TOWARDS THE SYNTHESIS OF JANUS-AT PHOSPHORAMIDITES

by

Wenbo Liu

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE
in
The Faculty of Graduate and Postdoctoral Studies
(Chemistry)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

August 2013

© Wenbo Liu, 2013
Abstract

A Janus heterocycle is one kind of heterocycle that possesses hydrogen bonding donors or acceptors at both sides compared to regular nuclear bases. This provides an interesting synthetic target as well as a tool for testing an intriguing biological hypothesis related to DNA recognition, namely can both strands in the DNA duplex be recognized simultaneously by a single oligomeric strand of Janus heterocycles? In order to test this hypothesis, oligonucleotides containing at least one Janus heterocycle must be synthesized. One challenge involved in the synthesis of oligonucleotides with a Janus heterocycle is the extremely low solubility of Janus heterocycle phosphoramidites in organic solvents, namely MeCN. To solve the solubility issue, two sulfone-masked Janus-AT phosphoramidites were designed and constructed in this thesis. In addition to questions of solubility, the other obstacle related with Janus heterocycle oligonucleotide construction is the potential for deglycosylation, analogous to depurination, which might occur under the acidic conditions used during solid phase synthesis of oligonucleotides. The carbocyclic analog of deoxyribose, in lieu of 2-deoxyribose was employed to construct Janus heterocycle phosphoramidites in order to obviate the potential for depurination-like elimination of the nucleobase from the sugar.
Preface

All the experiments described in this thesis were conducted by the author of this thesis. The author and Dr. Perrin conceived all the protocols employed in the thesis except for previously reported compounds. The computational calculation for the pKa of the Janus heterocycle along with general biophysical experiments concerning oligonucleotides was done by Dr. Eric Largy. All NMR spectra were obtained by the author and analyzed by the author in consultation with Dr. Perrin. The analysis of the research data was conducted by the author and Dr. Perrin, part of which was done with the help of Dr. Eric Largy.
Table of contents

Abstract ........................................................................................................................................... ii
Preface ............................................................................................................................................... iii
Table of contents ............................................................................................................................... iv
List of figures ...................................................................................................................................... vi
List of schemes ................................................................................................................................. viii
List of abbreviations ......................................................................................................................... x
Acknowledgements ........................................................................................................................... xi

Chapter 1: Background introduction ................................................................................................. 1
  1.1: Introduction to DNA, RNA, and central dogma ........................................................................ 1
  1.2: Molecular recognition of DNA and RNA ............................................................................. 4
  1.3: Introduction to Janus heterocycle ......................................................................................... 7
  1.4: Phosphoramidite technology used to construct oligonucleotide ......................................................... 13
  1.5: Previous results unpublished in Perrin’s lab and aims of this thesis ......................................... 15

Chapter 2: The synthesis of sulfone-masked Janus-AT-2 phosphoramidite................................. 17
  2.1 Introduction ............................................................................................................................... 17
  2.2 Retrosynthetic analysis of sulfone-masked Janus-AT-2 phosphoramidite .............................. 24
  2.3: Synthesis and discussion of sulfone-masked Janus AT-2 phosphoramidite ........................... 25
    2.3.1: Synthesis of 7 ..................................................................................................................... 25
    2.3.2: Synthesis of 8 ..................................................................................................................... 26
    2.3.3: The closure of synthesis of sulfone-masked Janus-AT-2 phosphoramidite .................... 27
  2.4: Conclusion ............................................................................................................................... 30

Chapter 3: The synthesis of sulfone-masked Janus-AT-1 phosphoramidite................................. 32
3.1 Introduction ............................................................................................................................................. 32
3.2: Retrosynthetic analysis of the sulfone-masked Janus-AT-1 phosphoramidite .... 38
3.3: Synthesis of the sulfone-masked Janus-AT-1 phosphoramidite.......................... 41
   3.3.1: Synthesis of 29 ............................................................................................................................. 41
   3.3.2: Synthesis of 24 ............................................................................................................................. 41
   3.3.3: Synthesis of 22 ............................................................................................................................. 44
   3.3.4: Closure of the sulfone-masked Janus-AT-1 phosphoramidite ...................... 45
3.4: Conclusion ............................................................................................................................................. 49

Chapter 4: Conclusions and future work................................................................................. 50

References: ............................................................................................................................................ 52

Appendix: Experimental details and characteristic spectrums ......................... 61
   A1: General Materials and equipment ................................................................................................. 61
   A2: Synthesis protocols and characteristic evidence for all compounds ............. 61
List of figures

Figure 1: Two Watson Crick base pairing patterns in DNA(R=deoxyribose) ................. 1
Figure 2: Two Watson Crick base pairing patterns in RNA(R=ribose) .......................... 2
Figure 3: Central dogma of molecular biology .................................................................. 3
Figure 4: Different backbones of DNA, RNA and PNA (left to right, respectively). ........ 4
Figure 5: Representative Janus heterocycles ................................................................... 8
Figure 6: Janus-wedge recognizing regular DNA base pairs ......................................... 10
Figure 7: McLaughlin’s work on Janus-Wedge triplex formation with CT mismatch ....... 12
Figure 8: McLaughlin’s work on regular DNA base pair ............................................... 13
Figure 9: Solid phase synthesis of oligonucleotide based on phosphoramidite ............. 14
Figure 10: First generation of Janus-AT phosphoramidite ............................................ 15
Figure 11: Two targets of this thesis .............................................................................. 16
Figure 12: Previous Janus-AT bases synthesized ......................................................... 19
Figure 13: Comparison between carbocyclic nucleoside and regular 2-deoxyribose nucleoside ................................................................. 20
Figure 14: Previously synthesized Janus-AT-1 phosphoramidite (left) and the new target contemplated in this thesis (right) ................................................................. 21
Figure 15: Sulfone-masked Janus-AT-2 phosphoramidite ............................................. 23
Figure 16: Acid-base equilibrium of Janus-AT-2 base .................................................... 34
Figure 17: Different hydrogen bond pattern between Janus-AT-2 and its protonated form ....................................................................................................................... 34
Figure 18: Three benzyl-sulfone masked Janus-AT base candidates with low basicity at N7 ....................................................................................................................... 35
Figure 19: pKa value of conjugated acid of nitrogen in Janus-AT-1 .........................36

Figure 20: New generation of sulfone-masked Jnaus-AT-1 phosphoramidite with less basicity .........................................................................................................................37

Figure 21: Mass spectrometry (ESI positive mode) to confirm the Zincke salt 35 ...........................................................................................................................................46

Figure 22: Mass spec information (ESI positive mode and negative mode) to confirm 34 ...........................................................................................................................................47

Figure 23: Sulfone-masked Janus-AT-2 phosphoramidite .....................................50

Figure 24: Sulfone-masked Janus-AT-1 phosphoramidite .....................................50
List of schemes

Scheme 1: Acid promoted elimination of a nucleobase, typically seen for purines
........................................................................................................................................19

Scheme 2: Previous method to construct Janus-AT-2 (R=butyl).........................22

Scheme 3: Retrosynthetic analysis of sulfone-masked Janus-AT-2

phosphoramidite .............................................................................................................24

Scheme 4: The scheme to synthesize 7.................................................................25

Scheme 5: The scheme to synthesize 8.................................................................26

Scheme 6: The closure of synthesis of sulfone-masked Janus-AT-2

phosphoramidite .............................................................................................................27

Scheme 7: Comparison of cyclization in the presence and absence of

triethylamine .................................................................................................................28

Scheme 8: Previous methodology to construct butyl Janus-AT-1 base(R=butyl)36

Scheme 9: Retrosynthetic analysis of sulfone-masked Janus-AT-1

phosphoramidite (Part A)...........................................................................................38

Scheme 10: Two possible disconnections for 23....................................................39

Scheme 11: Retrosynthetic analysis of the sulfone-masked Janus-AT-1

phosphoramidite (Part B)...........................................................................................40

Scheme 12: Attempt to synthesize 24 ......................................................................41

Scheme 13: A modified pathway to construct 23....................................................43

Scheme 14: Synthesis of 23 .....................................................................................44

Scheme 15: Synthesis of 22 .....................................................................................44

Scheme 16: Possible pathway to form Zincke salt 35 .............................................46

Scheme 17: Reaction between Zincke salt and MeOH ...........................................47
Scheme 18: Synthesis of 33 ................................................................. 48
Scheme 19: Synthesis of 21 ................................................................. 48
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DCM</td>
<td>dichloromethane</td>
</tr>
<tr>
<td>DFT</td>
<td>density function theory</td>
</tr>
<tr>
<td>DIPEA</td>
<td>diisopropylethylamine</td>
</tr>
<tr>
<td>DMAP</td>
<td>4-dimethylaminepyridine</td>
</tr>
<tr>
<td>DMF</td>
<td>dimethylformamide</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethylsulfoxide</td>
</tr>
<tr>
<td>DMT-CI</td>
<td>dimethoxytritylchloride</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>HOMO</td>
<td>Highest Occupied Molecular Orbital</td>
</tr>
<tr>
<td>LUMO</td>
<td>Lowest Unoccupied Molecular Orbital</td>
</tr>
<tr>
<td>MCPBA</td>
<td>meta-chloroperoxybenzoic acid</td>
</tr>
<tr>
<td>PNA</td>
<td>peptide nucleic acid</td>
</tr>
<tr>
<td>NMP</td>
<td>N-methylpyrrolidone</td>
</tr>
<tr>
<td>Py</td>
<td>pyridine</td>
</tr>
<tr>
<td>Red-Al</td>
<td>sodium bis(2-methoxyethoxy)aluminumhydride</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>TBAF</td>
<td>tetra-n-butylammonium fluoride</td>
</tr>
<tr>
<td>TBSCI</td>
<td>tert-butyldimethyilsilylchloride</td>
</tr>
<tr>
<td>THF</td>
<td>tetrahydrofuran</td>
</tr>
</tbody>
</table>
Acknowledgements

First and foremost, I would like to express my deep gratitude to my family, including my parents, my grandparents and my younger sister. It is their unconditional love and unselfish support that encourage me to walk on the scientific research pathway that is full of obstacles and frustration. Without them, it is for sure that there is no such thesis.

Then, I would like to thank my supervisor Dr. David Perrin. Over last three years, Dr. Perrin has taught me lot of things, which are crucial to being a good and fast chemist, and moreover a careful scientist. I always feel that I am lucky enough to have a chance to work in this laboratory at the outset of my scientific career. His enthusiasm, philosophy for science has affected me deeply; his encyclopedic knowledge of chemistry will be something I would like to emulate as a scientist. Thanks to his guidance and patience, I have changed into a qualified master graduate student from an undergraduate student who nearly knew nothing about organic chemistry.

