/z (M - H): calcd for $C_{45}H_{50}N_5O_8Si^-$, 816.343, found 816.43.

3'- *O*- alloxycarbonyl- 5'- *O*- dimethoxytrityl- *N²*- phenoxyacetyl- *O⁶*- (4-tert-butyl diphenylsiloxybutyl)-2'-deoxyguanosine (4a) : To a solution **3** (5.374 g, 6.82 mmol) in anhydrous dioxane (40 mL) was added 1-*tert*-butyldiphenylsilylbutanol (1.968 g, 9.63

mmol), triphenylphosphine (2.55 g, 9.72 mmol) followed by diisopropylazodicarboxylate (1.89 g, 9.34 mmol) introduced in a dioxane solution (4.4 mL) dropwise. The solvent was evaporated after 1h and the crude was taken up in DCM (50 mL) and washed twice with 5% sodium bicarbonate (50 mL). The organic layer was then dried over magnesium sulphate, filtered and evaporated to a solid. Flash chromatography with a short column on the crude mixture (sample:silica, 1:3, w/w) loaded in dichloromethane (DCM) and eluted with hexanes results in nearly all of the product and non-polar reaction components being eluted in the first fraction leaving the majority of triphenylphosphine oxide on the column as confirmed by thin layer chromatography (TLC). The resulting sample was purified by silica gel column chromatography using hexanes:ethylacetate (19:1 → 13:7). Triphenylphosphine oxide was precipitated from contaminated fractions with diethyl ether. The white solid removed by filtration and the diethyl ether portion evaporated and combined with non contaminated fractions to yield 5.16 g (78.9% yield) of a white foam; R_f (SiO$_2$): 0.67, hexanes:ethylacetate 2:3; ^{1}H NMR (DMSO-d_6, DMSO as reference): 10.52 (s, 1H, NH1), 8.36 (s, 1H, H8); 7.34-7.64 (m, 11H, Ar); 7.11-7.29 (m, 11H, Ar); 6.89-6.95 (m, 3H, Ar); 6.65-6.74 (dd, 4H, Ar); 6.40 (dd, 1H, H1', J = 7.2); 5.84-5.98 (m, 1H, allyl); 5.20-5.35 (m, 3H, allyl, H3'); 4.97-5.06 (m, 2H, CH$_2$OAr); 4.50-4.64 (m, 4H, vinylCH$_2$CO, PhOCH$_2$CO); 4.18-4.24 (m, 1H, H4'); 3.70 (t, 2H, CH$_2$OSi); 3.67 (s, 3H, OCH$_3$); 3.66 (s, 3H, OCH$_3$); 3.46 (dd, 1H, H5'); 3.23-3.27 (m, 1H, H2'); 3.15 (dd, 1H, H5"); 2.58-2.63 (m, 1H, H2"); 1.90 (q, 2H, CH$_2$); 1.67-1.76 (m, 2H, CH$_2$); 0.94 (s, 9H, SiC(CH$_3$)$_3$), ^{13}C NMR δ: 19.48, 25.68, 27.36, 29.29, 36.05, 55.64, 63.91, 65.07, 67.58, 68.07, 68.89, 79.13, 84.57, 84.68, 86.25, 113.61, 113.68, 115.20, 118.54, 119.32, 121.64, 127.25, 128.32, 128.56, 129.40, 129.56, 130.16, 130.28, 130.49, 132.15, 132.27, 132.73, 132.77, 132.80, 133.96,

134.16, 135.72, 136.15, 142.39, 145.56, 152.23, 153.08, 154.27, 158.63, 158.67, 158.71, 160.94, 168.16; MS (ESI), m/z (M + Na): calcd for $C_{63}H_{67}N_5O_{11}SiNa^+$, 1120.450, found 1120.39; m/z (M - H): calcd for $C_{63}H_{66}N_5O_{11}Si^-$, 1096.453, found 1096.41.

3' - O - alloxycarbonyl - 5' - O - dimethoxytrityl - N^2 - phenoxyacetyl - O^6-(7-tert-butyldiphenylsiloxyheptyl)-2'-deoxyguanosine (4b) : To a solution of **3** (3.48 g, 4.41 mmol) in anhydrous dioxane (20 mL) was added 1-*tert*butyldiphenylsilyloxyheptanol (1.80 g, 4.86 mmol), triphenylphosphine (1.76 g, 6.71 mmol) followed by diisopropylazodicarboxylate (1.37 g, 6.46 mmol) introduced in a dioxane solution (1.6 mL) dropwise. The syringe was rinsed with dioxane (0.9 mL, x3) and added to the reaction mixture. The solvent was evaporated after 1h and the crude was taken up in DCM (50 mL) and washed twice with 5% sodium bicarbonate (50 mL). The organic layer was then dried over magnesium sulphate, filtered and evaporated to a solid. Flash chromatography with a short column on the crude mixture (sample:silica, 1:3, *w/w*) loaded in dichloromethane (DCM) and eluted with hexanes results in nearly all of the product and non-polar reaction components being eluted in the first fraction leaving the majority of triphenylphosphine oxide on the column as confirmed by thin layer chromatography (TLC). The resulting sample was purified by silica gel column chromatography using a hexanes:ethylacetate gradient (19:1 → 13:7). Triphenylphosphine oxide was precipitated from contaminated fractions with diethylether. The white solid removed by filtration and the diethylether portion evaporated and combined with non contaminated fractions to yield 3.08 g (61.9% yield) of a white foam; R_f (SiO_2): 0.67, hexanes:ethylacetate 2:3; ^{1}H NMR ($CDCl_3$, $CHCl_3$ as reference): 8.68 (s, 1H, NH1); 8.10 (s, 1H, H8); 7.66-7.71 (m, 4H, Ar); 7.16-7.46 (m, 18H, Ar); 7.01-7.10 (m, 3H, Ar); 6.79 (dd, 4H, Ar); 6.48 (dd, 1H, H1', J = 5.7); 5.89-6.04

(m, 1H, allyl); 5.30-5.46 (m, 2H, allyl); 4.47-4.52 (m, 1H, H3'); 4.78 (s, 2H, CH$_2$OAr);

4.67 (dt, 2H, vinylCH$_2$CO); 4.59 (d, 2H, PhOCH$_2$CO); 4.35-4.39 (m, 1H, H4'); 3.78 (s, 6H,

OCH$_3$); 3.67 (t, 2H, CH$_2$OSi); 3.52 (dd, 1H, H5'); 3.41 (dd, 1H, H5"); 3.18-3.23 (m, 1H,

H2'); 2.71-2.76 (m, 1H, H2"); 1.89 (q, 2H, CH$_2$); 1.33-1.62 (m, 6H, (CH$_2$)$_3$); 1.06 (s, 9H,

SiC(CH$_3$)$_3$); ^{13}C NMR δ: 19.45, 25.77, 26.14, 27.36, 28.80, 29.25, 32.93, 38.36, 55.46,

63.11, 64.08, 68.20, 69.14, 78.95, 84.30, 84.46, 86.99, 113.44, 115.17, 119.07, 119.75,

122.54, 127.22, 128.16, 128.34, 128.57, 129.40, 129.56, 130.08, 130.30, 131.41, 132.27,

132.73, 132.77, 132.80, 133.96, 134.16, 135.71, 135.75, 140.11, 144.66, 151.50, 152.66,

154.43, 157.47, 158.81, 161.55; MS (ESI), m/z (M + Na): calcd for C$_{66}$H$_{73}$N$_5$O$_{11}$SiNa$^+$,

1062.497, found 1062.39; m/z (M - H): calcd for C$_{66}$H$_{72}$N$_5$O$_{11}$Si$^-$, 1138.500, found

1138.43.

3'- *O*- alloxycarbonyl- 5'- *O*- dimethoxytrityl- *N*^{2}- phenoxyacetyl- *O*^{6}- (4-hydroxybutyl)- 2'-deoxyguanosine (5a) : To a solution of **4a** (5.15 g, 5.38 mmol) in

anhydrous THF (10 mL) was added dropwise, 1M tetrabutylammonium fluoride (645 μL,

6.45 mmol) and the solution was allowed to stir at room temperature for 30 min. On

completion the reaction was concentrated, taken up in DCM (50 mL), washed twice with

5% sodium bicarbonate (50 mL). The organic layer was then dried over sodium sulphate

and evaporated to dryness, which was purified by silica gel column chromatography using

hexanes:ethylacetate (60:40 → 20:80) to yield 3.25 g (70.3% yield) of an off-white foam;

R$_f$ (SiO$_2$): 0.50, CH$_2$Cl$_2$:MeOH 10:1; ^{1}H NMR (DMSO-d_6, DMSO as reference): 10.58 (s,

1H, NH1), 8.41 (s, 1H, H8); 7.31-7.36 (m, 4H, Ar); 7.18-7.21 (m, 7H, Ar); 6.95-7.02 (m,

3H, Ar); 6.71-6.81 (dd, 4H, Ar); 6.45 (dd, 1H, H1', *J* = 6.9); 5.89-6.04 (m, 1H, allyl); 5.27-

5.41 (m, 3H, H3', allyl); 4.87-5.15 (m, 2H, CH$_2$OAr); 4.63-4.67 (dt, 2H, vinylCH$_2$CO);

4.59 (d, 2H, PhOCH$_2$CO); 4.53 (t, 1H, OH); 4.23-4.30 (m, 1H, H4'); 3.74 (s, 3H, OCH$_3$);

3.73 (s, 3H, OCH$_3$); 3.44-3.55 (m, 3H, **CH$_2$OH**, H5'); 3.26-3.35 (m, 1H, H5''); 3.20 (dd,

1H, H2'); 2.57-2.69 (m, 1H, H2''); 1.90 (q, 2H, CH$_2$); 1.56-1.68 (m, 2H, CH$_2$); ^{13}C NMR δ:

25.88, 29.64, 36.04, 55.66, 61.08, 65.06, 67.72, 68.03, 68.89, 79.13, 84.57, 84.68, 86.24,

113.61, 113.69, 115.20, 118.54, 119.32, 121.65, 127.27, 128.33, 130.18, 130.27, 130.44,

132.70, 136.15, 142.37, 145.56, 152.24, 153.03, 154.27, 158.64, 158.67, 158.71, 161.02,

168.11; MS (ESI), m/z (M + Na): calcd for C$_{47}$H$_{49}$N$_5$O$_{11}$Na$^+$, 882.333, found 882.23; m/z

(M - H): calcd for C$_{47}$H$_{48}$N$_5$O$_{11}^-$, 858.335, found 858. 39.

3'- *O*- alloxycarbonyl- 5'- *O*- dimethoxytrityl- *N*^{2}- phenoxyacetyl- *O*^{6}- (7-hydroxyheptyl)- 2'-deoxyguanosine (5b) : To a solution of **4b** (1.352 g, 1.20 mmol) in

anhydrous THF (10 mL) was added dropwise, 1M tetrabutylammonium fluoride (135 μL,

1.35 mmol) and the solution was allowed to stir at room temperature for 30 min. On

completion the reaction was concentrated, taken up in DCM (50 mL), washed twice with

5% sodium bicarbonate (50 mL). The organic layer was then dried over sodium sulphate

and evaporated to dryness, which was purified by silica gel column chromatography using

hexanes:ethylacetate (60:40 → 20:80) to yield 0.949 g (89.6% yield) of an off-white foam;

R$_f$ (SiO$_2$): 0.50, CH$_2$Cl$_2$:MeOH 10:1; ^{1}H NMR (DMSO-d_6, DMSO as reference): 10.55 (s,

1H, NH1), 8.36 (s, 1H, H8); 7.23-7.33 (m, 4H, Ar); 7.10-7.20 (m, 7H, Ar); 6.88-6.98 (m,

3H, Ar); 6.71 (dd, 4H, Ar); 6.40 (dd, 1H, H1', *J* = 5.7); 5.85-6.00 (m, 1H, allyl); 5.31-5.39

(m, 1H, H3'); 5.22-5.35 (m, 2H, allyl); 5.00 (t, 2H, CH$_2$OAr); 4.60 (d, 2H, vinylCH$_2$CO);

4.52 (d, 2H, PhOCH$_2$CO); 4.34 (s, 1H, CH$_2$OH); 4.18-4.26 (m, 1H, H4'); 3.69 (s, 6H,

OCH$_3$); 3.42 (dd, 1H, H5'); 3.30-3.41 (m, 2H, CH$_2$OH); 3.19-3.31 (m, 1H, H2'); 3.14 (dd,

1H, H5''); 2.56-2.64 (m, 1H, H2''); 1.80 (q, 2H, **CH$_2$CH$_2$OAr**); 1.21-1.47 (m, 6H, (CH$_2$)$_3$);

^{13}C NMR δ: 25.77, 26.14, 28.80, 29.25, 32.93, 38.36, 55.46, 63.11, 64.08, 68.20, 69.14, 78.95, 84.30, 84.46, 86.99, 113.44, 115.17, 119.07, 119.75, 122.54, 127.22, 28.16, 128.34, 130.08, 130.30, 131.41, 135.71, 135.75, 140.11, 144.66, 151.50, 152.66, 154.43, 15747, 158.81, 161.55; MS (ESI), m/z (M + Na): calcd for $C_{50}H_{55}N_5O_{11}Na^+$, 924.380, found 924.37; m/z (M - H): calcd for $C_{50}H_{54}N_5O_{11}^-$, 900.382, found 900.36.

1,4-{O^6-[3'-O-alloxycarbonyl-5'-O-dimethoxytrityl-N^2-phenoxyacetyl-2'-deoxy-guanidyl]}-butane (6a) : To a solution of compound **5a** (0.245 g, 0.285 mmol) in anhydrous dioxane (3.0 mL), triphenylphosphine (0.156 g, 0.595 mmol) and compound **3** (0.223 g, 0.283 mmol) were added. Diisopropylazodicarboxylate (0.117 g, 0.580 mmol) was introduced in a solution of dioxane (0.25 mL). The syringe was rinsed with dioxane (0.25 mL, x2) and added to the reaction mixture. The solvent was evaporated after 14 h and the crude was taken up in DCM (50 mL), washed twice with 5% sodium bicarbonate (50 mL). The organic layer was then dried over sodium sulphate and evaporated to dryness, which was purified by silica gel column chromatography using hexanes/ethylacetate (70:30 → 20:80) to yield 0.247 g (53.1% yield) of a white foam; R_f (SiO$_2$): 0.55, hexanes:ethylacetate 1:4; ^{1}H NMR (CDCl$_3$, CDCl$_3$ as reference): 8.83 (s, 2H, NH); 8.04 (s, 2H, H8); 7.15-7.47 (m, 20H, Ar); 6.98-7.15 (m, 6H, Ar); 6.74-6.87 (m, 8H, Ar); 6.49 (dd, 2H, H1', J = 5.7); 5.91-6.06 (m, 2H, allyl); 5.48-5.56 (m, 2H, H3'); 5.31-5.47 (m, 4H, allyl); 4.60-4.85 (m, 12H, PhOCH$_2$CO, vinylCH$_2$CO, CH$_2$OAr); 4.34-4.42 (m, 2H, H4'); 3.78 (s, 12H, OCH$_3$); 3.54 (dd, 2H, H5'); 3.44 (dd, 2H, H5"); 2.96-3.08 (m, 2H, H2'); 2.70-2.79 (m, 2H, H2"); 2.17 (s, 4H, **CH$_2$CH$_2$OAr**); MS (ESI), m/z (M + Na): calcd for $C_{90}H_{88}N_{10}O_{20}Na^+$, 1651.607, found 1651.53; m/z (M - H): calcd for $C_{90}H_{87}N_{10}O_{20}$, 1627.610, found 1627.34.