In addition, I am very grateful to Dr. Eric Largy, whose expertise in biophysical chemistry provided a rapid evaluation of Janus-AT containing oligonucleotides. This work was essential for validating the synthetic work of this thesis and resulted in the publication of some of this work in ChemBioChem. Finally, I am indebted to all the members in Perrin’s lab. It is because of you that I feel happy in this laboratory every day and enjoy the fantastic scientific research journey over last two years.
1.1: Introduction to DNA, RNA, and central dogma

DNA (deoxyribonucleic acid), whose structure was discovered in 1953, is arguably one of the most important and artistic molecules in nature.\(^1\) DNA consists of 2-deoxy-ribose, four types of nucleobases, and a phosphate-sugar alternating backbone. Of all three components, the chemistry of nucleobases is perhaps the most important aspect of DNA in chemical-biological terms as these hold the clue to both genetic information transfer as well as a means of rationally designing ligands that might interfere with the normal biological processes thereby forming the basis for treating diseases with a genetic component.

Indeed, DNA’s role as the reservoir of genetic information is realized by the nucleobases on each strand that are typically found as complementary base pairs. There are four kinds of DNA nucleobases in nature, which are adenine (A), thymine (T), guanine (G) and cytosine (C). According to the structure features of four bases, A is paired with T through two hydrogen bonds while G is paired with C through three hydrogen bonds (Figure 1). Different combinations and sequences of bases represent different genetic information.

![Two Watson Crick base pairing patterns in DNA](R=deoxyribose)

Figure 1: Two Watson Crick base pairing patterns in DNA (R=deoxyribose)
RNA (ribonucleic acid), similar to DNA, comprises ribose, a sugar-phosphate alternating backbone and four kinds of nucleobases. There are slight yet significant chemical differences between DNA and RNA as well as very different functions within the cell. The sugar contained in RNA is a ribose instead of a 2-deoxyribose found in DNA and four nucleobases in RNA are Adenine (A), Uracil (U), Guanine (G) and Cytosine (C) (Figure 2), with U differing from T by only a methyl group. DNA and RNA have very different secondary structures because DNA is formed by two complementary polynucleic stands while RNA is generally portrayed as a single stranded structure, which makes it possible to fold into much fewer regular structures, which are nevertheless discrete, each of which is based on the individual sequence of the RNA. Diverse roles of RNA have been discovered while the most frequent RNA functional roles are to code, decode and regulate genes’ transcription and expression.²

![Figure 2: Two Watson Crick base pairing patterns in RNA(R=ribose)](image)

The functional relationship between DNA and RNA is summarized in what is called the “central dogma of molecular biology” (Figure 3), which describes the flow of genetic information in biological systems that comprises the three most important biological macromolecules, DNA, RNA and protein. In this dogma,
genetic information in DNA can be transcribed into RNA, from which gene can be expressed in the form of protein as shown in Figure 3.  

\[
\text{DNA} \quad \xleftrightarrow{} \quad \text{RNA} \quad \rightarrow \quad \text{Protein}
\]

**Figure 3: Central dogma of molecular biology**

DNA’s double strand complementary structure along with extreme solvolytic stability compared to RNA is the foundation of DNA’s precise replication, transcription, which further affects RNA’s expression, rendering DNA suitable as the carrier of genetic information. RNA is biosynthesized as a single stranded polynucleotide, which means RNA is not as rigid as DNA. Gratifyingly, it is due to the flexibility of RNA that RNA plays different crucial roles in biological activities, such as diverse functions of messenger RNA (mRNA), transfer RNA (tRNA) and ribosomal RNA (rRNA).  

At the molecule level, the directionality of hydrogen bonding interactions between different base pairs provides the reliability of DNA’s transcription and expression. As shown above, purine bases highly favor to form base pairs with pyrimidines and vice versa. Mismatches between different base pairs result in unstable structures, which would not survive in cells. Research in bioorganic chemistry, starting with pioneering work by Khorana and Caruthers, led to the synthetic product of short pieces of DNA and RNA termed oligonucleotides. These advances led not only to the development of numerous probes for biology, new tools for sequencing, and a validation of both RNA and DNA structures, but also new approaches in therapeutic development starting with antisense and antigen
agents (*vide infra*). The synthetic approach to DNA and RNA was further extended to novel backbones of numerous kinds where both the ribose and polyphosphodiester backbone have been varied with increasingly diverse compositions. A notable example of this, first reported in 1991, was peptide nucleic acid (PNA), as the surrogate of DNA and RNA. Instead of a phosphate-sugar backbone found in DNA and RNA, PNA backbone is composed of repeating N-(2-aminoethyl)-glycine. PNA backbone is devoid of any negative charge, which brings special properties DNA and RNA lack. Recently, lots of synthetic PNA have been used in molecular biology procedures, diagnostic assays and antisense therapies. The comparison between DNA, RNA and PNA backbone is shown in Figure 4.

![Figure 4: Different backbones of DNA, RNA and PNA (left to right, respectively)](image)

### 1.2: Molecular recognition of DNA and RNA

Molecular recognition refers to the specific interaction between two or more molecules via non-covalent interactions such as hydrogen bonding, metal
coordination, hydrophobic forces, Van der Waals forces, and π-π stacking.\textsuperscript{6-7} Molecular recognition can account for many types of biological activities in living organisms, such as highly specific substrate interactions with enzymes as well as the flow of genetic information from DNA to RNA to proteins, a process that must be highly specific and which is marked by high affinity interactions between receptors and ligands. Not only in biological systems, molecular recognition is also finding wider and wider applications in supramolecular systems, which is a very attractive research field.\textsuperscript{5-6}

Recognizing DNA and RNA is of crucial importance in that they are involved in diverse biological processes as evidenced in the aforementioned central dogma. In regards to gene transcription and expression, the maturation of molecular biology as a chemical science provides researchers with a pretext for rationally designing molecules that could bind to defined targets on DNA in order to disrupt the regular biological processes that lead to the production of certain targets (e.g. viral coat proteins, oncoproteins). Therefore classes of molecules that could recognize a specific sequence of DNA or RNA would represent potential drugs. One major requirement to recognize DNA is to know sequences that are of interest for disrupting the processing of a specific gene whose protein target is responsible for certain pathology. Thanks to the human genome project and related studies on DNA sequences, the human gene sequences are now freely available, which makes the selection of DNA targets much easier.
So far, several strategies have been developed to recognize DNA and RNA with the explicit intention of interfering with some aspect of gene processing. These include the well-known antisense strategy along with a less applied anti-gene strategy that is directed specifically at DNA and therefore involves the development of ligands that recognize DNA through the major and minor grooves. The antisense strategy\textsuperscript{8-11} employs an oligonucleotide that is complementary to a specific RNA sequence, particularly messenger RNA (mRNA) such that it forms a standard double strand with mRNA. It is known that mRNA must be single stranded in order to function as the gene expression intermediary between DNA and protein. Therefore, once the mRNA of a specific gene is associated with the antisense oligonucleotide via standard base pairs, the ribosome-mediated flow of genetic information would be blocked and the expression of that gene will be inhibited. Recently, antisense therapy has witnessed significant success. Two antisense drugs have been approved and around 40 are under clinical trial, 20 of which are in advanced clinical trial (Phase 2 or 3).\textsuperscript{9-10} For DNA recognition, the first strategy involves designing oligonucleotides that recognize DNA in the major groove,\textsuperscript{12-13} whereby the third oligonucleotide forms a DNA triplex with the original DNA duplex. The basis of DNA triplex formation relies on the H-bonding nature of purines, which have the capacity to form both Watson-Crick pairing interactions as well as Hoogsteen pairing interactions simultaneously. Nevertheless for successful DNA recognition by triplex forming oligonucleotides, one stand of the target DNA duplex must be purine-rich (A or G).\textsuperscript{12-13} Due to this requirement, this strategy is essentially completely restricted to the recognition of
DNA strands that contain runs of purines all on one strand. Hence while the duplex itself is considered to bind to a triplex forming oligonucleotide, only one strand (the purine strand) provides all of the interactions that enable what appears to be duplex recognition. The minor groove binding strategy\textsuperscript{14-15} is based on the DNA minor groove binders, which generally show a preference for binding to A/T rich regions of DNA. DNA minor groove binders share two common structural characteristics: 1) they almost always possess a positive charge; 2) they present different aromatic and/or heteroaromatic rings that are linked in a polyamide bond. Brilliant reports have demonstrated DNA minor groove recognition, among which Peter Dervan’s polyamide is a classic example.\textsuperscript{13}

1.3: Introduction to Janus heterocycle
The term Janus heterocycle, was first coined by Nobel Prize laureate J.M-Lehn, after the two-faced Roman god.\textsuperscript{16-17} Generally speaking, the term Janus heterocycle is employed to describe heterocycles that are capable of forming hydrogen bonds on more than one so-called face of the heterocycle, in contrast to natural nucleobases that typically present only one face for functional hydrogen bonding (not counting the potential for simultaneous formation of Hoogsteen pairs). A Janus heterocycle therefore represents one type of artificial molecule that can be rationally designed with predefined hydrogen bonding donors and acceptors relative to regular nucleobases. For example, if one side is complementary to A (which has the character of T) the other side can complement C (which has the character of G base), this Janus heterocycle is
named Janus-TG. Typically, this Janus-TG heterocycle would not be capable of interrupting a Watson-Crick base pair because T and G are typically not associated. Taking this complementarity into account, there are two kinds of Janus heterocycles, one of which can form base pairs with itself whereas the other cannot. So far, several Janus heterocycles have been reported, including the seminal reported Janus-GA (reported by J. M. Lehn et al.), and subsequently, the Janus-GC (reported by Mascal et al.) and Janus-AT (reported by Perrin et al.)

(Figure 5).

Figure 5: Representative Janus heterocycles

Interestingly, due to special hydrogen-bonding donor and hydrogen-bonding acceptor characteristics, Janus heterocycles, particularly Janus-GC and Janus-AT that are self-complementary, can aggregate to form interesting non-covalent
supramolecules, the driving force of which being the formation of hydrogen bonding interactions as well as potential stacking interactions that favor the formation of higher order structures. For example, Fenniri et al.\textsuperscript{18,22} reported the formation of helical rosette nanotubes in water by employing Janus-GC heterocycle. Furthermore, his group also reported that the property of helical rosette nanotubes could be adjusted by changing the side chain attached to the Janus-GC heterocycle. It was shown that chirality of rosette nanotubes could be induced with the addition of chiral amino acids when the side chain of Janus-GC was a crown ether. In addition, when Janus-GC heterocycle was combined with the ribose forming Janus-GC nucleosides, Janus-GC nucleosides also have the capacity to self-aggregate forming the aforementioned rosette nanotubes but with apparent chiral properties owing to the presence of the ribose moiety.\textsuperscript{23} When the Janus-AT nucleoside\textsuperscript{23} was examined under similar conditions, instead of rosette cyclic structure, a linear sheet-like structure was observed. Although a given Janus-GC heterocycle or Janus-GC nucleoside can form the rosette nanotube superstructure, one significant limitation to the application of these kinds of rosette nanotubes is the lack of reproducibility of the identical nanotubes.\textsuperscript{21} In order to achieve pre-defined nanotubes, we have hypothesized that a certain number of Janus heterocycles should be incorporated into oligonucleotides to form a Janus heterocycle oligonucleotide, where sequence specificity can be modulated through suitable length as well as the incorporation of standard nucleosides. This is now being pursued by several laboratories.\textsuperscript{21,24}
Besides its utility in the construction of nanotubes, another potential application of a Janus heterocycle would be to act as a device by which one could recognize specific DNA sequences based on Watson-Crick base pairings. In the same seminal paper reporting Janus heterocycle, a Janus-wedge concept was also proposed by J.M Lehn.\textsuperscript{14-15} The figure below demonstrates vividly what Janus-wedge concept is (Figure 6).

\textbf{Figure 6: Janus-wedge recognizing regular DNA base pairs}

The driving force ($\Delta G$) for the Janus heterocycle insertion is likely to be due entirely to the enthalpy ($\Delta H$) that results from additional hydrogen bonds and as well as favorable stacking interactions. The question of entropy ($\Delta S$) is always more complicated as desolvation of a generally hydrophobic heterocycle into a coaxial pi-stacking arrangement would be favorable while there would be an overall loss of entropy as two molecules are associated as one. In order to begin
to test this and measure entropy and enthalpy, one needs oligonucleotides of suitable length to overcome the insertion energy barrier \( (\Delta G^\neq) \) in order to realize the thermodynamic gain \( (\Delta G) \) obtained from in the formation of additional hydrogen bonds and stacking interactions. In 2004, McLaughlin et al.\textsuperscript{25} reported prototypical triplex formation by utilizing Janus-Wedge recognizing strategy, in which CT mismatched base pairs, were recognized. Employing two mismatched DNA oligonucleotides and one peptide nucleic acid (PNA) oligomer\textsuperscript{5b, 5c, 5e, 26} as the Janus wedge, which consisted of eight Janus-CT heterocycles, the authors observed that the PNA Janus-Wedge could bind with these two oligonucleotides simultaneously to form a triplex. Although that work did not demonstrate recognition of regular double stranded DNA, it was still significant enough in that this triplex formation mechanism was totally different from those reported before in which the third strand was bound to the original DNA major groove from side rather than to insert into the original duplex, forming hydrogen bonds with both two DNA strands in the middle i.e. recognizing two strands simultaneously.
Figure 7: McLaughlin’s work on Janus-Wedge triplex formation with CT mismatch

Following this pioneering work, in 2008, it was reported by the same group\textsuperscript{27} that a Janus-Wedge triplex could not be formed at all when a Janus-wedge PNA consisting of eight Janus-AT and Janus-GC heterocycles (seven Janus-AT heterocycles and one Janus-GC heterocycle) with another two complementary oligonucleotides in which regular Watson-Crick base pairs were involved.
(a): structure of Janus-AT and Janus-GC used by McLaughlin; (b): double strand and Janus wedge used by McLaughlin

Figure 8: McLaughlin's work on regular DNA base pair

Based on this work, one conclusion is that more than eight Janus heterocycles must be incorporated into an oligonucleotide in order to realize the Janus wedge recognition.

1.4: Phosphoramidite technology used to construct oligonucleotide

As molecular biology has become broader and broader since the 1960s, in part due to its embrace of standard chemical approaches, many methods\textsuperscript{28} have been developed to chemically synthesize oligonucleotides, the most popular one being phosphoramidite technology\textsuperscript{26, 29}, which is the planned method for our
synthesis of Janus wedge oligonucleotide. The scheme below briefly demonstrates the protocol of phosphoramidite technology.\textsuperscript{30}

Figure 9: Solid phase synthesis of oligonucleotide based on phosphoramidite
1.5: Previous results unpublished in Perrin’s lab and aims of this thesis

One Janus-AT phosphoramidite shown below (the first generation of a Janus-AT phosphoramidite) in Figure 10 had been obtained in Perrin’s lab before. Unfortunately, the first generation of Janus-AT phosphoramidite failed to be incorporated into the oligonucleotide because of its extreme low solubility in regular organic solvents when anhydrous acetonitrile was employed as the choice solvent in the coupling step shown in Figure 9.

Figure 10: First generation of Janus-AT phosphoramidite

In generation of the first Janus-AT phosphoramidite, it is worthwhile to note that 2-(hydroxymethyl) cyclopentan-1-ol was employed instead of 2-deoxyribose in order to avoid the depurination-like elimination side reaction during acidic steps of the oligonucleotide synthesis. The details of the depurination will be discussed in the next chapter. Besides the use of the carbocyclic deoxyribose analog, the first problem I needed to solve was the problem of insolubility of the first generation of Janus-AT phosphoramidite. Two methods were proposed to solve this issue: 1) by using hydrophobic protecting groups to protect the exocyclic amine groups: 2) by constructing a new Janus-AT precursor phosphoramidite,
which is soluble enough for the oligonucleotide synthesis. However, because of the special structural features of Janus-AT-1 bases, there are two electron-withdrawing nitrogen atoms in each six-member ring, which tremendously decreases the nucleophilicity of two exocyclic amine groups. This conclusion is based on the fact that no suitable protecting group could be found to protect the exo-cyclic amines.

In summary, to recognize more specifically with high affinity a double stranded DNA sequence using Janus wedge, we would need to introduce several Janus bases to an oligonucleotide. In this regard, aims of the thesis are: (1) to synthesize new Janus-AT heterocycles that would facilitate the construction of soluble Janus-AT phosphoramidites; (2) to combine the carbocyclic 2-deoxyribose analog with new Janus-AT bases to afford new variants of soluble Janus-AT phosphoramidites; (3) to introduce the new Janus-AT phosphoramidites into the oligonucleotides; (4) elaboration of two new Janus-AT phosphoramidites shown in Figure 11 where $X = \text{N or CH}$.

![Figure 11: Two targets of this thesis](image-url)
Chapter 2: The synthesis of sulfone-masked Janus-AT-2 phosphoramide

2.1 Introduction

Hydrogen bonding is a fundamental concept in supramolecular chemistry that, despite being relatively weak, can account for many physical phenomena such as the folding of macromolecules (DNA, RNA and protein). For example, The study of hydrogen bonding interactions, has given rise to many reports which describe the assembly of two different monomers possessing two complementary recognition sites, (e.g. melamine and cyanuric or barbituric acid) to build up diverse architectures. Nevertheless, there are fewer examples to investigate the assembly of monomers that possess self-complementary recognition sites within a single monomer. Two remarkable examples for self-assembly are based on Janus-GC heterocycles and Janus-AT heterocycles.

As introduced previously (Chapter 1), the synthesis of Janus heterocycles may provide useful tools that can ultimately be employed to recognize regular DNA base pairs. To date, several groups have conducted preliminary research into the synthesis and self-assembly of various self-complementary heterocycles. However, none of them has provided oligomers in order to examine the formation of triplex based upon the recognition of regular Watson-Crick base pairs, i.e. A-T and C-G by utilizing a Janus heterocycle. While we cannot know exactly why, to date this has never been fully investigated, we suspect that the dearth of data in this regard may be due to the general difficulty in preparing oligonucleotides of
sufficient length and stability to provide the consequent energetics needed to overcome the nucleation barrier to form thermodynamically stable J-loop triplexes. Indeed, while several Janus heterocycles with alkyl and amino acid appendages have been reported, to date, there has been only one report of a Janus-AT heterocycle that was appended onto deoxyribose for further elaboration to the phosphoramidite. Yet this example, which by definition would constitute a Janus-AT nucleoside, lacked an exocyclic amino group (Janus-AT-3 in Figure 12). As such, this would result in one fewer H-bond to the target thymidine. Given the few examples of Janus heterocycles, it is well recognized that Janus-GC and Janus-AT heterocycles exhibit very low solubility in regular organic solvents, which would likely impede the construction of oligonucleotides bearing the same constituent nucleobases. In addition, the extensive H-bonding capabilities of these Janus heterocycles could cause aggregation and consequent chain termination during solid phase synthesis of the oligonucleotides containing a Janus AT or GC heterocycle. Our proposed solution to this critical problem of solubility, in the context of a deoxyribose analog (vide infra), represents the major topic of the following two chapters.

As we tackled the question of solubility, we carefully examined the few examples of Janus-AT heterocycles (Figure 12) that had been synthesized previously.
In addition to our goal of creating a suitable Janus-AT heterocycle that would be more soluble prior to global deprotection, we also needed to guard against the potential for an acid-catalyzed depurination-like elimination of the heterocycle from deoxyribose (Scheme 1). The deglycosylation of these heterocycles could represent a serious side reaction that is analogous to depurination, which indeed occurs facilely with purines under acidic conditions, as featured in Scheme 1.

Since acid is used during the synthesis of oligonucleotides and because we had no knowledge \textit{a-priori} as to the acid sensitivity of deoxyribosyl-Janus-AT nucleosides, we speculated that both Janus-AT-1 and Janus-AT-2 (Figure 12) might, under acidic conditions, undergo an analogous reaction that would undermine the aims of this project. The chemistry related to depurination inspires
and convinces us that depurination could be entirely obviated if a carbon atom replaces the ring oxygen atom in deoxyribose. Hence, initially we postulated that the 2-(hydroxymethyl)cyclopentan-1-ol could be employed to create a Janus-AT nucleoside analog that absolutely cannot undergo depurination. Indeed this carbocyclic analog of deoxyribose had been used previously to create standard DNA oligonucleotides and these exhibited excellent properties that included higher association equilibrium constants and increased nuclease stability.\(^{34}\) (Figure 13)

![Chemical structures](image)

**Figure 13: Comparison between carbocyclic nucleoside and regular 2-deoxyribose nucleoside**

Based on the successful synthesis of the butylated Janus-AT-1 (Figure 12),\(^{32}\) Mr. Abid Hasan in our lab indeed produced the corresponding carbocyclic Janus-AT-1 phosphoramidite (Figure 14, left). Unfortunately, but not unexpectedly, because of its insolubility in acetonitrile, which is the choice solvent for solid phase synthesis, and its very low solubility even in DMSO, DMF and NMP, oligonucleotide synthesis met with unmitigated failure.\(^{35}\) Therefore, a more soluble phosphoramidite based on Janus-AT-2 (Figure 14, right) was approached.
Figure 14: Previously synthesized Janus-AT-1 phosphoramidite (left) and the new target contemplated in this thesis (right)

Our choice for using Janus-AT-2 and our thinking as to its increased solubility was based in part on previous work where Dr. Asadi\textsuperscript{32i} had configured a pyrimidopyridine by similar approaches via a sulfone precursor, which, in the last step would be ammonolyzed (Scheme 2). Given that oligonucleotides are deprotected with concentrated ammonium hydroxide or methanolic ammonia, Janus-AT-2 seemed like a logical first-step towards creating a soluble phosphoramidite. Moreover, given the success in appending the carbocyclic analog of deoxyribose onto a pyrimidopyrimidine (Janus-AT-1) via conjugate Michael-type addition, and noting that both butylated Janus-AT-1 and Janus-AT-2 were prepared by conjugate addition of an amine, we hypothesized that the question of solubility could be addressed most readily by preparation of Janus-AT-2.\textsuperscript{32i} Hence the synthesis would commence by adapting Scheme 2.
In adopting this route, we further took into consideration the following concerns:
(1) R can be a structural motif which should be identical or similar to 2-deoxyribose such as a carbocycle shown in Figure 13; (2) based on earlier work with the butylated heterocycles in Scheme 2, we hypothesized that Janus-AT-2 could be obtained upon ammonolysis of precursor 43 in Scheme 2, following introduction within an oligonucleotide; (3) we hypothesized that precursor 43 which exhibits significantly higher solubility in organic solvents than that of Janus-AT-2, would represent an appropriate surrogate phosphoramidite that would eventually afford the desired Janus-AT nucleobase. However, in order to further increase the solubility of 43, a benzyl thioether was chosen in place of the
methylthioether, for its increased solubility as well as a distinct set of signals in the $^1$H-NMR spectrum that facilitated identification of late-stage products following purification. Taking all the concerns above into account, we proposed the second generation of sulfone-masked Janus-AT-2 phosphoramidite, which was expected to have higher solubility in organic solvent and devoid of depuration problem (Figure 15). The synthesis of this phosphoramidite is now detailed.