1,7-{O^6-[3'-O-alloxycarbonyl-5'-O-dimethoxytrityl-N^2-phenoxyacetyl-2'-deoxy-guanidyl]}-heptane (6b) : To a solution of compound **5b** (0.279 g, 0.309 mmol) in anhydrous dioxane (3.0 mL), triphenylphosphine (0.164 g, 0.624 mmol) and compound **3** (0.246 g, 0.313 mmol) were added. Diisopropylazodicarboxylate (0.124 g, 0.614 mmol) was introduced in a solution of dioxane (0.25 mL). The syringe was rinsed with dioxane (0.25 mL, x2) and added to the reaction mixture. The solvent was evaporated after 14 h and the crude was taken up in DCM (50 mL), washed twice with 5% sodium bicarbonate (50 mL). The organic layer was then dried over sodium sulphate and evaporated to dryness, which was purified by silica gel column chromatography using hexanes/ethylacetate (70:30 → 20:80) to yield 0.343 g (66.3% yield) of an off-white foam; R_f (SiO$_2$): 0.50, hexanes:ethylacetate 1:4; ^{1}H NMR (CDCl$_3$, CDCl$_3$ as reference): 8.67 (s, 2H, NH); 7.98 (s, 2H, H8); 7.13-7.40 (m, 22H, Ar); 6.97-7.08 (m, 6H, Ar); 6.73-6.80 (m, 8H, Ar); 6.45 (dd, 2H, H1', J = 6.0); 5.88-6.03 (m, 2H, allyl); 5.44-5.53 (m, 2H, H3'); 5.28-5.44 (m, 4H, allyl); 4.60-4.85 (s, 4H, CH$_2$OAr); 4.64-4.67 (dt, 4H, vinylCH$_2$CO); 4.58 (d, 4H, PhOCH$_2$CO); 4.30-4.38 (m, 2H, H4'); 3.76 (s, 12H, OCH$_3$); 3.50 (dd, 2H, H5'); 3.39 (dd, 2H, H5"); 2.95-3.06 (m, 2H, H2'); 2.67-2.75 (m, 2H, H2"); 1.91 (s, 4H, **CH$_2$CH$_2$OAr**); 1.51 (s, 6H, (CH$_2$)$_3$); MS (ESI), m/z (M + Na): calcd for C$_{93}$H$_{94}$N$_{10}$O$_{20}$Na$^+$, 1693.654, found 1693.3; m/z (M - H): calcd for C$_{93}$H$_{93}$N$_{10}$O$_{20}$, 1627.657, found .

1,4-{O^6-[5'-O-dimethoxytrityl-N^2-phenoxyacetyl-2'-deoxyguanidyl]}-butane (7a) : To a solution of compound **6a** (0.239 g, 0.147 mmol) in anhydrous THF (~3 mL) was added triphenylphosphine (14.6 mg, 0.056 mmol) and palladium (0) tetrakistriphenylphosphine (30.2 mg, 0.026 mmol). A stock solution (1:1) of butylamine (731.4 mg, 10.0 mmol) and formic acid (461.4 mg, 10.0mmol) was prepared. The butylamine/formic acid solution

(84.5 mg, 0.709 mmol) was suspended in THF (~1 mL) and added via syringe to a vigorously stirred solution which was allowed to stir at room temperature for 40 min. On completion the reaction was concentrated, taken up in DCM (25 mL), washed twice with 5% sodium bicarbonate (20 mL). The organic layer was then dried over magnesium sulphate and evaporated to dryness and purified by silica gel column chromatography using hexanes:ethylacetate:ethanol (20:80:0 → 20:80:6) to yield 0.182 g (84.5% yield) of an off-white solid. ^{1}H NMR (CDCl$_3$, CDCl$_3$ as reference): 8.84 (s, 2H, NH); 8.05 (s, 2H, H8); 7.15-7.47 (m, 22H, Ar); 6.96-7.14 (m, 6H, Ar); 6.77-6.86 (m, 8H, Ar); 6.67 (dd, 2H, H1', J = 6.3); 4.79-4.87 (m, 2H, H3'); 4.63-4.76 (m, 6H, CH$_2$OAr, 3OH') 4.63 (d, 4H, PhOCH$_2$CO); 4.35-4.44 (m, 2H, H4'); 3.77 (s, 12H, OCH$_3$); 3.32-3.70 (dd, 4H, H5', H5"); 2.49-2.80 (m, 4H, H2', H2"); 2.18 (s, 4H, (CH$_2$)$_2$); MS (ESI), m/z (M + Na): calcd for C$_{82}$H$_{80}$N$_{10}$O$_{16}$Na$^+$, 1483.565, found 1483.33; m/z (M - H): calcd for C$_{82}$H$_{79}$N$_{10}$O$_{16}$, 1459.568, found 1459.49.

1,7-{O^6-[5'-O-dimethoxytrityl-N^2-phenoxyacetyl-2'-deoxyguanidyl]}-heptane (7b) : To a solution of compound **6b** (0.328 g, 0.196 mmol) in anhydrous THF (~3 mL) was added triphenylphosphine (20.0 mg, 0.076 mmol) and palladium (0) tetrakistriphenylphosphine (47.1 mg, 0.041 mmol). A stock solution (1:1) of butylamine (731.4 mg, 10.0mmol) and formic acid (461.4 mg, 10.0 mmol) was prepared. The butylamine/formic acid solution (100.5mg, 0.843mmol) was suspended in THF (~1 mL) and added via syringe to a vigorously stirred solution which was allowed to stir at room temperature for 40 min. On completion the reaction was concentrated, taken up in DCM (25 mL), washed twice with 5% sodium bicarbonate (20mL). The organic layer was then dried over magnesium sulphate and evaporated to dryness and purified by silica gel column chromatography using

hexanes:ethylacetate:ethanol (20:80:0 → 20:80:6) to yield 0.130 g (77.3% yield) of an off-white solid; R_f (SiO_2): 0.70, ethylacetate; The organic phase from work-up (CH_2Cl_2/3% $NaHCO_3$ (aq.)) was dried over $MgSO_4$ followed by flash chromatography (hexanes/ethylacetate/ethanol, 1:4:0 → 4:16:1). ^1H NMR ($CDCl_3$, $CDCl_3$ as reference): 8.88 (s, 2H, NH); 8.06 (s, 2H, H8); 7.15-7.48 (m, 22H, Ar); 6.96-7.14 (m, 6H, Ar); 6.77-6.86 (m, 8H, Ar); 6.69 (dd, 2H, H1', J = 6.3); 4.76-4.94 (m, 2H, H3'); 4.59-4.76 (m, 6H, CH_2OAr, 3OH') 4.59 (d, 4H, $PhOCH_2CO$); 4.25-4.34 (m, 2H, H4'); 3.77 (s, 12H, OCH_3); 3.33-3.74 (dd, 4H, H5', H5"); 2.50-2.84 (m, 4H, H2', H2"); 1.92 (s, 4H, **CH_2CH_2OAr**); 1.46-1.58; (m, 6H, $(CH_2)_3$); MS (ESI), m/z (M + Na): calcd for $C_{85}H_{86}N_{10}O_{16}Na^+$, 1525.612, found 1525.48; m/z (M - H): calcd for $C_{85}H_{85}N_{10}O_{16}^+$, 1501.615, found 1525.53

1,7-{O^6-[3'-O-([β-2-cyanoethyl-N,N'-diisopropyl]phosphoramidite)-5'-O-dimethoxytrityl-N^2-phenoxyacetyl-2'-deoxyguanidyl]}-heptane (8b) : Compound **7b** (0.100 g, 0.066 mmol) was dissolved in DCM (1.0 mL) and diisopropylethylamine (34.4 mg, 0.201 mmol) followed by β-2-cyanoethyl-N,N,N',N'-tetraisopropylphosphor-diamidite (59.8 μL, 0.198 mmol) were added. After 4 h the reaction was found to be 50% complete and additional diisopropylethylamine (24 mg, 0.140 mmol) followed by β-2-cyanoethyl-N,N,N',N'-tetraisopropylphosphordiamidite (39.4 mg, 0.131 mmol) were added. The reaction was quenched after 4h by the addition of DCM (25 mL) and the solution was extracted with sodium bicarbonate (5 %, 3 x 50 mL). The organic layer was dried over sodium sulphate and evaporated to afford the crude product which was precipitated from hexanes to yield 0.065 g (52.0% yield) of product as an off-white foam; R_f (SiO_2): 0.47, hexanes:ethylacetate:ethanol (10:10:1); ^1H NMR (acetone-d_6, acetone as reference): 9.29 (s, 1H, NH); 8.20 (s, 2H, H8); 7.12-7.46 (m, 22H, Ar); 6.94-7.06 (m, 6H, Ar); 6.69-6.83 (m,

8H, Ar); 6.49 (dd, 2H, H1'', J = 6.6); 5.01-5.09 (m, 4H, CH$_2$OAr); 4.94-5.03 (m, 2H, H3')

4.54-4.61 (m, 4H, PhOCH$_2$CO); 4.23-4.31 (m, 2H, H4'); 3.74 (s, 12H, OCH$_3$); 3.46-3.68

(dd, 1H, H5', H5''); 3.12-3.25 (m, 2H, H2'); 2.60-2.86 (m, 2H, H2'',); 2.72 (t, 2H,

NCH(CH$_3$)$_2$); 1.89 (s, 4H, **CH$_2$CH$_2$OAr**); 1.45-1.58 (m, 6H, (CH$_2$)$_3$); 1.08-1.28 (m, 12H,

NCH(CH$_3$)$_2$); ^{31}P NMR (121.3 MHz, acetone-d_6, ppm) δ: 154.75, 154.97; MS (ESI), m/z

(M + Na): calcd for C$_{85}$H$_{86}$N$_{10}$O$_{16}$Na$^+$, 1925.828, found 1926.5.

1-{O^6-[3'-O-alloxycarbonyl-5'-O-dimethoxytrityl-N^2-phenoxyacetyl-2'-deoxyguanidyl]}

-4-{O^6-[-3'-O-tertbutyldimethylsilyl-5'-O-dimethoxytrityl-N^2-phenoxyacetyl-2'-

deoxyguanidyl]}-butane (10a) : To a solution of compound **5a** (1.699 g, 1.976 mmol) in

anhydrous dioxane (15 mL), triphenylphosphine (0.810 g, 3.087 mmol) and compound **9**

(1.575 g, 1.926 mmol) were added. Diisopropylazodicarboxylate (0.642 g, 3.018 mmol)

was introduced in a solution of dioxane (5 mL). The syringe was rinsed with dioxane (1

mL, x2) and added to the reaction mixture. The solvent was evaporated after 14 h and the

crude was taken up in DCM (50 mL) and washed twice with 5% sodium bicarbonate (50

mL). The organic layer was then dried over magnesium sulphate and evaporated to dryness,

which was purified by silica gel column chromatography using hexanes/ethylacetate (1:0 →

1:5) to yield 2.07 g (63.1% yield) of a white foam; R$_f$ (SiO$_2$): 0.45, CH$_2$Cl$_2$:MeOH 20:1; ^{1}H

NMR (CDCl$_3$, CH$_2$Cl$_2$ as reference): 8.82 (s, 1H, NH); 8.78 (s, 1H, NH); 8.06 (s, 1H, H8);

8.00 (s, 1H, H8); 6.67-7.45 (m, 22H, Ar); 6.94-6.09 (m, 6H, Ar); 6.64-6.80 (m, 8H, Ar);

6.42 (dd, 2H, H1', H1'', J = 5.7, J = 6.9); 5.86-6.02 (m, 1H, allyl); 5.43-5.51 (m, 1H, H3');

5.27-5.44 (m, 2H, allyl); 4.78 (s, 4H, CH$_2$OAr); 4.71 (d, 4H, PhOCH$_2$CO); 4.57-4.69 (m,

2H, vinylCH$_2$CO); 4.52-4.64 (m, 1H, H3'); 4.31-4.39 (m, 1H, H4'); 4.03-4.11 (m, 1H,

H4'); 3.75 (s, 12H, OCH$_3$); 3.30-3.56 (dd, 4H, H5', H5', H5'', H5''); 3.00 (m, 1H, H2');

2.62-2.76 (m, 2H, H2', H2"); 2.38-2.49 (m, 1H, H2"); 2.13 (s, 4H, $(CH_2)_2$); 0.85 (s, 9H,

$SiC(CH_3)_3$); 0.04 (s, 3H, $SiCH_3$); 0.00 (s, 3H, $SiCH_3$); ^{13}C NMR δ -4.69, -4.55, 18.08,

25.50, 25.87, 38.16, 41.01, 53.57, 55.31, 63.61, 63.97, 67.22, 67.26, 68.10, 68.19, 68.98,

69.14, 72.63, 77.35, 78.81, 84.20, 84.33, 84.40, 86.58, 86.83, 87.06, 113.25, 113.29,

115.02, 118.89, 119.00, 119.47, 119.56, 119.72, 122.28, 126.98, 127.07, 127.95, 128.01,

128.19, 128.23, 128.50, 128.59, 129.90, 130.15, 131.28, 132.03, 132.07, 132.15, 132.29,

135.58, 135.61, 135.83, 135.87, 140.10, 140.32, 144.54, 144.68, 151.30, 151.38, 152.53,

152.59, 154.28, 157.37, 157.42, 158.61, 158.66, 161.22, 161.27, 166.57; MS (ESI), m/z (M

+ Na): calcd for $C_{92}H_{98}N_{10}O_{18}SiNa^+$, 1681.672, found 1681.47; m/z (M - H): calcd for

$C_{92}H_{97}N_{10}O_{18}Si^-$, 1657.676, found 1657.60.

1-{O^6-[3'-O-alloxycarbonyl-5'-O-dimethoxytrityl-N^2-phenoxyacetyl-2'-deoxyguanidyl]}

-7-{O^6-[3'-O-tert-butyldimethylsilyl-5'-O-dimethoxytrityl-N^2-phenoxyacetyl-2'-

deoxyguanidyl]}-heptane (10b) : To a solution of **5b** (1.73 g, 1.92 mmol) in anhydrous

dioxane (10.0 mL), triphenylphosphine (0.765 g, 2.91 mmol) and compound **9** (1.54 g, 1.88

mmol) were added. Diisopropylazodicarboxylate (0.592 g, 2.78 mmol) was introduced

dropwise in a solution of dioxane (3 mL). The syringe was rinsed with dioxane (1 mL, x2)

and added to the reaction mixture. The solvent was evaporated after 30 min and the crude

was taken up in DCM (50 mL), washed twice with 5% sodium bicarbonate (50 mL). The

organic layer was then dried over sodium sulphate and evaporated to dryness, which was

purified by silica gel column chromatography using hexanes/ethylacetate (1:0 → 1:4) to

yield 2.32 g (71.1% yield) of an off-white foam; R_f (SiO$_2$): 0.70, CH$_2$Cl$_2$:MeOH 10:1; ^{1}H

NMR (DMSO-d_6, DMSO as reference): 10.58 (s, 1H, NH); 10.53 (s, 1H, NH); 8.39 (s, 1H,

H8); 8.35 (s, 1H, H8); 7.10-7.34 (m, 22H, Ar); 6.88-6.97 (m, 6H, Ar); 6.64-6.80 (m, 8H,

Ar); 6.40 (dd, 1H, H1', J = 6.0); 6.34 (dd, 1H, H1', J = 6.6); 5.85-6.00 (m, 1H, allyl); 5.31-

5.39 (m, 1H, H3'); 5.21-5.35 (m, 2H, allyl); 4.97-5.03 (m, 4H, PhOCH$_2$CO); 4.61-4.72 (m,

1H, H3'); 4.60 (d, 2H, vinylCH$_2$CO); 4.34 (t, 4H, CH$_2$OAr); 4.17-4.26 (m, 1H, H4'); 3.69-

3.77 (m, 1H, H4'); 3.68 (s, 12H, OCH$_3$); 3.46 (dd, 1H, H5'); 3.10-3.32 (m, 4H, H2', H5',

H5''); 2.86-2.98 (m, 1H, H2'); 2.54-2.65 (m, 1H, H2''); 2.27-2.38 (m, 1H, H2'); 1.80 (s, 4H,

CH$_2$CH$_2$OAr); 1.41; (s, 6H, (CH$_2$)$_3$); 0.78 (s, 9H, SiC(CH$_3$)$_3$); 0.00 (s, 3H, SiCH$_3$); -0.05 (s,

3H, SiCH$_3$); ^{13}C NMR δ -4.62, -4.48, 18.16, 25.95, 26.09, 28.98, 29.33, 29.56, 29.90,

38.27, 41.11, 55.39, 63.68, 64.04, 68.04, 68.09, 68.15, 68.23, 69.07, 72.74, 78.91, 84.25,

84.39, 84.47, 86.65, 86.90, 87.15, 113.32, 113.37, 115.11, 119.04, 119.15, 119.65, 122.45,

127.06, 127.14, 128.03, 128.08, 128.27, 128.31, 128.78, 129.99, 130.21, 131.35, 135.66,

135.69 135.91, 135.95, 140.04, 140.24, 144.61, 144.74, 151.34, 151.41, 152.54, 152.61,

154.36, 157.41, 157.47, 158.69, 158.74, 161.44, 161.48, 166.42; ; MS (ESI), m/z (M + Na):

calcd for C$_{94}$H$_{104}$N$_{10}$O$_{18}$SiNa$^+$, 1723.719, found 1724.46; m/z (M - H): calcd for

C$_{94}$H$_{103}$N$_{10}$O$_{18}$Si$^-$, 1699.723, found 1699.59.