![Figure 15: Sulfone-masked Janus-AT-2 phosphoramidite](image)
2.2 Retrosynthetic analysis of sulfone-masked Janus-AT-2 phosphoramidite

Scheme 3: Retrosynthetic analysis of sulfone-masked Janus-AT-2 phosphoramidite

Sulfone-masked Janus-AT phosphoramidite 1 was achieved from 2 according to a standard phosphoramidite synthesis protocol (Scheme 3). After standard deprotection protocol of silyl groups, 2 was prepared from 3. Thanks to a known methodology, 3 was synthesized from 6 after several steps, which could be easily obtained from 7 and 9.
2.3: Synthesis and discussion of sulfone-masked Janus AT-2 phosphoramidite

2.3.1: Synthesis of 7

The scheme employed to synthesize 7 has been reported in the literature. Therefore, after slight modifications to the scheme below, compound 7 was successfully obtained in an overall yield of 60%. (Scheme 4)

Scheme 4: The scheme to synthesize 7

It has been found that over the synthesis of 11 and 13, the use of shelf solvents (THF and MeOH) instead of anhydrous would not decrease the yields. Therefore, shelf solvents rather than dry solvents in these two steps were employed in the synthesis although dry solvents had been employed in reported protocol. Besides this, for the transformation from 15 to 7, a new protocol that employed sealed tube instead of the regular reflux apparatus was developed, which was much easier to handle when argon environment was required in this transformation. It
is noteworthy that from 15 to 7, reaction was totally suppressed when isopropanol or tert-butanol was used as the solvent instead of water.

2.3.2: Synthesis of 8

\[
\begin{align*}
\text{O} & \text{O} \\
\text{NH} & \\
\text{2} & \\
+ & \\
\text{HOOC-CN} & \\
\text{16} & \text{(1 eq)} \quad \text{POCl}_3 (0.5 \text{ eq}), \text{cat DMF} \\
& \text{toluene, } 80^\circ\text{C} \quad 80\% \\
\text{O} & \text{O} \\
\text{N} & \\
\text{H} & \text{O} \\
\text{CN} & \\
\text{17} & \text{(1 eq)} \quad \text{acetic anhydride} \\
\text{triethyl orthoester (1 eq)} \quad \text{reflux} \\
& 62\% \\
\text{O} & \text{O} \\
\text{N} & \\
\text{H} & \\
\text{CN} & \\
\text{8} & \text{(1 eq)} \quad \text{16} \\
\end{align*}
\]

Scheme 5: The scheme to synthesize 8

Although compound 9 is commercially available, it turned out that home-made 9, prepared according to the protocol in Scheme 5, is much cheaper than that sold by Sigma-Aldrich, especially when considering the convenient filter purification method instead of chromatography. Based on Scheme 5, 40 grams of 8 can be obtained. In the synthesis of 8, it turned out two isomers would be formed as shown on TLC (Rf=0.53 for one isomer, Rf=0.73 for another isomer when 10% MeOH in EtOAc was used as the TLC solvent and the bottom isomer is the major one). Fortunately, both stereoisomers were able to be employed in the following steps to get intermediate 18 (in Scheme 6).
2.3.3: The closure of synthesis of sulfone-masked Janus-AT-2 phosphoramidite

Scheme 6: The closure of synthesis of sulfone-masked Janus-AT-2 phosphoramidite
With compounds 7 and 8 in hand, they were combined in refluxing EtOH in the presence of triethylamine. It had been observed that without triethylamine, the reaction would stop at intermediate 19 (structure proposed based on Mass spec information and E or Z conformation was not further investigated) without cyclizing to form 18 (Scheme 7). It was postulated that triethylamine might act either as a general base to catalyze the reaction or as a nucleophile catalyst to facilitate ring closure.\(^{36}\)

\[\text{Scheme 7: Comparison of cyclization in the presence and absence of triethylamine}\]

After 18 was obtained, the two hydroxyl groups were protected by TBS groups, the choice of which was based on the relative stability of this protecting group under basic conditions since ammonium hydroxide would be employed in a subsequent step. Prior to ammonolysis, the vinylogous proton of the exo-methyl
group was deprotected by potassium tert-butoxide to create the extended enolate, reacted with carbon disulfide to form 5. When 5 was transferred to a sealed tube with excess ammonium hydroxide, the mixture was heated to 100°C and allowed to react overnight to give 4 (Caution: high pressure and use blast shield).

Following successful synthesis of 4, several bases including sodium hydroxide and triethylamine were employed in an attempt to create the benzylthioether. It turned out that triethylamine, which has a much higher solubility than that of sodium hydroxide in MeOH, worked well to achieve this transformation with the added benefit that the work-up was much easier than using sodium hydroxide. When sodium hydroxide was as a base, a viscous sticky residue was obtained, which made the following purification difficult. However, when triethylamine was used, a clean homogenous solution was observed over the reaction and after the solvent was removed, a white solid was obtained, which proved very suitable for column purification. Following this step, m-CPBA was employed to oxidize the thioether to sulfone 19, which was subsequently deprotected by the use of TBAF in THF to afford 2. Because of the high solubility of TBAF in both organic solvent and aqueous layer, excessive TBAF could not be washed away from the organic layer and when the desired deprotection product was purified by column chromatography, TBAF would elute with the product. It was found that addition of lithium perchlorate to the aqueous solution, removed the TBAF from the organic layer leaving a product whose purity was shown by TLC and 1H-NMR. With the
key intermediate 2 in hand, desired phosphoramidite 1 was obtained by standard protocols.34

2.4: Conclusion

Phosphoramidite 1 was successfully synthesized after 17 steps at an overall yield of 7.25%, in which the longest linear sequence consisted of 15 steps. Whereas only one step was used to construct the T face, seven steps were used for the A face, which lowered the overall efficiency. Thus, in the future, one might consider methods toward a more efficient method to construct the A face were highly needed. As expected, a second generation of sulfone-masked Janus-AT phosphoramidite might have a higher solubility in acetonitrile, which permitted its incorporation into an oligonucleotide. After the cleavage and deprotection in ammonia, oligonucleotides have been obtained. MALDI-TOF data acquired by Dr. Eric Largy showed chemical purity of the desired oligonucleotide with the expected Janus-AT heterocycle (data not shown). Biophysical data, also acquired by Dr. Eric Largy, indicated that in the context of an oligonucleotide with only one Janus-AT nucleoside, the Janus-AT heterocycle formed strong base pairing interactions with T while less favorable interactions were formed with A and much less favorable interactions were found with G and C as would be expected. Nevertheless, high level DFT calculations by Dr. Largy (data not shown) suggested that the diaminopyridine face would likely be protonated at pH values slightly lower than 7 giving rise to undesired H-bonding interactions within the DNA helix along, with other tautomers. Therefore, we sought to design a new Janus-AT phosphoramidite that would be significantly less basic than the one
above. Hence we turned our attention to Janus-AT-1 in the context of a thioether analog that would afford the analogous Janus-AT phosphoramidite, which is the topic of next chapter.
Chapter 3: The synthesis of sulfone-masked Janus-AT-1 phosphoramidite

3.1 Introduction

As discussed before, the sulfone-masked Janus-AT-2 phosphoramidite reported in Chapter 2 has a satisfactory solubility in CH$_3$CN, which enabled incorporation into oligonucleotides. Global deprotection of the oligonucleotide in the presence of concentrated methanolic ammonia resulted in ammonolysis of the benzylsulfone group with concomitant installation of the desired exocyclic amine (See discussion in Chapter 2) to provide the desired Janus-AT-2 within an oligonucleotide. MALDI-TOF analysis performed by Dr. Largy (data not shown) confirmed the overall composition of the oligonucleotide containing the Janus-AT-2, along with its chemical homogeneity. As previously noted, biophysical experiments, including UV-melting experiments and circular dichroism studies along with computational calculations involving density functional theory (DFT) calculations (conducted by Dr. Eric Largy)$^{37}$ provided rather ambiguous melting data as to the ability of the Janus-AT-2 to recognize a complementary T (data not shown). Moreover melting data where the complementary base was a C gave two transitions$^{37}$ where normally a mismatched situation gives only one. After the data were analyzed, the conclusion, specifically drawn from a mismatched pairing with C, is that the Janus-AT-2 heterocycle in the context of an oligonucleotide might exist as two non-equivalent species suggestive of two potentially different oligonucleotides instead of the only one we pre-designed.
5’-GAGCGATGJATGTAGCCAG-3’ (oligo-1)

(Pre-designed oligonucleotide)

3’-CTCGCTACCCATCGGTC-5’ (oligo-2)

(Regular oligonucleotide used to test recognition capacity of Janus-AT)

To further analyze and explain the UV melting results, p$K_a$ calculations on the isolated Janus-AT-2 via DFT were executed by Dr. Eric Largy, the result of which showed that p$K_a$ of conjugated acid of N$_7$ (Figure 16) is 6.79.$^{38}$ According to the Henderson-Hasselbalch equation, the ratio of conjugate acid to base was 0.62 when pH of the buffer was 7.0, which meant that 38.3% of the Janus-AT-2 was protonated. A direct consequence of this protonation is a radically different hydrogen bonding donor & acceptor pattern relative to the original Janus-AT-2 heterocycle (DAD versus DDD where D means hydrogen bonding donor and A means hydrogen bonding acceptor) (Figure 17). Although proton exchange in solution is very rapid such that it is impossible to isolate protonated molecules from unprotonated ones, within the context of the double helix (from which transitions to the single-stranded state are probed by temperature dependent UV melting and CD experiments), it is possible that protonated and unprotonated forms can exist in two separate populations that may give non-time averaged yet overlapping data sets suggestive of two major forms. This would be consistent with the p$K_a$ that was calculated to be near physiological pH such that the Janus-AT-2 heterocycle within the oligonucleotide would be hemiprotonated such that a portion of the oligonucleotide, once protonated would pair differently with the
complementary strand. In order to fully explore Janus-AT heterocycles’ recognition capacity for DNA base pair, we hypothesized that there would be chemical means for reducing the basicity of the ring nitrogen and thereby eliminating the irritation of pKa at physiological pH.

1,3,7 refers to the order number of nitrogen in Janus-AT heterocycle

**Figure 16: Acid-base equilibrium of Janus-AT-2 base**

**Figure 17: Different hydrogen bond pattern between Janus-AT-2 and its protonated form**
One approach to solve this problem is to decrease the basicity of the N$_7$ in Janus-AT-2 (Figure 16). Methods to decrease the basicity of N$_7$ include: (1) Introduction of an electron-withdrawing group into the original Janus-AT-2, such as –F (2) Removal of the electron-donating group, such as exo-cyclic NH$_2$ group in Janus-AT-2. (3) Exploration of Janus-AT-1, which has two ring nitrogens in the heterocycle (a pyrimidine) instead of the pyridine in Janus-AT-2. Taking all the concerns into account, we proposed three possible candidates to displace sulfone-masked Janus-AT-2 nucleosides, which may avoid hemiprotonation at physiological pH (Figure 18).

\[ \text{Figure 18: Three benzyl-sulfone masked Janus-AT base candidates with low basicity at N}_7 \]

Considering the methodological precedent of generating the Janus-AT-1 heterocycle as elaborated in Scheme 8,\textsuperscript{39} we opted to explore the synthesis of 36 in Figure 18, which would, upon global deprotection, provide an oligonucleotide that would contain the Janus-AT-1 heterocycle. Notably this target was entirely unknown at the undertaking of this thesis.
Scheme 8: Previous methodology to construct butyl Janus-AT-1 base(R=butyl)

Hence likewise, the basicity of Janus-AT-1 was not known. Therefore, computational calculations were executed by Dr. Eric Largy as previously described and gratifyingly, it was found the new Janus-AT-1 heterocycle has no protonation problem in pH=7 buffer (pK\text{a} values were shown in Figure 19. After calculation via the method described before, the protonation form was absolutely negligible)\textsuperscript{1} which confirmed the suitability of the Janus-AT-1 heterocycle at physiological pH without undesired protonation.

Figure 19: pK\text{a} value of conjugated acid of nitrogen in Janus-AT-1

Based on these calculations, we opted to further explore a route that would enable the introduction of Janus-AT-1 into an oligonucleotide. Nevertheless, as noted in the introduction, Mr. Abid Hasan had prepared the corresponding
phosphoramidite of Janus-AT-1 yet its insolubility in MeCN prevented its entry into oligonucleotide synthesis. Hence, we sought to generate the sulfone-masked Janus-AT-1 phosphoramidite shown in Figure 20:

![Chemical Structure](image)

**Figure 20: New generation of sulfone-masked Jnaus-AT-1 phosphoramidite with less basicity**

Based on the ensemble of these considerations, we needed to revise our synthetic routes to create a new sulfone-masked Janus-AT-1 phosphoramidite as shown in Figure 20. The synthesis of this target comprises the rest of this chapter.
3.2: Retrosynthetic analysis of the sulfone-masked Janus-AT-1 phosphoramidite

Based on the standard protocol to synthesize phosphoramidites,\textsuperscript{40} 21 could be achieved by standard phosphatidylating procedures starting with 22, which could be obtained by oxidizing 23 (Scheme 9). Hence the critical and unknown challenge in this endeavor was the creation of 23. Retrosynthetic analysis of 23 provided two potential disconnections as shown in Scheme 10:
Scheme 10: Two possible disconnections for 23

After a paper evaluation of Disconnection A and Disconnection B for 23 in Scheme 10 respectively, it was concluded that Disconnection A was superior to Disconnection B based on two reasons: (1) Disconnection A was more convergent than Disconnection B, (2) for Disconnection B, the step to combine 24 with 27 to obtain 26 could be perceived as risky in that the nitrogen and sulfur in thiourea are both nucleophiles which could hypothetically react with 24 by...
displacing the methylthioether. Therefore, we opted for Disconnection A. Hence
the complete retrosynthetic analysis for 23 was shown in Scheme 11.

Scheme 11: Retrosynthetic analysis of the sulfone-masked Janus-AT-1

phosphoramide (Part B)

It was expected that 23 could be constructed from 24 and 25. Compound 24
could be obtained from 7 and 29. As described in Chapter 2, 7 could be obtained
via a robust protocol and 29 could be obtained from 9 after two steps described
in Scheme 8. As to the synthesis of 25, it was expected that 27 would react with benzylbromide.

3.3: Synthesis of the sulfone-masked Janus-AT-1 phosphoramidite

3.3.1: Synthesis of 29

Compound 29 could be obtained by following the same protocol described in Scheme 8. By condensation with carbondisulfide, 30 can be obtained over two steps at an overall yield of 77% from 9. Although 30 is a salt, by using mixture solvents, it could be converted to 29 successfully.

3.3.2: Synthesis of 24

Based on Scheme 8, it was expected that 24 would be obtained as shown in Scheme 12.
However, it turned out that only the cyclization intermediate 31 instead of desired intermediate 24 could be obtained despite the use of a variety of conditions including various nucleophilic catalysts such as PPh$_3$, DABCO. It is noteworthy that the geometry (E or Z) of intermediate 31 was not further investigated but it is preferred to be drawn as Z shown in Scheme 31 based on the steric hindrance in E configuration. That only intermediate 31 could be isolated was probably due to steric hindrance at β position of enone in 31 for the transformation from intermediate 31 to 24 and further evidence,\textsuperscript{41} regarding such steric hindrance can be found in the literature. Possibility that differences between the similar ring closure reactions described in Scheme 7 and in Scheme 12 was caused by the methyl sulfur electronic effect could not be ruled out totally. However, it is well known that the sulfur atom is relatively bulky compared to nitrogen and carbon. Hence we hypothesized that steric hindrance could be alleviated when the methylthioether was replaced by another less bulky group such as nitrogen. Thus, it was decided that the construction of right ring instead of the left ring of 23 was executed first in order to remove the bulky methylthioether before the cyclization step (Scheme 13).
Scheme 13: A modified pathway to construct 23

Therefore, after 25 was obtained by combination of thiourea with benzylbromide (Scheme 14), it was reacted with intermediate 31 in the presence of NEt₃ while refluxing. To our surprise, 23 was obtained directly without the observation of 32 (Scheme 13). This exciting result confirmed the plausibility of our aforementioned steric proposal and was critical in the success of our synthesis of the new sulfone-masked Janus-AT-1 phosphoramidite.

Reaction details that were employed to achieve 23 were shown in Scheme 14.
It is noteworthy that the methodology to combine 25 with 31 was originally conceived by us. It is thanks to this highly efficient reaction that the right ring of this sulfone-masked Janus-AT-1 heterocycle was constructed only in 2 steps while 7 steps were used for the first generation of sulfone-masked Janus-AT-2 described in Chapter 2.

**3.3.3: Synthesis of 22**

With 23 in hand, it was straightforward to employ m-CPBA (3 eq) to oxidize 23, which was shown in Scheme 15.
It is noteworthy that the sulfoxide intermediate was observed by mass spectrometry, which was finally converted to 22 as the reaction went on. When 10% MeOH in ethyl acetate was employed as the TLC solvent mobile phase, it turned out 23 had a very close Rf value with that of 22, which suggested that mass spectrometry was the better way to monitor this reaction than TLC.

3.3.4 : Closure of the sulfone-masked Janus-AT-1 phosphoramidite

With 22 in hand, it seemed straightforward to follow the standard protocol\textsuperscript{4} to make the phosphoramidite, \textit{i.e.} to protect 5'-OH using dimethoxyltrityl group in pyridine, which was followed by protecting 3'-OH using 2-cyanoethyl N,N'-diisopropylchlorophosphoramidite. Unfortunately, when 22 was added into pyridine, the color of the resulting solution would become increasingly darker, progressing from colorless to yellow and finally purple especially with the application of gentle heating. TLC of the purple solution showed a highly UV-active spot at the baseline of TLC plate even when 50% MeOH in ammonium hydroxide was used as the mobile phase, which suggested the product was especially polar. Considering the structure features of 22 (highly electrophilicity of the right ring due to two pyridine-like nitrogen atoms and one electron-withdrawing sulfone group as well as a good leaving group-----benzyl-sulfone group) and the relatively strong nucleophilicity of pyridine, it was postulated that pyridine would displace the benzyl-sulfone group, which would result in an unusual Zincke salt.\textsuperscript{42} As expected, the presence of Zincke-type salt 35 was confirmed by mass spectrometry.
Scheme 16: Possible pathway to form Zincke salt 35

Figure 21: Mass spectrometry (ESI positive mode) to confirm the Zincke salt 35

In order to further confirm the existence of Zincke salt, it was hypothesized that when the Zincke salt was dissolved in MeOH, solvent would ultimately displace the pyridinium group, in keeping with the generally well-known instability of many Zincke salts and it was expected that the following reaction would occur as shown in Scheme 17.
Scheme 17: Reaction between Zincke salt and MeOH

Gratifyingly, after leaving the Zincke salt in MeOH for several days, mass spectrometry confirmed the presence of a new compound whose peak corresponded to the desired displacement product 34 as shown below in Figure 22 (cal: 324.1308, found: 324.1310).

Figure 22: Mass spec information (ESI positive mode and negative mode)

to confirm 34

In light of these findings, we opted for alternative solvents instead of the pyridine in order to install the DMT group. Towards this end, several bases were investigated as sponges to mop up the HCl released upon reaction with DMT-Cl.
After extensive screening, it turned out that 1,4-dioxane and DIPEA were a perfect team to give the desired dimethoxytritylated protected product 33 at a yield of 51% based on recovered starting material as shown in Scheme 18.

![Scheme 18: Synthesis of 33](image)

It is noteworthy that the commercially available DIPEA was still not suitable for this reaction because of the existence of primary and secondary amine in DIPEA. Therefore, the re-distillation of DIPEA in ninhydrin was required before use.

With 33 in hand and following standard protocol, the final sulfone-masked Janus-AT-1 phosphoramidite product could be obtained at a yield of 60%.

![Scheme 19: Synthesis of 21](image)
3.4: Conclusion

As for the synthesis of sulfone-masked Janus-AT-1 phosphoramidite, a scheme that consists of 15 total steps and 11 of the longest linear steps was developed. Compared to the method to construct the first Janus-AT phosphoramidite, this method was more efficient since only three steps were used to construct the fused ring system. Gratifyingly, this new synthesized sulfone-masked Janus-AT-1 phosphoramidite was sufficiently soluble in acetonitrile, which allowed incorporation of this new Janus-AT-1 into an oligonucleotide via solid phase synthesis method. Biophysical and computational tests are currently undergoing in our laboratory.
Chapter 4: Conclusions and future work

In this thesis, two sulfone-masked Janus-AT phosphoramidites (Figure 23&24) were successfully synthesized which were both soluble enough in acetonitrile and had been introduced into oligonucleotides. After ammonolysis deprotection, the desired Janus-AT heterocycles were recovered and homogeneity was confirmed by MALDI-TOF analysis that was done by Dr. Eric Largy (data not shown).

Figure 23: Sulfonemasked Janus-AT-2 phosphoramidite

Figure 24: Sulfone-maskaed Janus-AT-1 phosphoramidite
Currently, biophysical experiments including UV-melting, CD and computational calculations based on the Janus-AT-1 are undergoing in our laboratory by Dr. Eric Largy. According to the preliminary results, as expected, UV melting analysis based on oligonucleotide involved Janus-AT-1 presented expected transitions without any protonation problem and our designed Janus-AT-1 can successfully recognize A&T instead of G&C (data not shown). This exciting result gives us confidence on the journey to explore Janus heterocycles’ capacity to recognize regular DNA base pairs.