1-{O^6-[3'-O-alloxycarbonyl-5'-O-dimethoxytrityl-N^2-phenoxyacetyl-2'-deoxyguanidyl

]}-4-{O^6-[5'-O-dimethoxytrityl-N^2-phenoxyacetyl-2'-deoxyguanidyl]}-butane (11a) : To

a solution of compound **10a** (0.492 g, 0.297 mmol) in anhydrous THF (3 mL) was added

dropwise, 1M tetrabutylammonium fluoride in THF (0.327 μL, 0.327 mmol) and the

solution was allowed to stir at room temperature for 30 min. On completion the reaction

was concentrated, taken up in DCM (25 mL), washed twice with 5% sodium bicarbonate

(25 mL). The organic layer was then dried over sodium sulphate and evaporated to dryness,

which was purified by silica gel column chromatography using hexanes:ethylacetate (1:0 →

1:5) to yield 0.344 g (75.0% yield) of an off-white solid; R$_f$ (SiO$_2$): 0.39, ethylacetate; ^{1}H

NMR (CDCl₃, CH₂Cl₂ as reference): 8.86 (s, 1H, NH); 8.80 (s, 1H, NH); 8.36 (s, 1H, H8); 8.33 (s, 1H, H8); 7.12-7.43 (m, 22H, Ar); 6.96-7.08 (m, 6H, Ar); 6.74-6.82 (m, 8H, Ar); 6.60 (dd, 1H, H1', J = 6.6); 6.46 (dd, 1H ,H1", J = 6.0); 5.88-6.02 (m, 1H, allyl); 5.43-5.52 (m, 1H, H3'); 5.28-5.44 (m, 2H, allyl); 4.54-4.70 (m, 12H, H3', OH3', CH₂OAr, vinylCH₂CO, PhOCH₂CO); 4.31-4.39 (m, 1H, H4'); 4.17-4.26 (m, 1H, H4'); 3.77 (s, 12H, OCH₃); 3.32-3.56 (dd, 4H, H5', H5"); 2.94-3.06 (m, 1H, H2'); 2.67-2.79 (m, 2H, H2', H2"); 2.56-2.67 (m, 1H, H2"); 2.15 (s, 4H, **CH₂CH₂OAr**); ^{13}C NMR δ 14.32, 25.12, 38.19, 40.75, 55.31, 60.52, 63.97, 64.24, 67.26, 68.06, 68.99, 72.65, 78.81, 84.19, 84.28, 84..32, 86.52, 86.60, 86.83, 113.25, 113.30, 115.03, 118.88, 119.02, 119.59, 122.32, 122.40, 126.99, 127.07, 127.97, 128.01, 128.19, 128.23, 129.79, 129.92, 130.15, 131.27, 135.58, 135.61, 135.82, 135.89, 140.09, 140.49, 144.53, 144.71, 151.11, 151.35, 152.48, 152.57, 154.29, 157.26, 157.33, 158.61, 158.66, 161.19, 161.25; MS (ESI), m/z (M + Na): calcd for C₈₆H₈₄N₁₀O₁₈Na⁺, 1567.586, found 1567.37; m/z (M - H): calcd for C₈₆H₈₃N₁₀O₁₈⁻, 1543.589, found 1543.52.

1-{O^6-[3'-O-alloxycarbonyl-5'-O-dimethoxytrityl-N^2-phenoxyacetyl-2'-deoxyguanidyl]}-7-{O^6-[5'-O-dimethoxytrityl-N^2-phenoxyacetyl-2'-deoxyguanidyl]}-heptane (11b) : To a solution of compound **10b** (2.30 g, 1.35 mmol) in anhydrous THF (5 mL) was added dropwise, 1M tetrabutylammonium fluoride in THF (1.7 mL, 1.7 mmol) and the solution was allowed to stir at room temperature for 30 min. On completion the reaction was concentrated, taken up in DCM (25 mL), washed twice with 5% sodium bicarbonate (25 mL). The organic layer was then dried over sodium sulphate and evaporated to dryness and purified by silica gel column chromatography using hexanes:ethylacetate (1:0 → 1:5) to yield 1.86 g (86.6% yield) of an off-white solid; R_f

(SiO$_2$): 0.30, CH$_2$Cl$_2$:MeOH (20:1); ^{1}H NMR (DMSO-d_6, DMSO as reference): 10.59 (s, 1H, NH); 10.56 (s, 1H, NH); 8.38 (s, 1H, H8); 8.35 (s, 1H, H8); 7.25-7.37 (m, 8H, Ar); 7.11-7.24 (m, 14H, Ar); 6.90-6.99 (m, 6H, Ar); 6.67-6.82 (m, 8H, Ar); 6.36-6.47 (dd, 2H, H1', H1', J = 6.9, J = 6.6); 5.88-6.03 (m, 1H, allyl); 5.24-5.42 (m, 4H, H3', H3', allyl); 5.00-5.11 (m, 4H, CH$_2$OAr); 4.63 (d, 2H, vinylCH$_2$CO); 4.46-4.60 (m, 5H, OH3', PhOCH$_2$CO); 4.19-4.28 (m, 1H, H4'); 3.97-4.05 (m, 1H, H4'); 3.71 (s, 12H, OCH$_3$); 3.49 (dd, 1H, H5'); 3.08-3.42 (m, 4H, H2', H5', H5", H5"); 2.84-2.96 (m, 1H, H2'); 2.55-2.66 (m, 1H, H2"); 2.33-2.45 (m, 1H, H2"); 1.83 (s, 4H, **CH$_2$CH$_2$OAr**); 1.45; (s, 6H, (CH$_2$)$_3$); ^{13}C NMR δ: 26.12, 29.03, 29.23, 36.05, 55.65, 65.07, 67.71, 68.03, 68.89, 71.30, 79.12, 84.26, 84.57, 84.71, 86.04, 86.24, 87.02, 113.69, 115.17, 118.41, 118.53, 119.32, 121.61, 127.24, 128.33, 130.16, 130.31, 130.42, 132.70, 136.12, 136.15, 136.26, 136.29, 141.99, 142.37, 145.56, 145.67, 152.24, 152.35, 153.04, 153.13, 154.28, 158.63, 158.68, 158.71, 160.97, 161.00, 168.17 ; MS (ESI), m/z (M + Na): calcd for C$_{89}$H$_{90}$N$_{10}$O$_{18}$Na$^+$, 1609.633, found 1609.46; m/z (M - H): calcd for C$_{89}$H$_{89}$N$_{10}$O$_{18}$$^-$, 1586.716, found 1585.52.

1-{O^6-[3'-O-alloxycarbonyl-5'-O-dimethoxytrityl-N^2-phenoxyacetyl-2'-deoxyguanidyl]}-4-{O^6-[5'-O-dimethoxytrityl-N^2-phenoxyacetyl-2'-deoxyguanidyl-3'-O-(β-2-cyanoethyl-N,N'-diisopropyl)phosphoramidite]}-butane (12a) : Compound 11a (0.600 g, 0.388 mmol) was dissolved in THF (5.0 mL) and diisopropylethylamine (120 μL, 0.582 mmol) followed by N,N'-diisopropylamino cyanoethyl phosphoramidic chloride (130 μL, 0.330 mmol). After 4 h the reaction was found to be 50% complete and additional diisopropylethylamine (60 mg, 0.352 mmol) followed by N,N'-diisopropylamino cyanoethyl phosphoramidic chloride (110 μL, 0.330 mmol) were added. The reaction was quenched after 4 h by the addition of DCM (25 mL) and the solution was extracted with

108

sodium bicarbonate (5 %, 3 x 50 mL). The organic layer was dried over sodium sulphate and evaporated to afford the crude product which was precipitated from hexanes to yield 0.661 g (88.0% yield) of product as an off-white foam. This material was chromatographed using silica gel with hexanes:ethylacetate and 1% TEA to give 0.130 g (19.2% yield) of a white foam; R_f (SiO$_2$): 0.47, hexanes:ethylacetate (1:1); ^{1}H NMR (acetone-d_6, acetone as reference): 9.43 (s, 1H, NH); 9.40 (s, 1H, NH); 8.25 (s, 2H, H8); 7.09-7.51 (m, 22H, Ar); 6.96-7.04 (m, 6H, Ar); 6.70-6.84 (m, 8H, Ar); 6.53 (dd, 2H, H1',H1', J = 6.3); 5.94-6.09 (m, 1H, allyl); 5.52-5.61 (m, 1H, H3'); 5.26-5.46 (m, 2H, allyl) 5.04-5.15 (m, 4H, CH$_2$OAr); 4.73-4.78 (m, 4H, PhOCH$_2$CO); 4.68 (d, 2H, vinylCH$_2$CO); 4.35-4.44 (m, 1H, H4'); 3.30-3.38 (m, 1H, H4'); 3.77 (s, 12H, OCH$_3$); 3.33-3.74 (dd, 4H, H5', H5', H5", H5"); 3.15-3.27 (m, 1H, H2'); 2.60-2.80 (m, 5H, H2', H2", H2", NCH(CH$_3$)$_2$); 2.16 (s, 4H, CH$_2$CH$_2$OAr); 1.13-1.24; (m, 12H, NCH(CH$_3$)$_2$); ^{31}P NMR (121.3 MHz, acetone-d_6, ppm): 149.2, 149.5; MS (ESI), m/z (M + Na): calcd for C$_{95}$H$_{101}$N$_{12}$O$_{19}$PNa$^+$, 1767.694, found 1767.37; m/z (M - H): calcd for C$_{95}$H$_{100}$N$_{12}$O$_{19}$P$^-$, 1743.697, found 1743.33.

1-{O^6-[3'-O-alloxycarbonyl-5'-O-dimethoxytrityl-N^2-phenoxyacetyl-2'-deoxyguanidyl]}-7-{O^6-[5'-O-dimethoxytrityl-N^2-phenoxyacetyl-2'-deoxyguanidyl-3'-O-(β-2-cyanoethyl-N,N'-diisopropyl)phosphoramidite]}-heptane (12b) : Compound **11b** (0.618 g, 0.389 mmol) was dissolved in THF (5mL) and diisopropylethylamine (120 μL, 0.660 mmol) followed by N,N'-diisopropylamino cyanoethyl phosphoramidic chloride (130 μL, 0.582 mmol) were added forming a precipitate after 5 minutes. The atmosphere was then exchanged with argon. The reaction was rotary evaporated after 1h then solvated in DCM (25 mL) and extracted with sodium bicarbonate (5 %, 2 x 25 mL). The organic layer was dried over sodium sulphate and evaporated to afford the crude product which was

precipitated from hexanes and was filtered to yield 0.657g (97.3 % yield) of an off-white

powder. This material was chromatographed using silica gel with hexanes:ethylacetate and

1% TEA to give 0.168 g (24.1% yield) of a white foam; R_f (SiO$_2$): 0.27, 0.41

hexanes:ethylacetate:TEA (5:1:1%); ^{1}H NMR (acetone-d_6, acetone as reference): 9.38 (s,

1H, NH); 9.34 (s, 1H, NH); 8.24 (s, 1H, H8); 8.23 (s, 1H, H8); 7.15-7.48 (m, 22H, Ar);

6.96-7.09 (m, 6H, Ar); 6.70-6.85 (m, 8H, Ar); 6.53 (dd, 2H, H1',H1', J = 6.6); 5.92-6.09

(m, 1H, allyl); 5.49-5.64 (m, 1H, H3', H3', allyl); 5.26-5.45 (m, 2H, allyl); 4.96-5.17 (m,

4H, CH$_2$OAr); 4.65-4.77 (d, 2H, vinylCH$_2$CO); 4.53-4.64 (m, 5H, H3', PhOCH$_2$CO); 4.18-

4.27 (m, 1H, H4'); 3.28-3.35 (m, 1H, H4'); 3.77 (s, 12H, OCH$_3$); 3.34-3.74 (m, 4H, H2',

H5', H5", O(CH$_2$)$_2$CN); 3.14-3.26 (m, 1H, H2'); 2.94 (t, 2H, NCH(CH$_3$)$_2$); 2.61-2.82 (m,

2H, H2", H2"); 1.93 (s, 4H, **CH$_2$CH$_2$OAr**); 1.56; (s, 6H, (CH$_2$)$_3$); 1.13-1.24 (m, 12H,

NCH(**CH$_3$**)$_2$); ^{31}P NMR (121.3 MHz, acetone-d_6, ppm): 149.2, 149.5; MS (ESI), m/z (M +

Na): calcd for C$_{98}$H$_{107}$N$_{12}$O$_{19}$PNa$^+$, 1809.741, found 1809.47; m/z (M - H): calcd for

C$_{98}$H$_{106}$N$_{12}$O$_{19}$P$^-$, 1785.744, found 1785.33.

1-{O^6-[3'-O-$tert$-Butyldimethylsilyl-5'-O-dimethoxytrityl-N^2-phenoxyacetyl-2'-

deoxyguanidyl]}-7-{O^6-[5'-O-dimethoxytrityl-N^2-phenoxyacetyl-2'-deoxyguanidyl]}-

heptane (13) : To a solution of compound **10b** (0.522 g, 0.307 mmol) in anhydrous THF (3

mL) was added triphenylphosphine (24.4 mg, 0.093 mmol) and palladium (0)

tetrakistriphenylphosphine (35.2 mg, 0.030 mmol). A stock solution (1:1) of butylamine

(731.4 mg, 10.0 mmol) and formic acid (461.4 mg, 10.0 mmol) was prepared. The

butylamine/formic acid solution (78.7 mg, 0.66 mmol) was suspended in THF (~1 mL) and

added via syringe to a vigorously stirred solution which was allowed to stir at room

temperature for 40 min. On completion the reaction was concentrated, taken up in DCM

(25 mL), washed twice with 5% sodium bicarbonate (25 mL). The organic layer was then

dried over sodium sulphate and evaporated to dryness and purified by silica gel column

chromatography using hexanes:ethylacetate (1:0 → 1:4) to yield 0.153 g (76.0% yield) of

an off-white solid; R_f (SiO$_2$): 0.33 hexanes:ethylacetate (1:4); ^{1}H NMR (CHCl$_3$, CH$_2$Cl$_2$ as

reference): 8.76 (s, 1H, NH); 8.67 (s, 1H, NH); 8.03 (s, 1H, H8); 8.01 (s, 1H, H8); 7.14-

7.44 (m, 22H, Ar); 6.97-7.09 (m, 6H, Ar); 6.74-6.84 (m, 8H, Ar); 6.60 (dd, 1H, H1', J =

6.6); 6.43 (dd, 1H, H1', J = 6.4); 4.40-4.85 (m, 10H, H3', H3', PhOCH$_2$CO, CH$_2$OAr);

4.17-4.26 (m, 1H, H4'); 4.05-4.14 (m, 1H, H4'); 3.77 (s, 6H, OCH$_3$); 3.75 (s, 6H, OCH$_3$);

3.45 (dd, 1H, H5'); 3.00-3.38 (m, 3H, H5", H5', H5"); 3.21 (s, 1H, 3'OH); 2.56-2.83 (m,

3H, H2', H2", H2'); 2.40-2.51 (m, 1H, H2"); 1.90 (s, 4H, **CH$_2$CH$_2$OAr**); 1.52; (s, 6H,

(CH$_2$)$_3$); 0.86 (s, 9H, SiC(CH$_3$)$_3$); 0.86 (s, 6H, (CH$_3$)$_2$); ^{13}C NMR δ: -4.63, -4.48, 18.15,

25.88, 25.94, 26.08, 28.97, 29.31, 40.83, 41.11, 53.63, 55.37, 55.39, 63.67, 64.32, 68.03,

68.11, 68.23, 72.73, 84.36, 84.47, 86.60, 86.65, 86.68, 87.15, 113.32, 114.78, 115.11,

119.14, 119.20, 122.43, 122.53, 127.06, 128.04, 128.31, 129.58, 129.98, 130.02, 130.21,

135.91, 135.96, 140.23, 140.42, 144.74, 144.79, 151.16, 151.34, 152.50, 152.54, 157.33,

157.49, 158.69, 161.43, 166.52; MS (ESI), m/z (M + Na): calcd for C$_{91}$H$_{100}$N$_{10}$O$_{16}$SiNa$^+$,

1639.698, found 1639.50; m/z (M - H): calcd for C$_{91}$H$_{99}$N$_{10}$O$_{16}$Si$^-$, 1615.702, found

1615.67.

1-{O^6-[3'-O-*tert*-Butyldimethylsilyl-5'-O-dimethoxytrityl-N^2-phenoxyacetyl-2'-

deoxyguanidyl]}-7-{O^6-[5'-O-dimethoxytrityl-N^2-phenoxyacetyl-2'-deoxyguanidyl-3'-

O-(β-2-cyanoethyl-N,N'-diisopropyl)phosphoramidite]}-heptane (14) : Compound **13**

(0.142 g, 0.088 mmol) was dissolved in THF (1.3mL) and diisopropylethylamine (26 μL,

0.150 mmol) followed by N,N'-diisopropylamino cyanoethyl phosphoramidic chloride (30

μL, 0.132 mmol) were added. After 1 h the reaction was found to be 50% complete and additional diisopropylethylamine (13 μL, 0.075 mmol) followed by N,N'-diisopropylamino cyanoethyl phosphoramidic chloride (15 μL, 0.066 mmol) were added whereupon a precipitate formed. The reaction was quenched after 2h by the addition of DCM (25 mL) and the solution was extracted with sodium bicarbonate (5 %, 3 x 50 mL). The organic layer was dried over sodium sulphate and evaporated to afford the crude product which was precipitated from hexanes to yield 0.138 g (86.3 % yield) of product as an off-white foam.

R_f (SiO$_2$): 0.72, 0.83 hexanes:ethylacetate (5:1); ^{1}H NMR (acetone-d_6, acetone as reference): 9.31 (s, 1H, NH); 9.29 (s, 1H, NH); 8.21 (s, 1H, H8); 8.20 (s, 1H, H8); 7.37-7.44 (m, 4H, Ar); 7.12-7.34 (m, 18H, Ar); 6.92-7.05 (m, 6H, Ar); 6.69-6.81 (m, 8H, Ar); 6.48 (dd, 1H, H1', J = 6.0); 6.44 (dd, 1H, H1', J = 6.3); 5.01-5.07 (m, 4H, ArOCH$_2$); 4.91-5.00 (m, 1H, H3'); 4.84-4.90 (m, 1H, H3'); 4.54-4.60 (m, 4H, PhOCH$_2$CO); 4.24-4.31 (m, 1H, H4'); 3.98-4.05 (m, 1H, H4'); 3.74 (s, 12H, OCH$_3$); 3.30-3.71 (m, 8H, H5', H5", OCH$_2$CH$_2$CN); 3.13-3.24 (m, 2H, H2'); 2.96-3.12 (m, 2H, NCH(CH$_3$)$_2$); 2.57-2.63 (m, 1H, H2"); 2.42-2.47 (m, 1H, H2"); 1.80-1.98 (m, 4H, **CH$_2$CH$_2$OAr**); 1.51 (s, 6H, (CH$_2$)$_3$); 1.12-1.23 (m, 12H, NCH(**CH$_3$**)$_2$); 0.85 (s, 9H, SiC(CH$_3$)$_3$); 0.07 (s, 3H, SiCH$_3$); 0.02 (s, 3H, SiCH$_3$); ^{31}P NMR (121.3 MHz, acetone-d_6, ppm): 154.76, 154.56; MS (ESI), m/z (M + Na): calcd for C$_{100}$H$_{117}$N$_{12}$O$_{17}$PSi Na$^+$, 1839.806, found 1840.47; m/z (M - H): calcd for C$_{100}$H$_{116}$N$_{12}$O$_{17}$PSi$^-$, 1815.903, found 1816.34.

5'-O-*tert*-Butyldiphenylsilyl-N^2-phenoxyacetyl-2'-deoxyguanosine (15) : To a solution of compound **1**, (2.44 g, 6.08 mmol) in anhydrous N',N'-dimethylformamide (40 mL) was added imidazole (1.68 g, 24.7 mmol) followed by addition via a syringe of *tert*-butyldiphenylsilyl chloride (2.48g, 9.03mmol) and the mixture was allowed to stir at room

temperature for 18 h. The reaction mixture taken up in ethylacetate (90 mL) and washed thrice with 5% sodium bicarbonate (90 mL) and back extraction performed twice. The organic layer was then dried over sodium sulphate and evaporated to dryness, which was dried onto celite then purified by silica gel column chromatography using CH_2Cl_2:MeOH (100:0.4 → 100:1.4) to yield 2.80g (72.0% yield) of an off-white foam; R_f (SiO_2): 0.41, CH_2Cl_2:MeOH 10:1; ^{1}H NMR ($CDCl_3$, $CHCl_3$ as reference): 11.90 (s, 1H, NH1), 9.45 (s, 1H, NH), 7.94 (s, 1H, H8); 7.58-7.67 (m, 4H, Ar); 7.27-7.45 (m, 8H, Ar); 7.08 (t, 1H, Ar); 6.99 (d, 2H, Ar); 6.31 (dd, 1H, H1', J = 6.3,); 4.67-4.79 (m, 3H, $PhOCH_2CO$, H3'); 3.97-4.04 (m, 1H, H4'); 3.80-3.90 (dd, 2H, H5', H5"); 2.70 (s, 1H, OH3'); 2.49-2.60 (m, 2H, H2', H2"); 1.03 (s, 9H, $SiC(CH_3)_3$); ^{13}C NMR δ: 19.38, 27.08, 29.90, 40.94, 53.66, 64.32, 67.10, 71.82, 83.97, 87.47, 114.98, 121.68, 122.98, 128.00, 128.03, 130.12, 132.92, 132.98, 135.66, 135.76, 137.49, 146.61, 147.96, 155.71, 156.72, 170.25; MS (ESI), m/z (M + Na): calcd for $C_{34}H_{37}N_5O_6SiNa^+$, 662.241, found 662.16; m/z (M - H): calcd for $C_{34}H_{36}N_5O_6Si^-$, 638.224, found 638.35.

3'-*O*-Levulinyl-5'-*O*-*tert*-butyldiphenylsilyl-*N*^{2}-phenoxyacetyl-2'-deoxyguanosine (16) :

To a solution of compound **15**, (2.09 g, 3.26 mmol) in anhydrous dioxane (25 mL) was added DCC (2.35 g, 11.38 mmol) and DMAP (0.20 g, 1.62 mmol). Levulinic acid (2.26 g, 19.50 mmol) was added to a vigorously stirred solution and the mixture was allowed to stir at room temperature for 3 h. The reaction mixture was vacuum filtered and the solid rinsed with dioxane (4 mL, x2). The orange coloured solution was dried and taken up in DCM (70 mL) and washed with 5% sodium bicarbonate (40 mL, x3). The organic layer was then dried over magnesium sulphate, filtered and evaporated to dryness, which was then purified by silica gel column chromatography using CH_2Cl_2:MeOH (100:0.5 → 100:1.4) to yield

1.77g (73.8% yield) of an off-white foam; R_f (SiO$_2$): 0.54, CH$_2$Cl$_2$:MeOH 16:1; ^{1}H NMR

(CDCl$_3$, CHCl$_3$ as reference): 11.90 (s, 1H, NH1), 9.19 (s, 1H, NH), 7.88 (s, 1H, H8); 7.61-

7.64 (m, 4H, Ar); 7.27-7.45 (m, 8H, Ar); 7.10 (t, 1H, Ar); 7.00 (d, 2H, Ar); 6.27 (dd, 1H,

H1', J = 5.4); 5.51-5.60 (m, 1H, H3'); 4.58-4.63 (m, 2H, PhOCH$_2$CO); 4.18-4.24 (m, 1H,

H4'); 3.79-3.91 (dd, 2H, H5', H5"); 2.55-2.85 (m, 6H, CH$_2$, H2', H2"); 2.21 (s, 3H, CH$_3$);

1.06 (s, 9H, SiC(CH$_3$)$_3$); ^{13}C NMR δ: 19.41, 27.10, 27.87, 28.09, 30.00, 38.03, 38.12,

38.43, 64.10, 67.13, 75.00, 83.72, 85.41, 115.07, 122.13, 123.22, 128.05, 128.10, 130.20,

132.61, 132.8, 135.60, 135.77, 137.06, 146.41, 147.83, 155.48, 156.56, 169.72, 172.39,

206.68; MS (ESI), m/z (M + Na): calcd for C$_{39}$H$_{43}$N$_5$O$_8$SiNa$^+$, 760.277, found 760.16; m/z

(M - H): calcd for C$_{39}$H$_{42}$N$_5$O$_8$Si$^-$, 736.281, found 736.32.

1-{O^6-[3'-O-alloxycarbonyl-5'-O-dimethoxytrityl-N^2-phenoxyacetyl-2'-deoxyguanidyl]}

-7-{O^6-[3'-O-levulinyl-5'-O-tert-butyldiphenylsilyl-N^2-phenoxyacetyl-2'-

deoxyguanidyl]}-heptane (17): To a solution of compound **5b** (1.26 g, 1.40 mmol) and

triphenylphosphine (2.33 g, 8.92 mmol) in anhydrous dioxane (5.0 mL) was added **16** (1.05

g, 1.43 mmol). DIAD (0.453 g, 2.13 mmol) was added dropwise over 1 min via a syringe.

The syringe was rinsed with dioxane (0.2 mL, x4) and the solution was allowed to stir at

room temperature for 30 min. On completion the reaction was concentrated, taken up in

DCM (25 mL) and washed with 5% sodium bicarbonate (25 mL, x3). The organic layer

was then dried over sodium sulphate and evaporated to dryness, which was purified by

silica gel column chromatography using hexanes:ethylacetate (1:0 → 1:4) to yield 1.58g

(70.4% yield) of an off-white solid; R_f (SiO$_2$): 0.50, CH$_2$Cl$_2$:MeOH, 20:1; ^{1}H NMR (CDCl$_3$,

CH$_2$Cl$_2$ as reference): 8.75 (s, 1H, NH); 8.67 (s, 1H, NH); 8.03 (s, 1H, H8); 7.98 (s, 1H,

H8); 7.59-7.67 (m, 4H, Ar); 7.13-7.45 (m, 20H, Ar); 6.97-7.09 (m, 6H, Ar); 6.72-6.81 (m,

4H, Ar); 6.60 (dd, 1H, H1', J = 6.0); 6.45 (dd, J = 6.3, 1H, H1"); 5.87-6.03 (m, 1H, allyl);

5.53-5.59 (m, 1H, H3'); 5.43-5.52 (m, 1H, H3'); 5.27-5.44 (m, 2H, allyl); 4.76 (s, 4H,

CH_2OAr); 4.65 (d, 2H, vinylCH_2CO); 4.58 (t, 4H, CH_2OPh); 4.32-4.37 (m, 1H, H4'); 4.16-

4.22 (m, 1H, H4'); 3.81-4.01 (dd, 2H, H5', H5"); 3.76 (s, 6H, OCH_3); 3.36-3.54 (dd, 2H,

H5', H5"); 2.95-3.05 (m, 1H, H2'); 2.56-2.85 (m, 7H, H2', H2", H2", $(CH_2)_2$); 2.18-2.23

(m, 3H, CH_3); 1.86-1.94 (m, 4H, **CH_2CH_2OAr**); 1.66; (s, 2H, (CH_2)); 1.42-1.58; (m, 6H,

$(CH_2)_3$); 1.06 (s, 9H, $SiC(CH_3)_3$); ^{13}C NMR δ 19.32, 25.96, 27.02, 28.06, 28.85, 29.19,

29.89, 37.98, 38.13, 38.53, 55.26, 63.92, 64.28, 67.95, 68.04, 68.93, 75.12, 78.78, 83.89,

84.15, 84.28, 85.42, 86.78, 113.25, 114.98, 118.86, 118.93, 119.51, 122.33, 127.02, 127.91,

127.96, 128.15, 129.45, 129.88, 130.01, 130.10, 131.24, 132.60, 132.83, 135.51, 135.57,

135.67, 139.64, 139.93, 144.50, 151.29, 151.33, 152.49, 152.56, 154.23, 157.30, 158.62,

161.37, 16.35, 172.29, 206.41; MS (ESI), m/z (M + Na): calcd for $C_{89}H_{96}N_{10}O_{18}SiNa^+$,

1643.657, found 1643.42; m/z (M - H): calcd for $C_{89}H_{95}N_{10}O_{18}Si^-$, 1619.660, found

1619.51.

1-{O^6-[5'-*O*-dimethoxytrityl-*N^2*-phenoxyacetyl-2'-deoxyguanidyl]}-7-{O^6-[3'-*O*-

levulinyl-5'-*O*-tert-butyldiphenylsilyl-*N^2*-phenoxyacetyl-2'-deoxyguanidyl]}-heptane

(18) : To a solution of compound **17** (0.348 g, 0.214 mmol) in anhydrous THF (2.5 mL)

was added triphenylphosphine (17.6 mg, 0.067 mmol) and palladium (0)

tetrakistriphenylphosphine (36.9 mg, 0.032 mmol). A stock solution (1:1) of butylamine

(731.4 mg, 10.0 mmol) and formic acid (461.4 mg, 10.0mmol) was prepared. The

butylamine/formic acid solution (0.113 g, 0.952 mmol) was suspended in THF (~1 mL) and

added via syringe to a vigorously stirred solution which was allowed to stir at room

temperature for 40 min. On completion the reaction was concentrated, taken up in DCM

(25 mL) and washed with 5% sodium bicarbonate (25 mL, x3). The organic layer was then dried over sodium sulphate and evaporated to dryness and purified by silica gel column chromatography using hexanes:ethylacetate (1:0 → 1:5) to yield 0.229 g (69.5% yield) of an off-white solid; R_f (SiO$_2$): 0.47, CH$_2$Cl$_2$:MeOH, 10:1; ^{1}H NMR (DMSO-d_6, DMSO as reference): 10.59 (s, 1H, NH); 10.57 (s, 1H, NH); 8.34 (s, 1H, H8); 8.33 (s, 1H, H8); 7.50-7.58 (m, 4H, Ar); 7.14-7.42 (m, 19H, Ar); 6.84-6.97 (m, 6H, Ar); 6.66-6.88 (dd, 4H, Ar, J = 9.0); 6.33-6.43 (dd, 2H, H1',H1', J = 6.3); 5.44-5.52 (m, 1H, H3'); 5.28-5.37 (d, 1H, H3'); 4.99-5.04 (m, 4H, CH$_2$OAr); 4.44-4.58 (m, 5H, PhOCH$_2$CO, 3OH'); 4.11-4.19 (m, 1H, H4'); 3.94-4.01 (m, 1H, H4'); 3.78-3.96 (dd, 2H, H5'); 3.69 (s, 6H, OCH$_3$); 3.10-3.32 (m, 3H, H2', H5', H5"); 2.81-2.90 (m, 1H, H2'); 2.76 (t, 2H, **CH$_2$**COCH$_3$); 2.52 (t, 2H, **CH$_2$**CO$_2$) 2.28-2.39 (m, 1H, H2"); 2.12 (s, 3H, COCH$_3$); 1.80 (s, 4H, **CH$_2$CH$_2$**OAr); 1.41 (s, 6H, (CH$_2$)$_3$), 0.93 (s, 9H, SiC(CH$_3$)$_3$); ^{13}C NMR δ: 19.44, 25.01, 26.11, 27.28, 28.46, 29.02, 29.20, 30.26, 36.32, 36.95, 38.17, 55.66, 64.80, 65.09, 67.63, 68.02, 71.30, 74.93, 84.28, 85.49, 86.04, 87.02, 113.69, 115.17, 118.41, 121.59, 127.21, 128.36, 128.44, 128.54, 130.14, 130.31, 130.41, 130.54, 131.85, 131.98, 133.35, 133.47, 135.69, 136.25, 136.29, 142.00, 145.66, 152.35, 152.45, 153.12, 153.20, 158.62, 158.67, 160.96, 161.01, 168.23, 172.64, 207.54; MS (ESI), m/z (M + Na): calcd for C$_{85}$H$_{92}$N$_{10}$O$_{16}$SiNa$^+$, 1559.635, found 1559.38; m/z (M - H): calcd for C$_{85}$H$_{91}$N$_{10}$O$_{16}$Si$^-$, 1535.639, found 1535.52.