Therefore, the following questions we would like to answer are: (1) Can we employ our Janus-AT to recognize the regular DNA base pair A/T instead of just one single base? (2) Can we execute the similar research on Janus-GC, i.e. to introduce the Janus-GC to an oligonucleotide to recognize DNA G/C base pair? (3) Finally, can we employ Janus-AT/GC to comprise a Janus-wedge to recognize any specific DNA target by forming a Janus-Wedge triplex? Research toward to answer the questions above will comprise the future work of this project.
References:


31. (a) Archer, E. A.; Goldberg, N. T.; Lynch, V.; Krische, M. J., Nanostructured Polymer Duplexes via the Covalent Casting of 1-Dimensional H-


35. Unpublished results in Perrin's laboratory.


37. Unpublished results in Perrin's laboratory.

38. The calculation was done by Dr. Eric Largy.


Appendix: Experimental details and characteristic spectrums

A1: General Materials and equipment

All the starting materials and solvents were purchased from commercially available sources. Deuterated solvents for H\textsuperscript{1}-NMR were purchased from Cambridge Isotope Laboratories, Inc. Further purification of some solvents for some reactions was executed unless otherwise noted. \textsuperscript{1}H-NMR spectra were recorded on a Bruker AV-300 spectrometer and calibrated to the residual protonated solvent at chemical shift 2.50 ppm for deuterated DMSO and 7.27 ppm for deuterated chloroform CDCl\textsubscript{3}. ESI and HRMS mass spectrometer information were obtained by UBC Mass Spectrometry facility. All the new compounds were confirmed by HRMS besides the low resolution ESI.

A2: Synthesis protocols and characteristic evidence for all compounds

Synthesis of 11:

\[
\begin{align*}
\text{HN}_2O & \quad \text{(Boc)}_2O, \text{DMAP, THF} \\
\text{10} & \quad \text{BocN}_2O, \text{DMAP, THF} \\
\text{11} & \quad \text{HN}_2O
\end{align*}
\]

\((1R)-(\text{2.00g, 18.3mmol}), \text{Di-tert-butyld} \text{dicarbonate (8.00g, 36.7mmol, 2eq) and DMAP (2.24g, 18.3mmol, 1eq) were dissolved in THF (60 mL). The mixture was stirred at room temperature until the reaction was complete, as determined by TLC (20% EtOAc in Petro ether, desired product Rf=0.50). The solvent was evaporated and the product was purified by flash column chromatography (Petroleum ether: EtOAc = 8:1) to yield}


the product as a white powder (3.8g, in near quantitative yield). $^1$H-NMR (300 MHz, CDCl$_3$) 1.49(s, 9H), 2.15(d, J=8.5Hz, 1H), 2.35(d, J=8.5Hz, 1H), 3.38(m, 1H), 4.94(m, 1H), 6.64(m, 1H), 6.89(m, 1H). ESI-MS: 232.4 [M+Na]$^+$. 

$^1$H-NMR (CDCl$_3$) and ESI of compound 11
Synthesis of 12:

Starting material (8g, 38.24mmol) was dissolved in dichloromethane (240 mL). 77% effective mCPBA (20 g, 2.5 eq.) was then added and the mixture was stirred at room temperature until the reaction was complete (TLC monitoring: 20% EtOAc in Petroleum ether, desired product Rf=0.49). The solution was washed by saturated aqueous Na₂CO₃ (three times 100 mL), saturated NaHCO₃ solution (100 mL) and water (200 mL). The organic phase was dried over anhydrous MgSO₄, filtered and concentrated under reduced pressure. The residue was purified by flash column chromatography (Petroleum ether: EtOAc = 8:1) to provide the epoxide as a white solid (7.1g, 83%) ¹H-NMR (300 MHz, CDCl₃): 1.51(s, 9H), 1.60(d, J=10.5Hz, 1H), 1.78(d, J=10.5Hz, 1H), 3.05(m, 1H), 3.59(m, 1H), 3.76(m, 1H), 4.60(m, 1H). ESI-MS: 225.8 [M+Na]⁺.
H¹-NMR(CDCl₃) and ESI of compound 12

Synthesis of 13:

Starting material (6.6g, 29.32mmol) was dissolved in methanol (250mL). Sodium borohydride (5.56g, 5eq) was added over several portions at 0°C. The mixture was stirred at room temperature for 20min, and then neutralized by 10% acetic acid in methanol by carefully monitoring the pH. The reaction was monitored by
TLC (Hexane: EtOAc=4:1, desired product Rf=0.22). The solvent was removed under reduced pressure and saturated NH₄Cl (100mL) was added to the residue, which was then extracted with chloroform (four times 100mL). The combined organic phases were dried over anhydrous MgSO₄, filtered and dried under reduced pressure, yielding product as a white solid (6.7g, near quantitative) ¹H-NMR (300MHz, CDCl₃) 1.42(m, 1H), 1.46 (s, 9H), 1.89( br, 1H), 2.12(m, 1H), 2.40(m,1H), 3.43(s, 2H), 3.65(m, 1H), 3.90(m, 1H), 4.27(m,1H), 5.83(br, 1H). ESI-MS: 252.3 [M+Na]⁺.

**H¹-NMR (CDCl₃) and ESI of compound 13**
Synthesis of 14:

Starting material (6.53g, 28.6mmol), imidazole (4.3g, 2.2eq) and tert-butyldimethylsilylchloride (9.5g, 2.2eq) were dissolved into DMF (60 mL). The clear solution was stirred for 3 h. The reaction was monitored by TLC (20% EtOAc in hexane, desired product Rf=0.65). Ethyl acetate (150 mL) was added to the mixture along with saturated aqueous Na₂CO₃ solution (twice 100mL) and brine (4 times 100mL) were used to wash the solution. The organic phase was dried over anhydrous MgSO₄, filtered and concentrated under reduced pressure. The crude product was purified by flash column chromatography (5% EtOAc in Petroleum ether) to yield product as a colorless oil (9.4g, 96%). ¹H-NMR (300MHz, CDCl₃) 0.11(s, 3H), 0.12(s, 3H), 0.91(s, 9H), 1.38(m, 1H), 1.41(s, 9H), 2.10(m, 1H), 2.35(m, 1H), 3.32(m, 1H), 3.36(m, 1H), 3.61(m, 1H), 3.87(m, 1H), 4.22(m, 1H), 5.67(br, 1H). ESI-MS: 366.4 [M+Na]⁺.
**Synthesis of 15:**

Starting material (8.55g, 24.5mmol) was dissolved in dry toluene (180mL) under argon. Red-Al (15 mL of a 65% solution in toluene, 2 eq) was then added drop wise via an oven-dried syringe, at 0°C. After 1h, a white solid precipitated. A
saturated aqueous NH₄Cl solution (50mL) was added and the mixture was stirred for 10 min. Ethyl acetate (100 mL) was added to the mixture and brine (4 times 50mL) was employed to wash the organic phase. Ethyl acetate (500 mL) was used to wash the aqueous phase. The combined organic phase was dried over anhydrous MgSO₄, filtered and the solvent was removed under reduced pressure, yielding 5 as a white powder (4.1g, 71%). The Rf value of the product is 0.89 when TLC solvent is MeOH:EtOAc=1:5. ¹H-NMR (300MHz, CDCl₃) 1.15(m, 1H), 1.42(s, 9H), 1.59(m, OH), 1.61(m,OH), 1.85(m, 1H), 2.04(m, 2H), 2.30(m, 1H), 3.59(m, 1H), 3.80(m, 1H), 4.24(m, 1H), 0.89(br, 1H). ESI-MS: 254.4 [M+Na]⁺.
H\textsuperscript{1}-NMR (CDCl\textsubscript{3}) and ESI of compound 15

Synthesis of 7:

![Chemical structure diagram]

Starting material (1.0g, 4.33mmol) was dissolved in water (24 mL) in a teflon-sealed tube and nitrogen was bubbled in the water for 20 min. The tube was then sealed and heated to 100°C overnight. The mixture was frozen and the water was removed by lyophilization over 2 days yielding 6 as an off-white solid (0.57 g, 100%). \textsuperscript{1}H-NMR (300MHz, DMSO-d\textsubscript{6}) 0.95(m, 1H), 1.45(m,1H), 1.62(m, 1H), 1.78(m, 1H), 1.97(m, 1H), 3.25-3.40(m, 4H), 3.87(m, 1H). HRMS (calcd). for
C₆H₁₄NO₂: 132.1022, found: 132.1025.

H¹-NMR (d⁶-DMSO) and ESI of compound 7

Synthesis of 9:

To a flame-dried 50mL round bottom flask 4.25g (50mmol) cyanoacetic acid and 4.9g (50mmol) urethane were added. The flask was flushed with Argon gas, using a balloon three times. To the flask 10mL dry toluene, 2.5mL POCl₃ (26mmol) and 1.2mL dry DMF were added. A clean solution was obtained after stirring for 5mins at room temperature. The flask was then equipped with a
condenser under N$_2$ flow and the solution was heated to 100°C until a white solid
was formed. The mixture was cooled down to room temperature and poured into
ice. The resulting suspension was filtered by a fritted funnel. The filtrate was
collected and dried on the high vacuum. The resulting white solid was pure as
evidenced by TLC (50% EtOAc in hexane, Rf=0.70) and H-NMR at a yield of
88% and directly used for the next step without further purification. $^1$H-NMR
(300MHz, CDCl$_3$) 1.33(t, 3H), 4.05(s, 2H), 4.25(q, 2H), 7.70(br, 1H). ESI-MS:
179.3 [M+Na]$^+$. 

![NMR spectrum image]
Synthesis of 8:

\[
\begin{align*}
&\text{9} \xrightarrow{\text{triethyl orthoacetate}} \text{8} \\
\text{HN}O &\quad \text{HN}O
\end{align*}
\]

To a flame-dried 50mL round bottom flask were added 3.12g (20mmol) starting material, 3.7mL (20mmol) triethyl orthoacetate and 8mL acetic anhydride were added and the flask was equipped with a condenser. The mixture were completely dissolved at 85°C and heated to 100°C for 1h. Then the mixture was cooled down using a water bath, which resulted in the formation of white crystals. Crystals were collected by filtration and washed using petroleum ether and diethyl ether very carefully and around 2.8g white product was isolated at a yield of 62%(Rf=0.56 of the major isomer when TLC solvent is 10% MeOH in EtOAc).\(^1\)H-NMR (300MHz, CDCl\(_3\)) 1.31(t, 3H), 1.56(t, 3H), 2.50(s, 3H), 4.25(q, 2H), 4.37(q, 2H), 9.07(br, 1H). ESI-MS: 249.4 [M+Na].
H¹-NMR (CDCl₃) and ESI of compound 8

**Synthesis of 18:**

To a 25mL round bottom flask 1.9g (8.4mmol) starting material, 0.86g (6.6mmol) 4-amino-2-(hydroxymethyl)cyclopentan-1-ol, 20mL absolutely dry EtOH and 2mL dry tritheyl amine were added. The mixture was heated to reflux until it was
shown there was no further conversion by TLC analysis (10% MeOH in DCM, desired product Rf=0.24) and the reaction was stopped and cooled down to room temperature and a white solid precipitated from the yellow solution, which was filtered by fritted funnel. Further purification was realized by recrystallization in dichloromethane. Product was finally isolated at a yield of 75% (1.31 g) as white powder. $^1$H-NMR (300MHz, DMSO-d$_6$) 1.65-1.9(m, 3H), 2.02(m, 1H), 2.35(m, 1H), 2.55(s, 3H), 3.55(m, 1H), 4.05(m, 1H), 4.52(m, 1H), 4.71(m, 2H), 11.85(br, 1H). HRMS (TOF) calcd. for C$_{12}$H$_{15}$N$_3$O$_4$Na: 288.0960, found:288.0964.