1-{O^6-[3'-O-levulinyl-5'-O-tert-butyldiphenylsilyl-N^2-phenoxyacetyl-2'-deoxyguanidyl]}-7-{O^6- [(5'-O-dimethoxytrityl-N^2-phenoxyacetyl-2'-deoxyguanidyl-3'-O-($β$-2-cyanoethyl-N,N'-diisopropyl)phosphoramidite]}-heptane (19) : Compound **18** (0.220 g, 0.143 mmol) was dissolved in THF (1.5 mL) and diisopropylethylamine (43 μL, 0.243 mmol) followed by N,N'-diisopropylamino cyanoethyl phosphoramidic chloride (48

μL, 0.215 mmol) were added. After 1 h the reaction was found to be 50% complete and additional diisopropylethylamine (22 μL, 0.122 mmol) followed by *N,N'*-diisopropylamino cyanoethyl phosphoramidic chloride (24 μL, 0.107 mmol) were added whereupon a precipitate formed. The reaction was quenched after 2h by the addition of DCM (10 mL) and the solution was extracted with sodium bicarbonate (5 %, 3 x 10 mL). The organic layer was dried over sodium sulphate and evaporated and silica gel chromatography performed using hexanes:ethylacetate (1:0 → 1:4) and TEA (1%) to yield 0.145 g (58.3% yield) of an off-white foam. R_f (SiO$_2$): 0.38, 0.50, hexanes:ethylacetate:TEA (1:4:1%); ^{1}H NMR (CDCl$_3$, CHCl$_3$ as reference): 10.58 (s, 1H, NH); 10.53 (s, 1H, NH); 8.36 (s, 1H, H8); 8.33 (s, 1H, H8); 7.08-7.33 (m, 20H, Ar); 6.86-6.96 (m, 6H, Ar); 6.64-6.78 (m, 8H, Ar); 6.33-6.43 (dd, 2H, H1',H1', J = 6.3); 5.15-5.24 (m, 2H, H3', H3'); 4.97-5.03 (m, 4H, PhOCH$_2$CO); 4.46-4.55 (m, 5H, OH3', CH$_2$OAr); 4.18-4.26 (m, 1H, H4'); 3.93-4.00 (m, 1H, H4'); 3.67 (s, 12H, OCH$_3$); 3.46 (dd, 1H, H5'); 3.03-3.32 (m, 4H, H2', H5', H5"); 2.82-2.91 (m, 1H, H2'); 2.55-2.64 (m, 1H, H2"); 2.36-2.42; (m, 2H, NCH(CH$_3$)$_2$); 2.28-2.36 (m, 1H, H2"); 1.80 (s, 4H, **CH$_2$CH$_2$OAr**); 1.41; (s, 6H, (CH$_2$)$_3$) 1.14-1.25; (m, 12H, NCH(**CH$_3$**)$_2$); ^{31}P NMR δ: 154.56, 154.76 MS (ESI), m/z (M + Na): calcd for C$_{94}$H$_{109}$N$_{12}$O$_{17}$PSiNa$^+$, 1759.743, found 1759.47; m/z (M - H): calcd for C$_{94}$H$_{108}$N$_{12}$O$_{17}$PSi$^-$, 1735.747, found 1735.53.

References

1. Vogel, E., W., Nivard, M.J., Ballering, L.A., Bartsch, H., Barbin, A., Nair, J., Comendador, M.,A., Sierra, L. M., Aguirrezabalaga, I., Tosal, L., Ehrenberg, L., Fuchs, R. P., Janel-Bintz, R., Maenhaut-Michel, G., Montesano, R., Hall, J., Kang, H., Miele, M., Thornale, J., Bender, K., Engelbergs, J., Rajewsky, M. F.; *Mutat. Res.*, **1996**, *353*, 177-218.
2. Loveless, A., Ross, W.C.J.; *Nature*, **1950**, *166*, 113-114.
3. Dronkert, M.L.G., Kannar, R.; *Mutat. Res.*, **2001**, *486*, 217-247.
4. Mc High, P., Spanswick, V., Hartley, J.; *J. Lancet Oncol.*, **2001**, *2*, 483-490.
5. Noll, D.M., McGregor-Mason T., Miller, P.S.; *Chem. Rev.*, **2006**, *106*, 277-301.
6. Westerhof, G.R., Down, J.D., Blokland, I., Wood, M., Boudewijn, A., Watson, A.J., McGown, A.T., Ploemacher, R.E., Margison, G.P.; *Exp. Hematology*, **2001**, *29*, 633-638.
7. Garyfallia Christodoulopoulos, McGill, Doctoral Thesis: Drug Resistance to Nitrogen Mustards in B-Cell Chronic Lymphocytic Leukemia, **1999**, pg. 3.
8. Puten, L.M.V., Lelieveld, P.; *Eur. J. Cancer*, **1971**, *5*, 11-20.
9. Vogel, E.W., Barbin, A., Nivard, M.J., Stack, H.F., Waters, M.D. Lohman, P.H.; *Mutat. Res.*, **1998**, *400*, 509-540.
10. Scott, B.R., Pathak, M.A., Mohn, G.R.; *Mutat. Res.*, **1976**, *39*, 29–74.
11. Basu, A.K., Marnett, L.J., Romano, L.J.; *Mutat. Res.*, **1984**, *129*, 39-46.
12. Chaudhary, A.K., Nokubo, M., Reddy, G.R., Yeola, S.N., Morrow, J.D., Blair, I.A., Marnett, L.J.; *Science*, **1994**, *265*, 1580-1582.
13. (a) Ali-Osman, F., Rairkar, A., Young, P.; *Cancer Biochem. Biophys.* **1995**, *14*, 231–241, (b) O'Connor, P.M., Kohn, K.W.; *Cancer Commun.*, **1990**, *2*, 387–394, (c) Kano, Y., Fujiwara, Y.; *Mutat. Res.*, **1981**, *81*, 365–375.
14. Noll, D.M., Noronha, A.M., Miller, P.S.; *J. Am. Chem. Soc.*, **2001**, *123*, 3405–3411.
15. a) Teicher, B.A., Cucchi, C.A., Lee, J.B., Flatow, J.L., Rosowsky, A., Frei, E., III.; *Cancer Res.*, **1986**, *46*, 4379-4383. b) Frei, E., III, Teicher, B.A., Holden, S.A., Cathcart, K.N.S., Wang, Y.; *Cancer Res.*, **1988**, *48*, 6417-6423. c) Teicher, B.A., Frei, E., III.; *Cancer Chemother. Pharmacol.*, **1988**, 21, 292-298.
16. Spielmann, H.P., Dwyer, T.J. Hearst, J.E. Wemmer, D.E.; *Biochemistry* **1995**, *34*, 12937- 12953.
17. Chaplin, RE. Chronic myelogenous leukemia. R.P. Gale (ed.). In: Leukemia Therapy, pp. 147-165. Boston: Blackwell Scientific Publications, 1986.
18. Pacheco, D.Y., Stratton, N.K., Gibson. N.W.; *Cancer Res.*, **1989**, *49*, 5108-5110.
19. Armstrong, D.K., Gordon, G.B., Hilton, J., Streeper, R.T., Colvin, O.M., Davidson, N.E.; *Cancer Res.*, **1992**, *52*, 1416-1421.
20. Waud, W.R. Plowman, J., Paul, K.D., Narayanan, V.L., Bailey, D.M., Harrison, S.D., Dykes, D.J., Laster, W.R., Griswold, D.P.; *Investigational New Drugs*, **1991**, *9*, 149-157.
21. Lukin, M., de la Santos, C.; *Chem. Rev.*, **2006**, *106*, 607-686.
22. Jeltsch, A.; *ChemBioChem.*, **2002**, 3, 274-293.
23. Bae, S.-H., Cheong, H.-K., Cheong, C., Kang, S., Hwang, D.S., Choi, B.-S.; *J. Biol. Chem.* **2003**, *278*, 45987-45993.

24. Streeper, R.T., Cotter, R.J., Colvin, M.E., Hilton, J., Colvin, M.O.; *Cancer Res.*, **1995**, *55*, 1491-1498.

25. Price, C.C., Gaucher, G.M., Koneru, P., Shibakawa, R., Sowa, J.R., Yamaguchi, M.; *Biochim. Biophys. Acta.*, **1968**, *166*, 327-359.

26. Rajski, S.R., Williams, R.M.; *Chem. Rev.*, **1998**, *98*, 2723-2795.

27. Margison, G. P., Santibanez-Koref, M. F.; *BioEssays* **2002**, *24*, 255-266.

28. Lawley, P.D., *Chemical Carcinogens*, 2nd Ed., Searle, C.E. Ed., American Chemical Society: Washington, DC, **1984**.

29. Bignami, M., O'Driscoll, M., Aquilina, G., Karran, P.; *Mutat. Res.* **2000**, *462*, 71-82.

30. Ojwang, J.O., Grueneberg, D.A., Loechler, E.L.; *Cancer Res.*, **1989**, *49*, 6529-6537.

31. Maxam. A.M., Gilbert, W.W.; *Methods Enzymol.*, **1980**, 65, 499-560.

32. Singer. B., Grunberger, D.; Molecular Biology of Mutagens and Carcinogens. New York: Plenum Press. **1983**.

33. Reynolds, V.L., Molineux, I.J., Kaplan, D.J., Swenson. D.H., Hurley, L.H.; B*iochemistry*, **1985**, 24, 6228-6237.

34. Bichara. M., Fuchs, R.P.P.; *J. Mol. Biol.*, **1985**, 183, 341-351.

35. Hashimoto. Y., Shudo. K.; *Biochem. Biophys. Res. Commun.*, **1983**, *116*, 1100-1106.

36. Tong, W.P, Ludlum, D.B.; *Biochim. Biophys. Acta.*, **1980**, *608*, 174-181.

37. Pacheco, D.Y., Cook, C., Hincks. J.R., Gibson, N.W.; *Cancer Res.*, **1990**, *50*, 7555-7558.

38. Koo, H.S., Crothers, D.M.; *Proc. Natl. Acad Sci. U.S.A.*, **1988**, *85*, 1763-1767.

39. Bellon, S.F., Lippard, S.J.; *Biophys. Chem.*, **1990**, *35*, 179-188.

40. Wilds, C.J., Booth, J.D., Noronha, A.; *Tet. Let.*, **2006**, *47*, 9125-9128.

41. de la Santos, C., Zaliznyak, T., Johnson, F.; *J. Biol. Chem.*, **2001**, *276*, 9077-9082.

42. Patel, D.J., Shapiro, L., Kozlowski, S.A., Gaffney, B.L. Jones, R.A.; *Biochemistry*, **1986**, *25*, 1027-1036.

43. Patel, D.J., Shapiro, L., Kozlowski, S.A., Gaffney, B.L. Jones, R.A.; *Biochemistry*, **1986**, *25*, 1036-1042.

44. Kalnik, M.W., Li, B.F.L., Swann, P.F., Patel, D.J.; *Biochemistry*, **1989**, *28*, 6170-6181.

45. Kalnik, M.W., Li, B.F.L., Swann, P.F., Patel, D.J.; *Biochemistry*, **1989**, *28*, 6182-6192.

46. Watson, J.D., Crick, F.H. C.; *Nature*, 1953, *171*, 964.

47. Guengerich, F.P.; *Chem. Rev.*, **2006**, *106*, 420-452.

48. Leonard, G.A., Thomson, J., Watson, W.P., Brown, T.; *Proc. Natl. Acad. Sci. U.S.A.*, **1990**, *87*, 9573-9576.

49. Lin, J-K., Miller, J.A., Miller, E.C.; *Cancer Res.*, **1977**, 37, 4430-4438.

50. Essigmann, J.M., Green, C.L., Croy, R.G., Fowler, K., Buchi, G.H., Vogan, G.N.; *Cold Spring Harbor Symp. Quant. Biol.*, **1983**, *47*, 327-337.

51. Boiteux, S., Belleney, J., Roques, B.P., Laval, J.; *Nucleic Acids Res.*, **1984**, *12*, 5429-5439.

52. Persmark, M., Guengerich, F.P.; *Biochemistry*, **1994**, *33*, 8662-8672.

53. Fink, D., Aebi, S., Howell, S.B.; *Clinical Cancer Research*, **1998**, *4*, 1-6.

54. Morgan, A.R.; *Trends Biochem. Sci.*, **1993**, *18*, 160-163.

55. Hunter, W.N, Brown, T., Kneale, G., Anand, N., Rabinovich, D., Kennard, O.; *J. Biol. Chem.*, **1987**, *262*, 9962-9970.

56. Lawley, P.D., Brookes, P.; *Nature*, **1961**, *192*, 1081-1082.

57. Sowers, L.C., Saw, B.R., Veigl, M.L., Sedwick, W.D.; *Mutat. Res.*, **1987**, 177, 201-218.

58. Ezaz-Nikpay, K., Verdine, G.L.; *Chem. Biol.,* **1994**, *1*, 235-240.

59. Colvin, M., (Chabner, B.A., Alkylating agents. In: B.A. Chabner and J.M. Collins (eds.), Cancer Chemotherapy: Principles and Practice, pp. 276-313. Philadelphia: J.B. Lippincott Co., **1990**.

60. Alvino, E., Castiglia, D., Caporali, S., Pepponi, R., Caporaso, P., Lacal, P.M., Marra, G., Fischer, F., Zambruno, G., Bonmassar, E., Jiricny, J., D'Atri, S.; *Intl. J. Oncololgy,* **2006**, *29*, 785-797.

61. Kolodner, R.D., Marsischky, G.T.; *Curr. Opin. Genet. Dev.,* **1999**, *9*, 89-96.

62. Hoeijmakers, J.H.; *Nature,* **2001**, *411*, 366-374.

63. Mitchell, R.J., Farrington, S.M., Dunlop, M.G, Campbell, H.; *Am. J. Epidemiol.,* **2002**, *156*, 885-902.

64. Aaltonen, L.A., Peltomäki, P., Leach, F.S., Sistonen, P., Pylkkänen, L., Mecklin, J-P., Järvinen, H., Powell, S.M., Jen, J., Hamilton, S.R., Petersen, G.M., Kinzler, K.W., Vogelstein, B., and dela Chapelle, A.; *Science,* **1993**, *260*, 812-816.

65. Thibodeau, S.N., Bren, G., Schaid, D.; *Science,* **1993**, *260*, 816-819.

66. Risinger, J.I., Berchuck, A., Kohier, M.F., Watson, P., Lynch, H.T., Boyd, I.; *Cancer Res.,* **1993**, 53, 5100-5103.

67. Merlo, A., Mabry, M., Gabrielson, E., Vollmer, R., Baylin, S.B., Sidransky, D.; *Cancer Res.,* **1994**, *54*, 2098-2101.

68. Shridhar, V., Siegfried, J., Hunt, J., del Mar Alonso, M., Smith D.I.; *Cancer Res.,* **1994**, *54*, 2084-2087.

69. Han, H-J., Yanagisawa, A., Kato, Y., Park, J-G., Nakamura, Y.; *Cancer Res.,* **1993**, *53*, 5087-5089.

70. Wooster, R., Cleton-iansen, A-M., Collins, N., Mangion, I., Cornelis, R.S., Cooper, C.S., Gusterson, B.A., Ponder, B.A.I., von Deimling, A., Wiestler, O.D., Cornelisse, C.I., Devilee, P., Stratton, M.R.; *Nat. Genet.,* **1994**, *6*, 152-156.

71. King R.W., Jackson, P.K., Kirschner, M.W.; *Cell,* **1994**, *79*, 563-571.

72. Koi, M., Umar, A., Chauhan, D.P., Cherian, S.P., Carethers, J.M., Kunkel, T.A., Boland, C.R.; *Cancer Research,* **1994**, *54*, 4308-4312.

73. Huang, J., Papadopoulos, N., McKinley, A.J., Farrington, S.M., Curti,s L.J., Wyllie, A.H., Zheng, S., Willson, J.K.V., Markowitz S.D., Morin, P., Kinzler, K.W., Vogelstein, B., Dunlop, M.G.; *Proc. Natl. Acad. Sci. USA,* **1996**, *93*, 9049-9054.

74. Hassan, M., Ehrsson, H.; *J. Pharm. Biomed. Anal.,* **1986**, *4*, 95-101.

75. Paborji, M., Waugh, W.N., Stella, V.J.; *J. Pharm. Sci.,* **1987**, *76*, 161-165.

76. Waud, W.R., Plowman, J., Pauli, K.D., Narayanan, V.L., Bailey, D.M.; *Eur. J. Clin. Pharmacol.,* **1989**, *36*, 525-530.

77. Hassan, M., Ehrsson, H.; *Eur. J. Drug Metab. Pharmacokinet.,* **1987**, *12*, 71-76.

78. Harkey, M.A., Czerwinski, M., Slattery, J., Kiem, H-P.; *Cancer Invest.,* **2005**, *1*, 19-25.

79. Merchand, D.H., Remmel, R.P., Abdel-Monem, M.M.; *Drug Metab. Dispos.,* **1988**, *16*, 85-92.

80. Brodfuehrer, J.I., Wilki, T.J., Powis, G.; *Cancer Chemother. Pharmacol.,* **1988**, *22*, 120-125.

81. Rasimas, J.J., Dalessio, P.A., Ropson, I.J., Pegg, A.E., Fried, M.G.; *Protein Sci,* **2004**, *13*, 301-305.

82. Verdemato, P.E., Brannigan, J.A., Damblon, C., Zuccotto, F., Moody, P.C.E., Lian, L.; *Nuc. Acids Res.,* **2000**, *28*, 3710-3718.

83. Noll, D.M., Clarke, N.D.; *Nuc. Acid Res.*, **2001**, *29, 19*, 4025-4034.

84. Pegg, A.E.; *Mutat. Res.,* **2000**, *462*, 83-100.

85. Daniels, D.S., Mol, C.D., Arvai, A.S., Kanugula, S., Pegg, A.E., Tainer, J.A.; *EMBO J.,* **2000**, *19*, 1719-1730.

86. Crone,T.M., Goodtzova, K., Pegg, A.E.; *Mutat. Res.*, **1996**, *363*, 15–25.

87. Friedman, H.S., Kokkinakis, D.M., Pluda, J., Friedman, A.H., Cokgor, I., Haglund, M.M, Ashley, D.M., Rich, J., Dolan, M.E., Pegg, A.E., Moschel, R.C., McLendon, R.E., Kerby, T., Herndon, J.E., Bigner, D.D., Schold, S.C.; *J. Clin. Oncol.*, **1998**, *16*, 3570-3575.

88. Spiro, T.P., Gerson, S.L., Liu, L., Majka, S., Haaga, J., Hoppel, C.L., Ingalls, S.T., Pluda, J.M., Willson, J.K.V.; *Cancer Res.*, **1999**, 59, 2402-2410.

89. Bailey, C.C., Marsden, H.B., Morris-Jones, P.H.; *Cancer*, **1978**, *42*, 74-76.

90. Tosal, L., Commendador, M.A., Sierra, L.M.; *Mol. Genet. Genomics*, **2001**, *265*, 327-335.

91. a) Gukathasan, R., Massoudipour, M., Gupta, I., Chowdhury, A., Pulst, S., Ratnam, S., Sanghvi, Y.S., Laneman, S.A.; *J. Organomet. Chem.*, **2005**, *690*, 2603-2607; b) Sanghvi, Y.S., Andrade, M., Deshmukh, R.R., Holmburg, L., Scozzari, A.N., Cole D.L. in: G. Hartmann, S. Endres (Eds.), Kluwar Academic Publishers, Norwell, MA, 1999, p. 3.

92. Beaucage, S.L., Caruthers, M.H.; *Tet. Let.*, **1981**, *22*, 1859-1862.

93. Erlanson, D.A., Chen, L., Verdine, G.L.; *J. Am. Chem. Soc.* **1993**, *115*, 12583-12584.

94. Zang, H., Harris, T.M., Guengerich, F.P.; *Chem. Res. Toxicol.*, **2005**, *18*, 389-400.

95. Devadas, B., Leonard, N.J.; *J. Am. Chem. Soc.*, **1986**, *108*, 5012-5014.

96. Bhat, B., Leonard, N.J., Robinson, H., Wang, A.H.-J.; *J. Am. Chem. Soc.*, **1996**, 118, 10744- 10751.

97. Li, H., Qui, Y., Moyroud, E., Kishi, Y.; *Angew. Chem. Int. Ed.*, **2001**, *40, 8*, 1471-1475.

98. Kumar, S., Johnson, W.S., Tornasz, M.; *Biochem.*, **1993**, *32*, 1364-1372.

99. Speilmann, H.P., , Sastry, S.S., Hearst, J.E.; *Proc. Natl. Acad. Sci. U.S.A.*, **1992**, *89*, 4514-4518.

100. Kobertz, W.R., Essigmann, J.M.; *J. Am. Chem. Soc.*, **1997**, *119*, 5960-5961.

101. Nechev, L.V., Kozekov, I, Harris, C.M., Harris T.M.; *Chem. Res. Toxicol.*, **2001**, *14*, 1506-1512.

102. Nechev, L.V., Harris, C.M., Harris T.M.; *Chem. Res. Toxicol.*, **2000**, *13*, 421-429.

103. Zang, H., Harris, T.M., Guengerich, F.P.; *J. Biol. Chem.*, **2005**, *280*, 1165-1178.

104. Ni, J., Liu, T., Kolbanovskiy, A., Krezeminski, J., Amin, S., Geacintov, N.E.; *Anal. Biochem.,* **1998**, *264*, 222-229.

105. Gray, D.M., Ratliff, R.L., Vaughan, M.R.; *Methods in Enzymology, Vol. 211,* Academic Press, Inc., pg 389, **1992**.

106. Gray, D.M., Liu, J.J., Ratliff, R.L., Allen, F.S.; *Biopolymers*, **1981**, *20*, 1337-1382.

107. www.bio.davidson.edu/courses/molbio/molstudents/spring2003/holmberg/oligonucl eotide 08/05/2007

108. Pullman, B., Pullman, A.; *Nature*, **1963**, *199*, 467-469.

109. a) Schirrmacher, R., Schirrmacher, E., Mühlhausen, U., Kaina, B., Wängler, B.; *Curr. Org. Synth.* **2005**, *2*, 215-230; b) Nasutavicus, W.A., Love, J.; *J. Heterocyclic*

Chem., **1974**, *11*, 77-78; c) Igi, T. Hayashi; Eur. Pat. Appl. EP 543095 A2 930526; *Chem. Abstr.*, **1993**, *119*, 180812; d) Hanson, C. PCT Int. Appl. WO 92113859 A1 920820; *Chem. Abstr.*, **1992**, *117*, 233720. e) Killen, R. J. PCP Int. Appl. WO 9315075 A1 930805; *Chem. Abstr.*, **1994**, *120*, 30779; f) Zhou, J.; Tsai, J.-Y.; Bouhadir, K.; Shevlin, P.B. *Synth. Commun.*, **1999**, *29*, 3003-3009; g) Bowels, W.A.; Schneider, F.H.; Lewis, L.R.; Robins, R.K. *J. Med. Chem.*, **1963**, *6*, 471-477; h) Kilburis, J.; Lister, J. H. *J. Chem. Soc. (C)*, **1971**, 3942-3947; i) Linn, J. A.; McLean, E. W.; Kelly, J. L. *J. Chem. Soc., Chem. Commun.*, **1994**, 913-914; j) Höfle , G., Steglich, W., Vorbrüggen, H.; *Angew. Chem.*, **1978**, *90*, 602-615; k) Hassner, A.; Krepski, L. R.; Alexanian, V. *Tetrahedron*, **1978**, *34*, 2069-2076; l) Scriven, E. F. V. *Chem. Soc. Rev.*, **1983**, *13*, 129-161.