$^1$H-NMR (d$_6$-DMSO) and ESI of compound 18
Synthesis of 6:

To a flame dried 25mL round bottom flask 1.31g (4.94mmol) starting material, 1.86g (12.3mmol) tert-butyldimethylsilyl chloride, 0.84g (12.3mmol) imidazole and 10mL dry DMF were added. After stirring for 5mins, a yellow clear solution was obtained. After around another 3 hours, when TLC analysis (10% MeOH in DCM) showed there was no starting material left, the reaction was stopped. 50mL ethyl acetate was added to dilute the mixture and 100mL brine was used to wash the organic layer, which was then dried over anhydrous MgSO\(_4\). The solvent was removed under reduced pressure and yellow oil was obtained. Further purification was done by flash column chromatography (ethyl acetate: petroleum ether=1:2), which led to a white solid at a yield of 96% (50% EtOAc in petro ether, Rf=0.81). \(^{1}\)H-NMR (300MHz, CDCl\(_3\)) 0.03-0.08(dd, 12H), 0.86-0.91(dd, 18H), 1.72-1.83(m, 1H), 1.99-2.11(3H, m), 2.52-2.60(1H, m), 2.63(3H, s), 3.57(m, 1H), 3.68(m, 1H), 4.35(m, 1H), 4.73(m, 1H). HRMS (TOF) calcd. for
$C_{24}H_{44}N_3O_4Si_2$: 494.2870, found: 494.2870

**$^1$H-NMR ($CDCl_3$) and ESI of compound 6**

**Synthesis of 5:**

To a flame dried 100mL round bottom flask 2.0g (4.1mmol) starting material and 40mL dry THF were added which resulted in a clean colorless solution. Then 1.61g (14.35 mmol) KOt-Bu in 10mL dry THF was transferred to the solution.
aforementioned via a syringe. The mixture was stirred for 10 mins at room temperature, resulting in a clear yellow solution. Then 0.62mL CS₂ (10.25mmol) in 5mL dry THF was added dropwise, which gave a clear dark red solution. When TLC analysis showed there was no starting material left after around 2 hours, the reaction was stopped. Solvent was removed under reduced pressure, which led to an orange solid. Then 15mL water was added into the solid, which led to slurry. Following this, glacial acetic acid was employed to acidify the solution to pH=5. The yellow solid formed was filtered by fritted funnel, and the solid was washed with 10mL petroleum ether and 10mL cold diethyl ether. The resulting yellow solid was collected and dried under vacuum at a yield of 92%, which was directly used for the next step without further purification (30% EtOAc in petroleum ether, Rf=0.39). ¹H-NMR (300MHz, CDCl₃) 0.03-0.10(m, 12H), 0.85-0.95(m, 18H), 1.80-1.95(m, 2H), 2.01-2.17(m, 3H), 2.45-2.60(m, 1H), 3.60-3.67(m, 1H), 3.70-3.77(m, 1H), 4.37-4.77(m, 1H), 4.91-5.05(m, 1H), 7.04(s, 1H),
10.69 (br, 1H). HRMS (TOF) calcd. for C_{25}H_{44}N_{3}O_{4}S_{2}Si_{2}: 570.231, found: 570.2318.

H¹-NMR (CDCl₃) and ESI of compound 5

**Synthesis of 4:**

To a 48 mL thick Teflon sealed pressure tube 1.99g starting material and 30mL ammonium hydroxide solution were added. It was heated to 100 °C for 12 hours.
and subsequently cooled, which resulted in orange slurry. The slurry was filtered and an orange solid was obtained, which was dried under vacuum. TLC (30% EtOAc in Petroleum ether, Rf=0.53) and NMR analysis showed there was no impurity and the solid was directly used for the next step and reaction yield was 75%. $^1$H-NMR (300MHz, DMSO-$d_6$) 0.05-0.07(m, 12H), 0.80-0.91(s, 18H), 1.88-1.98(m, 2H), 2.27-2.41(m, 1H), 3.51-3.60(m, 2H), 4.35-4.44(m, 1H), 4.85-5.00(m, 1H), 6.45(s, 1H), 7.05(br, 1H), 8.75(br, 1H). HRMS (TOF) calcd. for C$_{25}$H$_{45}$N$_4$O$_4$SSi$_2$:553.2700, found:553.2700.
Synthesis of 3:

To a 100mL round bottom flask 1.43g (2.6mmol) starting material and 30 mL MeOH were added, which led to a clear orange solution. NaOH, 0.13g (3.25mmol), first dissolved in 2mL water was transferred to the aforementioned round bottom flask. The mixture was stirred for 5 mins at room temperature. Then 0.3mL benzyl bromide in 3mL MeOH was added dropwise. Once TLC analysis (20% EtOAc in petroleum ether, desired product Rf=0.37) showed there was no starting material left (after ~1.5 hours), the reaction was stopped. Solvent was removed under reduced pressure and 30mL water and 30mL ethyl acetate were added. The organic layer was dried over anhydrous MgSO₄. Filtered and removed the ethyl acetate under reduced pressure. Further purification was done by flash chromatography (20% EtOAc in petroleum ether). Product was obtained as a yellow foam at a yield of 82%. $^1$H-NMR (300MHz, CDCl₃) 0.05-0.10(m, 12H), 0.87-0.93(dd, 18H), 1.62(br, 1H), 1.75-1.87(m, 1H), 2.01-2.10(m, 3H), 2.50-2.62(m, 1H), 3.6-3.65(m, 1H), 3.7-3.75(m, 1H), 4.36-4.40(m, 1H), 4.42(s, 2H), 4.87-4.97(m, 1H), 6.30(s, 1H), 7.25-7.45(m, 5H), 8.28(br, 1H). HRMS (TOF)
calcd. for C$_{32}$H$_{51}$N$_4$O$_4$Si$_2$: 643.3170, found: 643.3169.

H$^1$-NMR (d$^6$-DMSO) and ESI of compound 3

Synthesis of 19:

```
3

TBSO

O

N

HN

CO

NH$_2$

N

TBSO

S

H

CHCl$_3$

mCPBA

19

TBSO

O

N

HN

CO

NH$_2$

N

TBSO

S

H

CHCl$_3$
```
To a 100mL round bottom flask 1.36g (2.12mmol) starting material, 20mL chloroform and 1.7g 77% effective m-CPBA were added. The mixture was stirred at room temperature. When TLC analysis (50% EtOAc in Petroleum ether) showed there was no more starting material left and desired product was confirmed by mass spectrometry analysis, the reaction was stopped. The solvent was removed under reduce pressure, which resulted in a yellow solid. 60mL ethyl acetate was added into the flask to dissolve the solid and saturated NaHCO₃ solution was employed to wash the organic layer until there was no more m-CPBA left as evidenced by TLC analysis (50% EtOAc in Petroleum ether, desired product Rf=0.84). Organic solvent was dried over anhydrous MgSO₄. Further purification was done by flash column chromatography (20% EtOAc in Petroleum ether) and the product was obtained at a yield of 60% as a white powder. ¹H-NMR (300MHz, CDCl₃) 0.05-0.12(m, 12H), 0.90-0.97(m, 18H), 1.75-1.85(m, 1H), 1.95-2.10(m, 3H), 2.40-2.50(m, 1H), 3.55-3.65(m, 1H), 3.70-3.82(m, 1H), 4.35-4.45(m, 1H), 4.63(s, 2H), 4.90-5.02(m, 1H), 6.00(br, 1H), 6.95(s, 1H), 7.27-7.37(m, 5H), 8.60-8.75(m, 1H), 9.18(br, 1H). ESI-MS: 675.4 [M+H⁺], 697.4 [M+Na⁺]. HRMS (TOF) calcd. for C₃₂H₅₁N₄O₆S₂Si₂: 675.3068, found: 675.3066.
**Synthesis of 2:**

To a 100mL round bottom flask 0.65g (0.96mmol) starting material, 10mL TBAF and 20mL THF were added. The mixture was stirred at room temperature for around 20 hours. When TLC analysis (50% EtOAc in Petroleum ether) showed...
there was no more starting material left, reaction was stopped. The solvent was removed under reduced pressure and 60mL ethyl acetate was added into the residue and saturated ammonium chloride solution was used to wash the organic layer. Anhydrous MgSO₄ was used to dry the organic layer and the solvent was removed under reduced pressure. Further purification was done by flash column chromatography (20% EtOAc in Petroleum ether), which resulted in a white solid at a yield of 100%. Rf value of the product is 0.21 when TLC solvent is 10% MeOH in chloroform.¹H-NMR (300MHz, DMSO-d₆) 1.70-1.90(m, 2H), 1.91-1.99(m, 2H), 2.30-2.40(m, 1H), 3.51-3.61(m, 1H), 4.02-4.10(m, 1H), 4.42-4.47(m, 1H), 4.78(s, 2H), 4.95-5.02(m, 1H), 6.85(s, 1H), 7.20-7.40(m, 5H), 8.03(br, 1H), 8.53(br, 1H), 11.87(br, 1H). ESI-MS: 447.4 [M+H]⁺. HRMS (TOF) calcd. for C₂₀H₂₃N₄O₆S:447.1338, found:447.1334.
H¹-NMR (d⁶-DMSO) and ESI of compound 2

Synthesis of 1:

To a 10mL flame-dried round bottom flask 0.123g (0.28mmol) starting material was added. 2*5mL dry pyridine was used to co-evaporate with the starting material. 0.1g dimethoxyltrityl chloride (1.1eq) was added into the flask, which was dissolved into 5mL dry pyridine. After 4 hours, the reaction was stopped and the desired 5-OH protected intermediate as a white powder was purified by flash column chromatography (1% MeOH in EtOAc and the silica gel was neutralized with triethylamine before using and Rf=0.33 when 10% MeOH in chloroform was used as the TLC solvent). After the DMT-protected intermediate was obtained, 2.5mL, dry DCM was employed to co-evaporate with the intermediate. 5mL dry
DCM was added into the round-bottom flask with the intermediate, which resulted in a colorless clear solution. 28µL (1.2eq) DIPEA and 37µLN,N-diisopropylcyanoethylchlorophosphoramidite (1.2eq) were added into the clean solution at 0 °C. When TLC analysis (60% EtOAc in petroleum ether, desired product Rf=0.50) showed the intermediate was completely consumed, the reaction was stopped. Another 20mL DCM was added and 50mL saturated NaHCO₃ solution was employed to wash the organic layer. After drying over anhydrous NaSO₄, DCM was removed under reduced pressure. Desired phosphoramidite as a white foam was obtained after flash column chromatography (50% EtOAc in petroleum ether and the silica gel was neutralized before using) at a yield of 58% over two steps. ¹H-NMR (300MHz,CDCl₃) 1.0-1.24(m, 15H), 2.10-2.25(m, 3H), 2.41-2.49(m, 1H), 2.50-2.57(m, 2H), 2.72-2.80(m, 6H), 3.20-3.25(m, 1H),3.30-3.35(m, 2H),3.80(dd, 6H), 4.60(s, 2H), 4.77(br, 1H), 5.90(s, 1H), 6.77-6.80(m, 4H), 6.90(dd, 1H), 7.30-7.37(m, 10H), 7.42-7.47(m, 2H), 8.02(br, 1H), 8.62(br, 1H).³¹P-NMR (120MHz,CDCl₃) 149.3, 149.7. HRMS (TOF) calcd. for C₅₀H₅₆N₆O₉SP: 947.3567, found: 947.3559.
ESI of intermediate 11

P$^{31}$-NMR and ESI of 1
Synthesis of 30:

To a 250mL round bottom flask 12g (77mmol) starting material, 10.7g (77mmol, 1eq) and 163mL regular DMF were added. After stirring for around 4 hours, the reaction mixture changed into cloudy form clear, in which 12 mL CS$_2$ was added. After another 2 hours, the reaction was stopped and 300 mL ethanol was added to the cloudy mixture, which was filtered by fritted funnel. The resulting solid was washed by 200mL diethyl ether. 18.5g 30 as a slightly yellow solid was obtained at a yield of 77% whose Rf value is 0.52 when TLC solvent is 20% MeOH in EtOAc. $^1$H-NMR ($d^6$-DMSO): 1.23(t, 3H), 4.02(q, 2H), 14.90(s, 1H)
Synthesis of 29:

To a 500mL round bottom flask 18.5g starting material (60mmol), 140mL water, 60mL acetonitrile were added, which resulted in a yellow solution. By a dropping funnel, 8mL MeI in 30mL acetonitrile was added into the solution over 1 hour. After the addition was finished, the solution was heated to 80 °C. The reaction was monitored by TLC carefully (50% EtOAc in Petroleum ether, desired product Rf=0.44). When all the starting material was consumed, the reaction was stopped. Most of the volatile was removed under reduce pressure and the residue was extracted by ethyl acetate. The organic layer was dried and ethyl acetate was removed under reduced pressure. The residue was purified by flash column chromatography. (25% EtOAc in Petroleum ether) 9.4g slightly yellow solid was obtained at an isolated yield of 60%. Rf value of the product is 0.65 when the solvent is 50% ethyl acetate in petroleum ether. H^1-NMR (CDCl₃): 1.25(t, 3H), 2.60(s, 3H), 2.80(s, 3H), 4.25(q, 2H), ESI (M+Na⁺): 283.3.
$^1$H-NMR (CDCl$_3$) and ESI of 29
Synthesis of 31:

To a 250 mL round bottom flask 0.8g (6.1mmol) carbocyclic amine, 1.77g (7.23mmol, 1.2eq) another starting material were added. Around 150 mL isopropanol was added into the flask until a clean solution was obtained. The reaction was heated to 80 °C and refluxed over night for around 20 hours. When TLC analysis showed the carbocyclic amine was totally consumed, the reaction was stopped and all the volatiles were removed under high vacuum. Via dry loading method, the product was purified and 1.15g product as a sticky brown foam was obtained at an isolable yield of 55% (gradient solvent from pure EtOAc to 10% MeOH in EtOAc). Rf value of the product is 0.36 when 10% MeOH in ethyl acetate was used as the solvent. H1-NMR (d6-DMSO): 1.23(t, 3H), 1.33(m, 1H), 1.75(m, 1H), 1.80(m, 3H), 1.95(s, 1H), 2.25(m, 1H), 2.60(s, 3H), 3.40(m, 2H), 3.90(br, 1H), 4.20(q, 2H), 4.50(m, 1H), 4.60(m, 1H), 4.70(d, 1H), 9.80(s, 1H), 10.55(br, 1H), ESI(M+H+) 344.3; (M+Na+) 366.3 found.
Synthesis of 25:

To a 100mL round bottom flask 6g thiourea, 50mL ethanol and 15mL 37% HCl were added. After stirring for several minutes, a clear solution was obtained. The reaction continued for another 30 mins before all the volatiles were removed under reduced pressure. To the residue 50mL MeOH and 11 mL benzyl bromide
(1.2eq) were added. When TLC analysis (50% MeOH in EtOAc was used as TLC solvent and Rf=0.19) showed there was no more change, the reaction was stopped and the solvent was removed under reduced pressure, which resulted in a white solid. The product was purified by recrystallization in ethyl acetate. 14.8g white solid was obtained at a yield of 93%. H$^1$-NMR (d$^6$-DMSO): 4.5(s, 2H), 7.3-7.5(m, 5H), 9.2(d, 4H), ESI(M+H$^+$):167.3
Synthesis of 23:

To a 100mL round bottom flask 1.1g (3.2mmol) starting material, 0.8g (3.84mmol, 1.2 eq) benzyl protected thiourea and 60mL isopropanol were added, which resulted in a yellow solution. Following this, 2mL regular triethylamine was added into the solution. The reaction was heated to 100 °C and the refluxed over night. When TLC analysis showed the starting material was totally consumed, reaction was stopped and all the volatiles was removed under high vacuum. Via dry loading method, the product was purified by flash column chromatography (gradient solvent from pure solvent to 10% MeOH in EtOAc). 0.81g white solid product was got at an isolable yield of 61%. Rf value of product is 0.61 when 10% MeOH in ethyl acetate was used as a TLC solvent. H\(^1\)-NMR(d\(^6\)-DMSO): 1.60(m, 2H), 1.80(m, 2H), 2.45(m, 1H), 3.55(d, 1H), 4.10(br, 1H), 4.35(s, 2H),

H\(^1\)-NMR(d\(^6\)-DMSO) and ESI of 25

![Diagram of synthesis](image-url)
4.50(br,1H), 4.60(br, 1H), 5,75(br, 1H), 7.3-7.5(m, 5H), 8.25(s, 1H), 8.40(s, 1H), 11.5(s, 1H), ESI(M+Na\(^+\)): 438.2 found, (M+H\(^+\)): 416.2 found. (M-H\(^+\)): 414.2 found

\( ^1\)H-NMR (d\(^6\)-DMSO) and ESI of 23
Synthesis of 22:

To a 500mL round bottom flask 0.4g (0.96mmol) starting material 300mL chloroform and 50mL isopropanol were added, which resulted in a clear solution. 1g m-CPBA (77% effective, 4.48mmol, 4.7eq) was added into the solution. Mass spectrometry was used to monitor the reaction until the starting material was no longer detected. After around 20 hours, all the volatiles were evaporated under reduced pressure. Via dry loading method, the residue was purified using flash column chromatography to obtain product 0.31g at a yield of 72% (gradient solvent from pure EtOAc to 10% MeOH in EtOAc) as a white solid. The Rf value of product is 0.67 when 10% MeOH in ethyl acetate was used as a TLC solvent. \( ^1H \)-NMR(\(d^6\)-DMSO): 1.60(m, 1H), 1.85(m, 3H), 2.40(m, 1H), 3.52(br, 1H), 4.10(br, 1H), 4.45(br, 1H), 4.55(br, 1H), 5.35(s, 2H), 5.60(m, 1H), 7.32-7.55(m, 5H), 7.90(s, 1H), 8.60(s, 1H), 8.80(s, 1H), 11.60(s, 1H) HRMS: calcd for \(C_{19}H_{21}N_5O_6NaS\): 470.1110, found: 470.1104
H¹-NMR(d⁶-DMSO) and ESI of 22

Synthesis of 33:

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]

\[
\begin{align*}
\text{HN} & \quad \text{N} \\
\text{N} & \quad \text{O} \\
\text{NH}_2 & \quad \text{S} \\
\text{OH} & \quad \text{O}
\end{align*}
\]
To a 100mL flame dried round bottom flask with stir bar 0.3g (0.67mmol) starting material and 50mL distilled 1,4-dioxane were added, resulting in a clear solution. 0.25g (0.74mmol, 1.1eq) DMT-Cl and 0.5mL redistilled DIPEA were added into the clear solution. The reaction continued over night (20 hours) at room temperature. Although TLC analysis (50% EtOAc in Petroleum ether) showed the starting material was still not totally consumed, the reaction was interrupted. All volatile were removed under removed under high vacuum and the residue was purified via dry loading method flash column chromatography (25% EtOAc in Petroleum ether). 0.17g product as a white solid was obtained and 0.1g starting material was recovered, which was recycled to synthesize the product. The yield was 51% (b.r.s.m). The Rf value of product is 0.75 when 25% petroleum ether in ethyl acetate was as the TLC solvent. $^1$H-NMR($d^6$-DMSO): 1.25(m, 1H), 1.70(m, 1H), 2.10(br, 3H), 2.80(m, 1H), 3.25(m, 1H), 3.75(s, 6H), 4.20( m, 1H), 4.70( d, 1H), 4.90(s, 2H), 5.75(br, 1H), 6.75-6.95(m, 5H), 7.20-7.50(m, 13H), 8.60( s, 1H), 9.0(s, 1H), 11.80(s, 1H).HRMS: calcd for C$_{40}$H$_{38}$N$_5$O$_8$S: 748.2441, found: 748.2440
H$^1$-NMR(d$^6$-DMSO) and ESI of 33
Synthesis of 21:

To a 100mL round bottom flask with the starting material 0.2g (0.27mmol) and stir bar, 80mL dry DCM was added, resulting in a clear solution. 92 µLDIPEA (0.53mmol, 2eq) was added into the clean solution, which was followed by the addition of oven-dried 4Å molecular sieves. The round bottom flask was then placed into an acetonitrile dry-ice bath, in which 90 µL(0.40mmol, 1.5eq) 2-cyanoethyl N,N-diisopropylchlorophosphoramidite was added. The acetonitrile dry ice bath was removed and the reaction continued for another 3 hours, which was monitored by TLC analysis (20% EtOAc in petroleum ether). When TLC showed all the starting material was consumed, the reaction was stopped. The volatiles were removed under reduced pressure and the residue was purified by flash column chromatography (60% yield) (20% EtOAc in petroleum ether). The Rf value of the product as a colorless oil is 0.8 when 75% ethyl acetate in petroleum ether was used as the solvent for TLC. HRMS: (M-H\(^+\)): calcd for C\(_{49}\)H\(_{55}\)N\(_7\)O\(_9\)SP 948.3520, found: 948.3529. \(^{31}\)P-NMR: 148.5, 148.7
$^{31}$P-NMR and ESI of 21