110. Tomasz, M., Olsen, J., Mercado, C.M.; *Biochemistry*, **1972**, *11*, 1235-1241.
111. Palus, E.; M.Sc. Thesis, Concordia University, 2007.
112. Beranek D.T.; *Mutation Res.*, **1990**, *231*, 11-30.
113. Zhong, M., Nowak, I., Robins, M.J.; *J. Org. Chem.*, **2006**, *71*, 7773-7779.
114. Zhou, J., Tsai, J-Y., Bouhadir, K., Shevlin, P.B.; *Synth Commun* , **1999**, *29*, 3003-3009.
115. Jenny, T.F., Benne, S.A.; *Tet. Lett.*, **1992**, *33*, 6619-6620.
116. Gao, X., Gaffney, B.L., Hadden, S., Jones, R.A.; *J. Org. Chem.*, **1986**, *51*, 755-758.
117. Mitsunobu, O.Y.M.; *Bull. Chem. Soc. Jpn.*, **1967**, *40*, 2380-2382.
118. Wilds, C.; Concordia University, *Private communication.*
119. Wu, T., Ogilvie, K.K., Pon, R.T.; *Nucleic Acids Res.*, **1989**, *17*, 3501-3517.
120. Hayakawa, Y, Wakabayashi, S., Kato, H, Noyori, R.; *J. Am. Chem. Soc.*, **1990**, *112*, 1691-1696.
121. Rigby, J.H., Moore, T.L., Rege, S.; *J. Org. Chem.*, **1986**, *51*, 2402-2404.
122. Trichtinger, T., Charubala, R., Pfleiderer, W.; *Tet. Lett.*, **1983**, *24*, 711-714.
123. Himmelsbach, F., Schulz, B.S., Trichtinger, T., Charubala, R., Pfleiderer, W; *Tetrahedron*, **1984**, *40*, 59-72.
124. Otte, R.D., Sakata, T., Guzei, I.A., Lee, D.; *Org. Lett.*, **2005**, *7*, 495-498.
125. Marchán, V., Cieslak, J., Livengood, V., Beaucage, S.L.; *J. Am. Chem. Soc.*, **2004**, *126*, 9601-9610.
126. Harwood E.A., Hopkins, P.B., Sigurdsson, B.R.; *J. Org. Chem.*, **2000**, *65*, 2959-2964.
127. Zhu, Q., Delaney, M.O., Greenberg, M.M.; *Bioorg. Med. Chem. Lett.*, **2001**, *11*, 1105–1107.
128. Noll D.M., da Silva, M.W., Noronha, A.M., Wilds, C.J., Colvin, O.M., Gamcsik, M.P., Miller, P.S.; *Biochemistry*, **2005**, *44*, 6764-6775.
129. Noronha, A.M., Noll D.M., Wilds, C.J., Miller, P.S.; *Biochemistry*, **2002**, *41*, 760-771.
130. Noll D.M., Noronha, A.M., Wilds, C.J., Miller, P.S.; *Frontiers in Bioscience*, **2004**, *9*, 421-437.
131. Wilds, C.J., Noronha, A.M., Robidoux, S., Miller, P.S.; *J. Am. Chem. Soc.*, **2004**, *126*, 9257-9265.
132. Wilds, C.J., Noronha, A.M., Robidoux, S., Miller, P.S.; *Nucleos. Nucleot. Nuc. Acids.*, **2005**, *24*, 965-969.
133. Trempe, J.F., Wilds, C.J., Denisov A.Y., Pon, R.T., Damha, M.J., Gehring, K.; *J. Am. Chem. Soc.*, **2001**, *123*, 4896-4903.
134. Puglisi, J.D., Tinoco, I., Jr.; *Methods Enzymol.*, **1989**, *180*, 304-325.
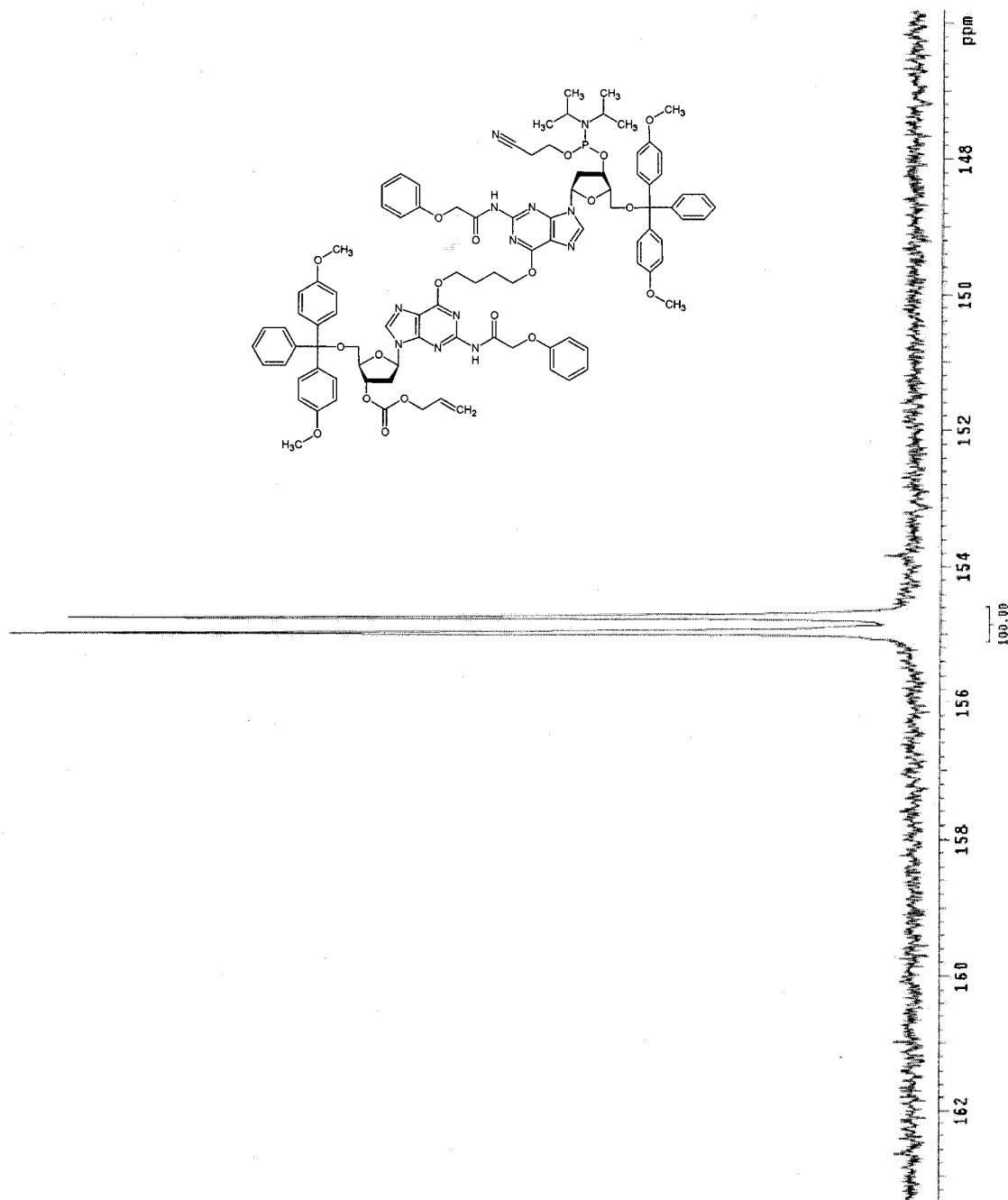
300 MHz ^{1}H NMR spectrum of compound **6a**
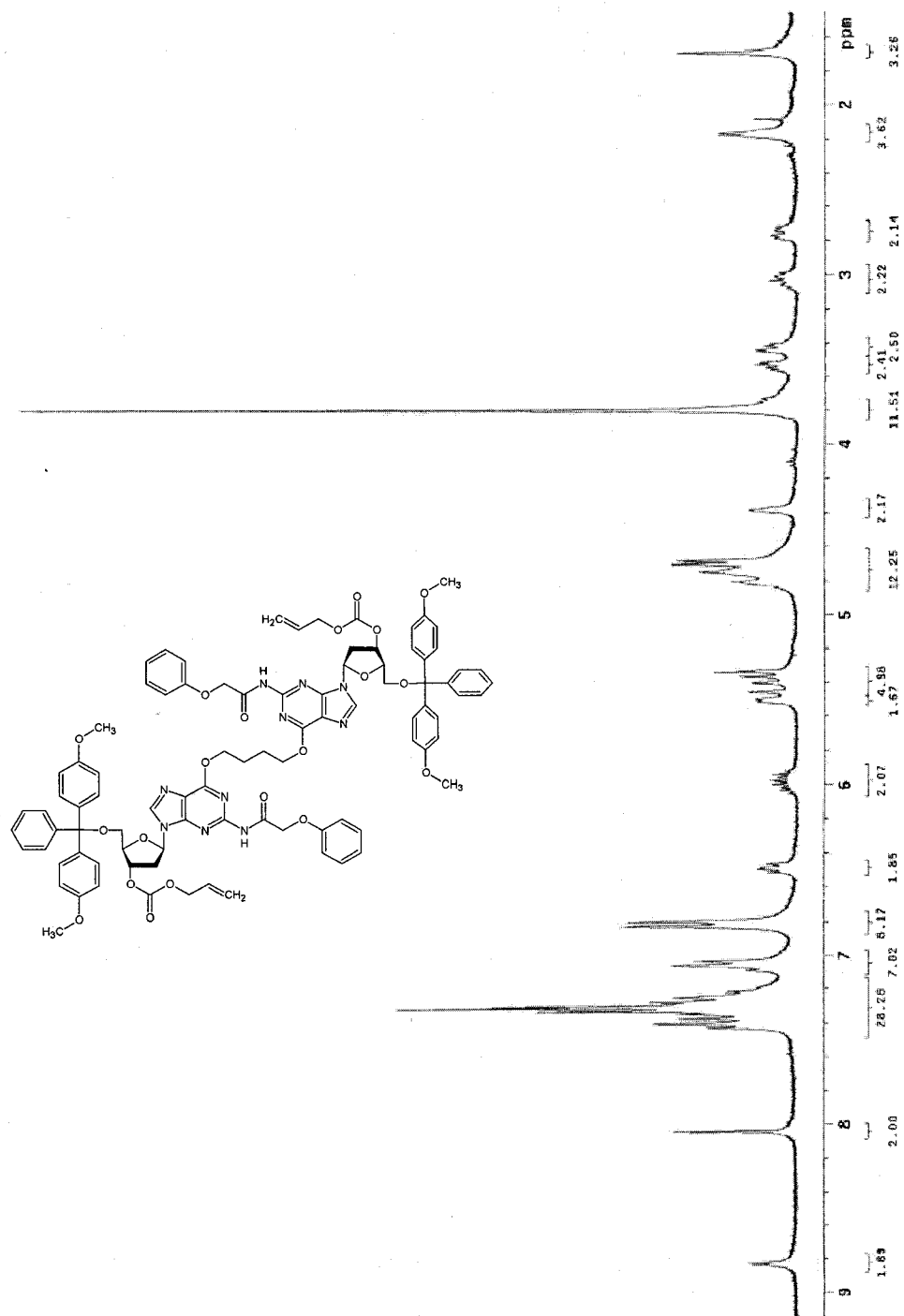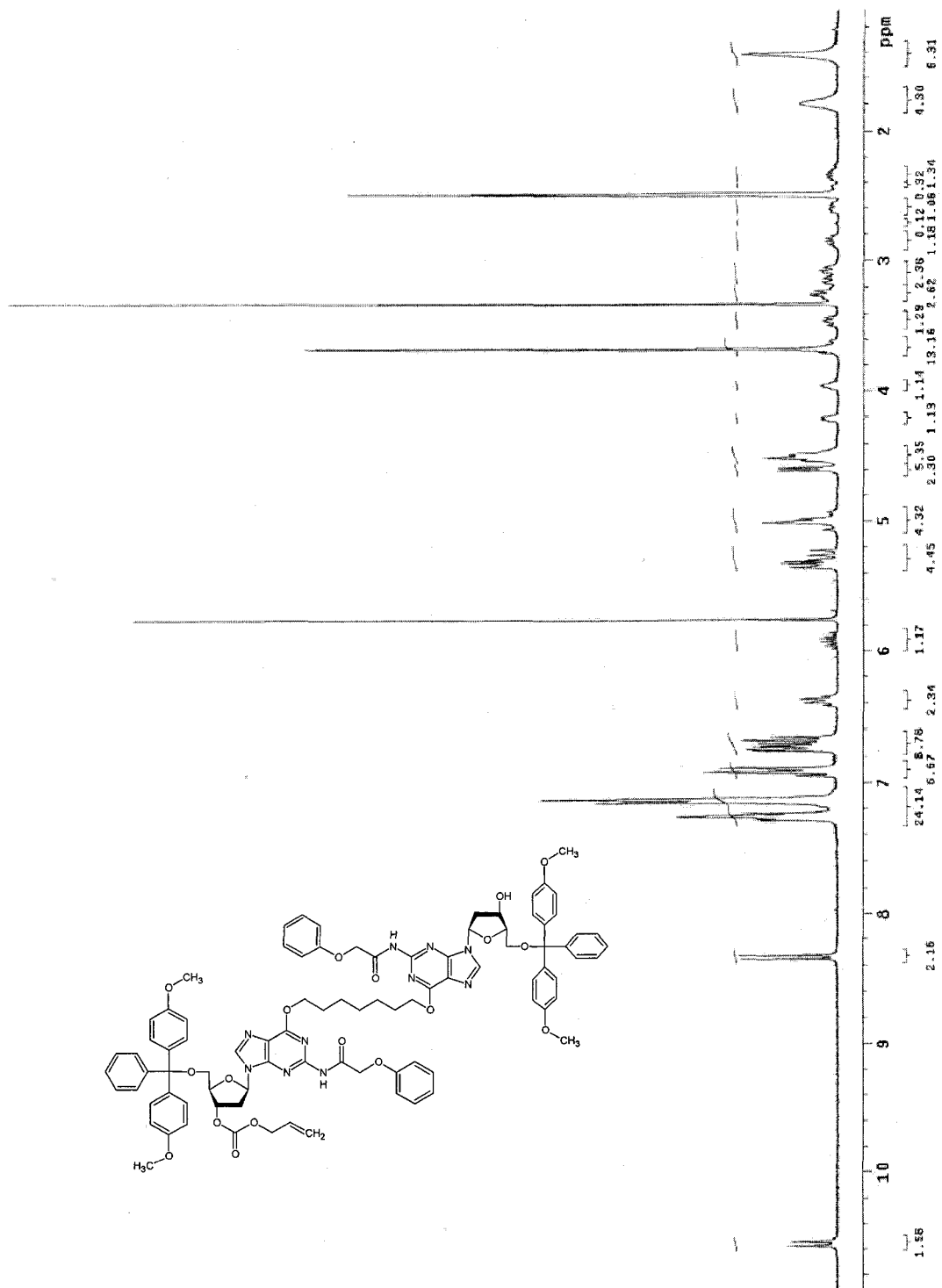
App-i

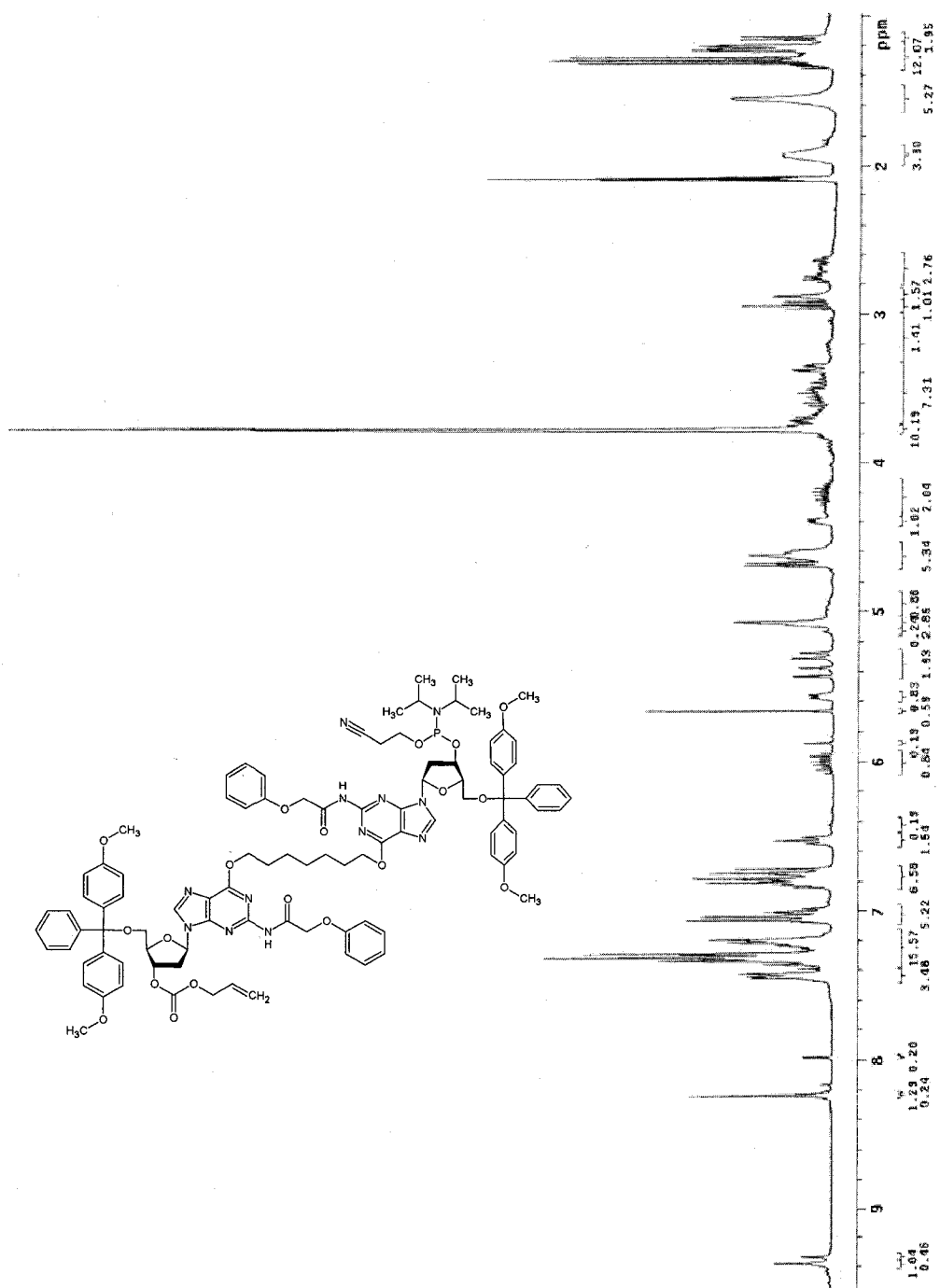# 300 MHz ^{1}H NMR spectrum of compound 7b

App-ii

300 MHz ^{1}H NMR spectrum of compound **11a**

App-iii

300 MHz ¹H NMR spectrum of compound **12a**

App-iv

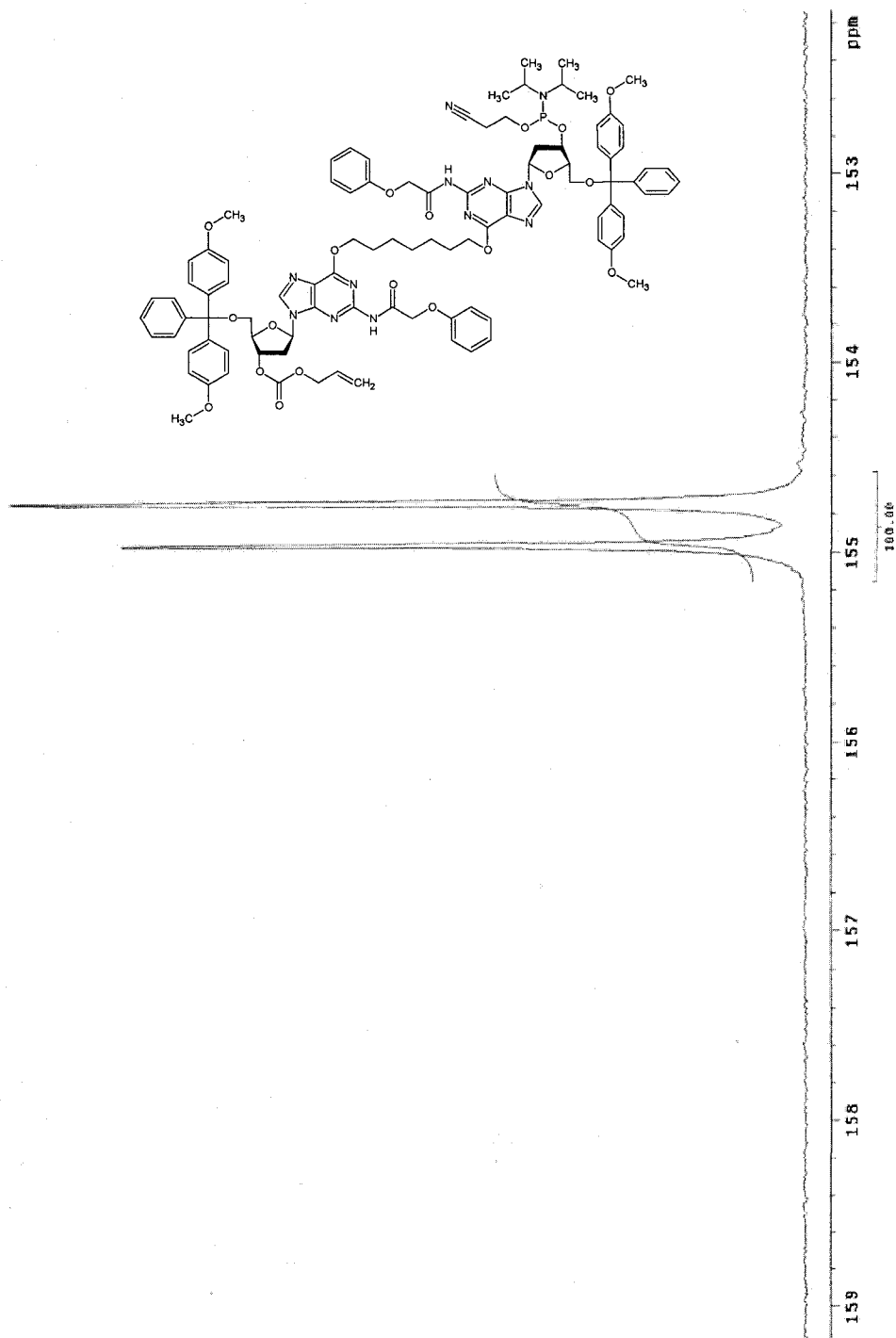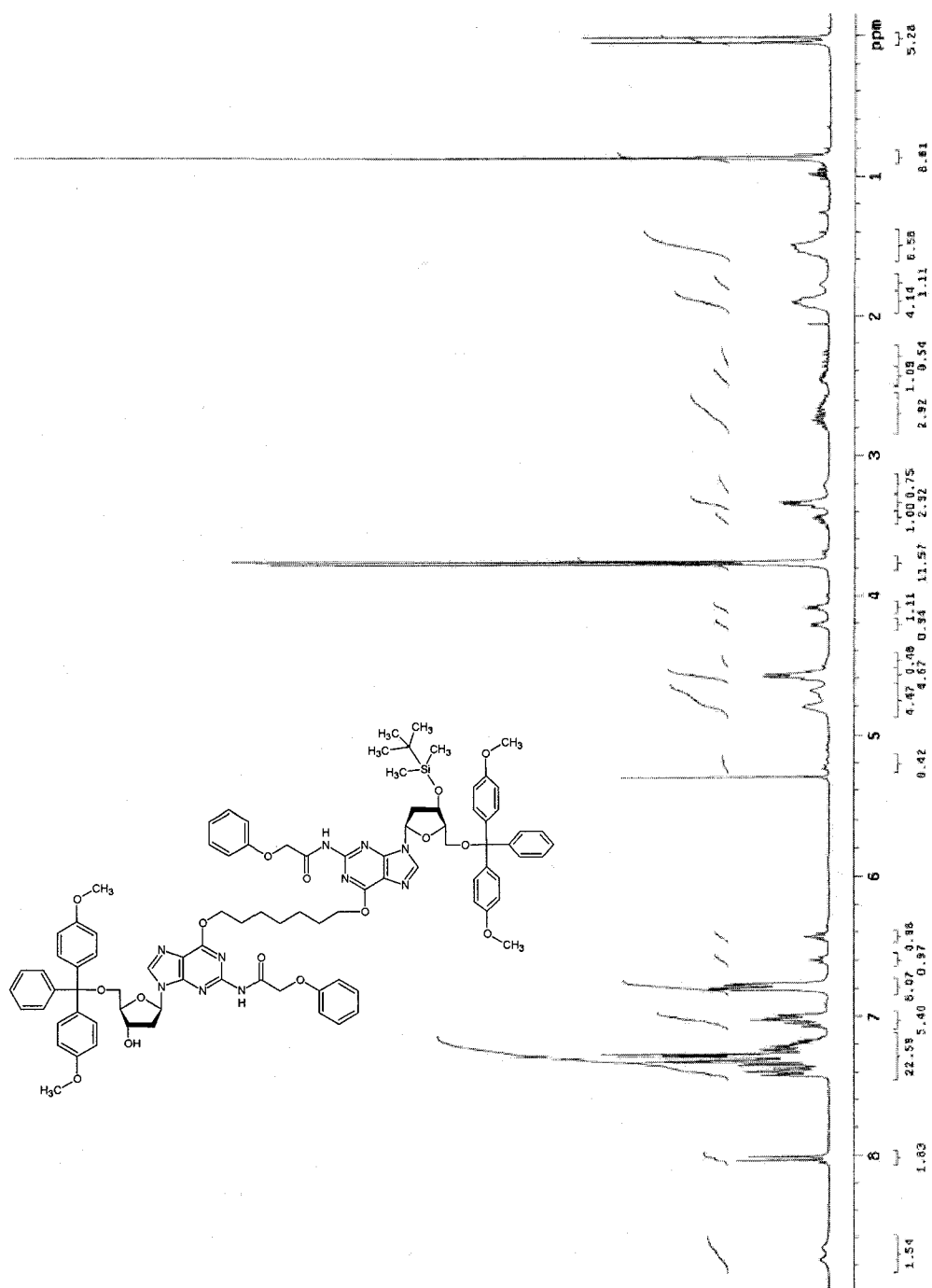121 MHz ^{31}P NMR spectrum of compound **12a**

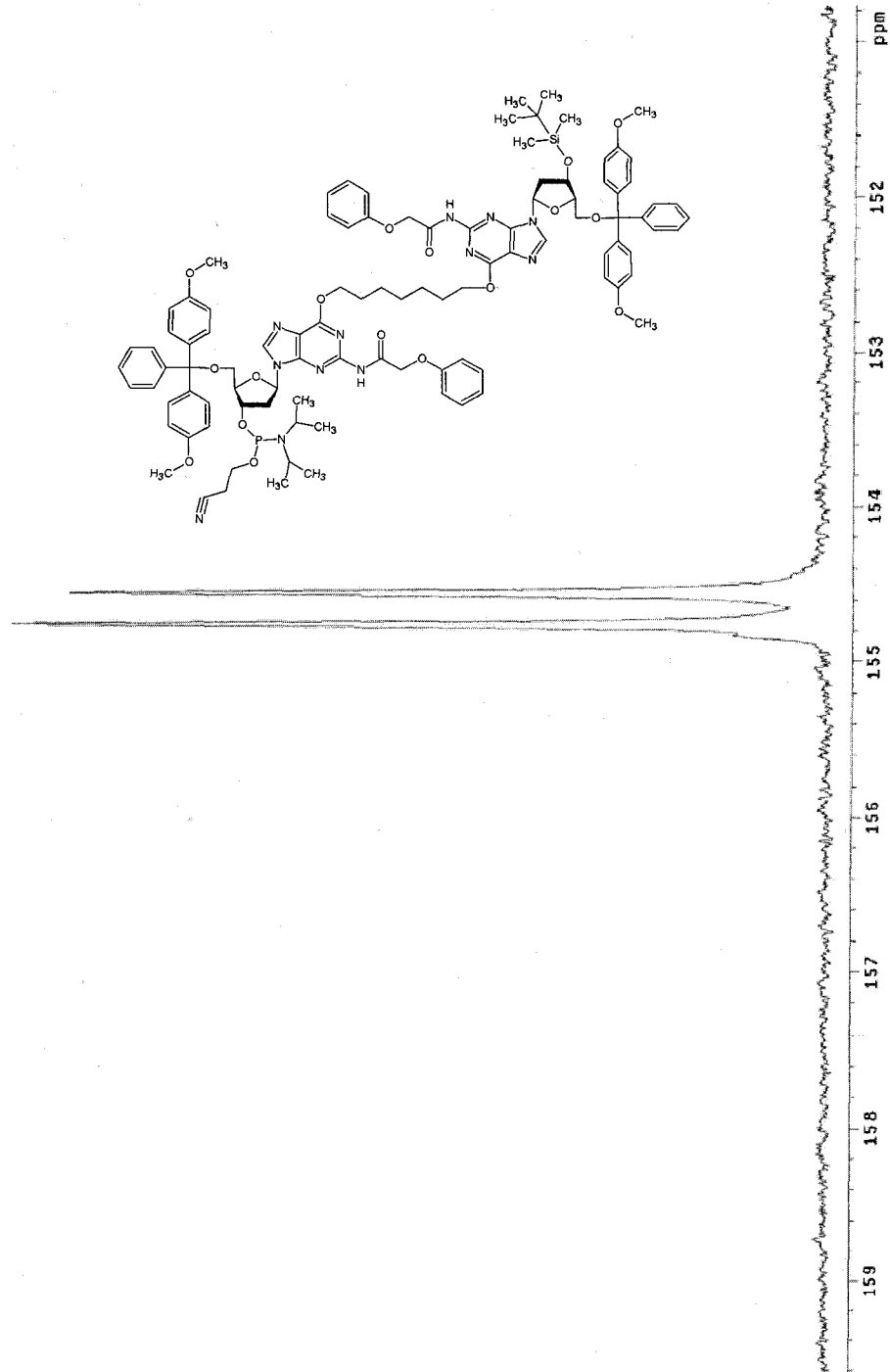

App-v
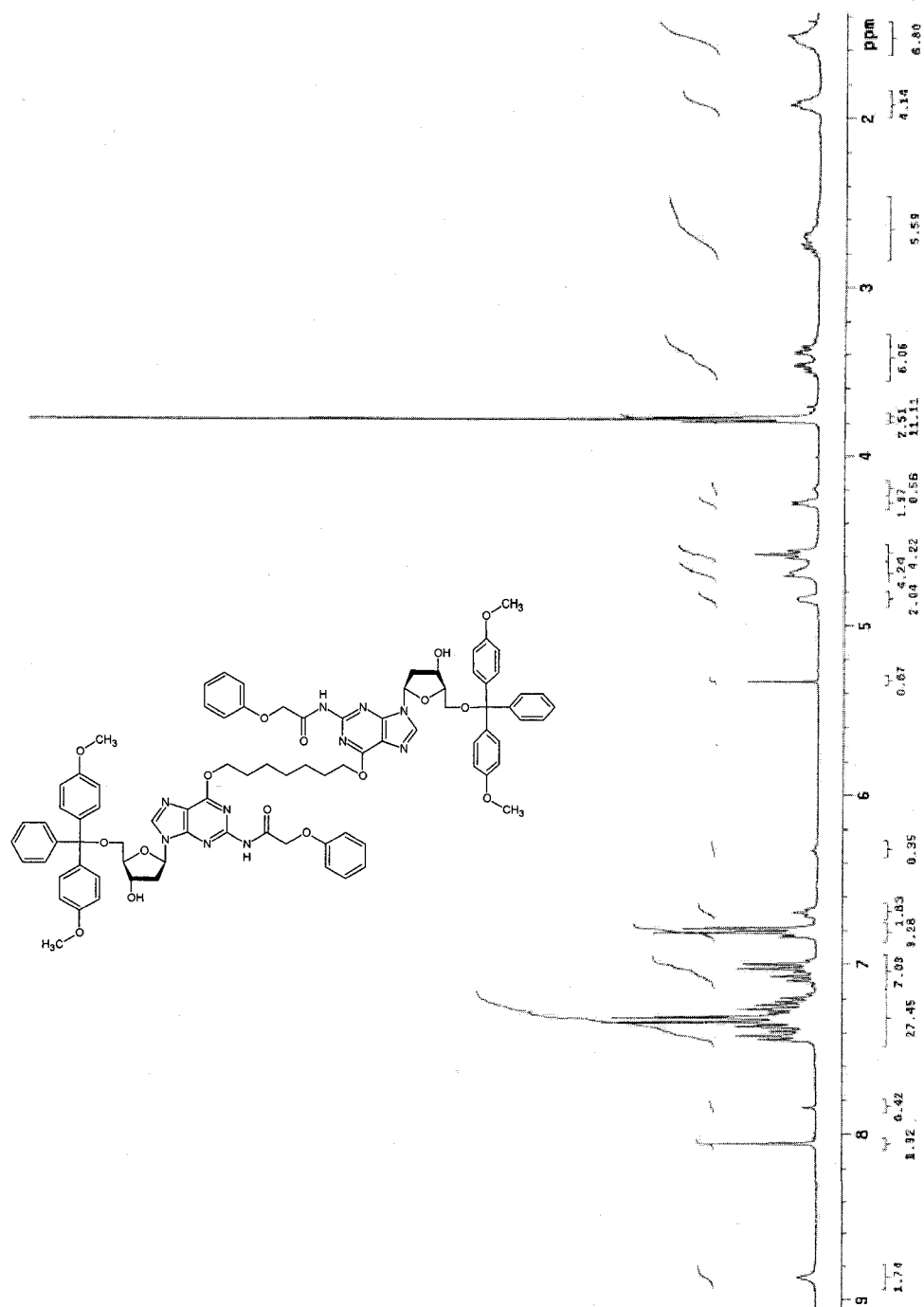
300 MHz ^{1}H NMR spectrum of compound 11b

App-vi

300 MHz ¹H NMR spectrum of compound 12b

121 MHz ^{31}P NMR spectrum of compound **12b**

App-viii

300 MHz ^{1}H NMR spectrum of compound 13

App-ix

300 MHz ¹H NMR spectrum of compound **14**
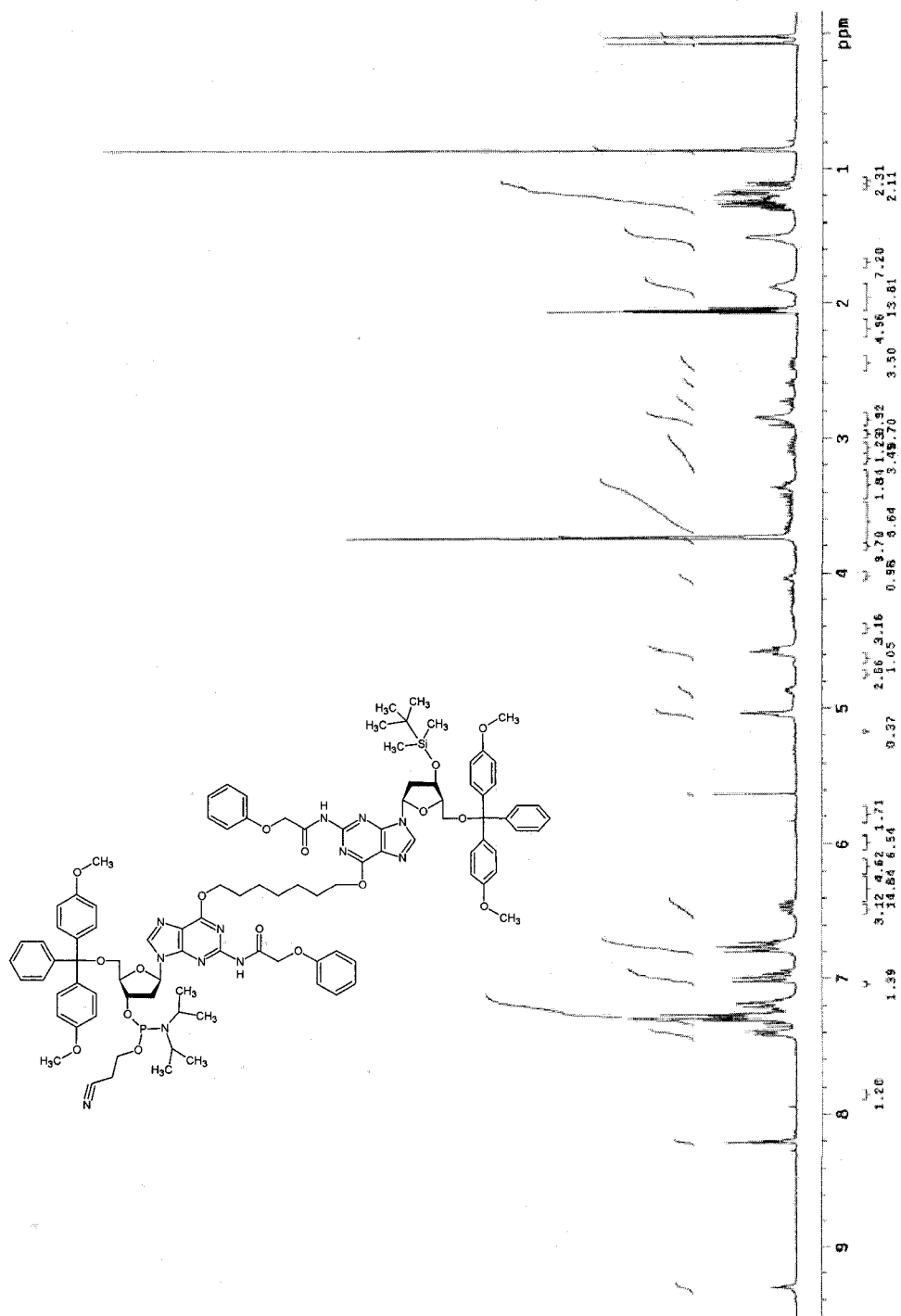
App-x

121 MHz ^{31}P NMR spectrum of compound **14**

300 MHz ^{1}H NMR spectrum of compound **18**

App-xii

MALDI-TOF of **22**

App-xiii