CHEMICAL AND ENZYMATIC SYNTHESIS OF DNA AND RNA CONTAINING SELENIUM FOR X-RAY CRYSTALLOGRAPHY USING MULTIWAVELENGTH ANOMALOUS DISPERSION

by

NICOLAS CARRASCO

A dissertation submitted to the Graduate Faculty in Biochemistry in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2005

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

This manuscript has been read and accepted for the Graduate Faculty in Biochemistry in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Sep. 30, 2004
Date

Zhen Huang (Zhen Huang)
Chair of Examining Committee

December 2, 2004
Date

Lesley Davenport
Executive Officer

L. Davenport
Richard S. Magliozzo
William D. Hein
Richard V. Swanson
Supervisory Committee

The City University of New York

Abstract

CHEMICAL AND ENZYMATIC SYNTHESIS OF DNA AND RNA CONTAINING SELENIUM FOR X-RAY CRYSTALLOGRAPHY USING MULTIWAVELENGTH ANOMALOUS DISPERSION

by

Nicolas Carrasco

Adviser: Professor Zhen Huang

A novel approach has been developed to synthesize DNA and RNA oligonucleotides containing selenium for structure determination by X-ray crystallography using Multiwavelength Anomalous Diffraction (MAD). The impetus for the project has been based on a challenging aspect in the field of nucleic acid X-ray crystallography: the elucidation of the three-dimensional structures of DNA and RNA molecules is hampered by the difficulty to create suitable heavy atom derivatives for crystallographic phasing. To address this problem, a chemical and an enzymatic approach have been investigated to introduce the selenium functionality into both DNA and RNA covalently. In the chemical approach, the solid phase synthesis via the phosphoramidite chemistry was utilized to synthesize relatively short oligonucleotides. To this end, synthetic procedures were designed to incorporate the selenium functionality at either the 5'- or

2'- positions of thymidine and uridine phosphoramidites. These phosphoramidite derivatives were then utilized to introduce selenium at specific sites in selected DNA and RNA oligonucleotides. In all cases, the selenium functional group is presented in the form of methyl selenide in order to confer superior stabilities to the corresponding products. Similarly, insight into the enzymatic synthesis of phosphoroselenoate DNA and RNA has been achieved via the synthesis of nucleoside triphosphates harboring selenium at the α-phosphate, in which selenium replaces one of the non-bridging oxygen atoms. The synthesis of these nucleoside triphosphate analogs was performed via a phosphite intermediate that was oxidized with 3H-1,2-benzothiaselenol-3-one. This enzymatic strategy was of particular interest because, in contrast to the solid phase approach, it facilitated the synthesis of long DNA and RNA molecules using the Klenow DNA polymerase and the T7 RNA polymerase, respectively. The presence of selenium in the phosphoroselenoate DNA and RNA were confirmed using selenium-enhanced resistance to enzymatic digestion. The data demonstrated the practicability of the chemical and enzymatic approaches to introduce the selenium functionality into DNA and RNA as a scatter center for MAD phasing in nucleic acid

X-ray crystallography. The development and maturity of this selenium derivatization method should significantly advance the field of nucleic acid structure analysis, biomedine, and genomic research.

Dedication

It is with the deepest love that I dedicate this work to both my mother, Ercilia Carrasco and my friend Alexandra Quezada, for their love, caring, understanding, and encouragement that they have always given me throughout my education.

Acknowledgements

I would like to thank with great appreciation those talented and dedicated individuals who guided, supported, and encouraged me throughout the completion of this dissertation. In particular, I would like to express my deepest thanks to my advisor, Professor Zhen Huang, for his dedication, inspiration, and valuable insights that enriched my educational experience and knowledge.

In addition, I would like to thank Professor William Hersh at Queens College for providing me with the facility to conduct experiments in his laboratory significant for the completion of my thesis.

I would also like to thank all the financial institutions that funded my research; those include the LS-AMP Program, the Gates and Melinda Foundation, and the AGEP Program at The Graduate Center.

Table of Contents Page

Abbreviations

A	adenine
Ac	acetyl
Ac_2O	acetic anhydride
ATP	adenosine triphosphate
ATPαSe	ATP with selenium at the α-phosphate group
BTSe	3H-1,2-benzothiaselenol-3-one
C	cytosine
DMF	N,N-dimethylformamide
DMTr	4,4'-dimethoxytrityl
DNA	deoxyribonucleic acid
eq.	equivalents
G	guanine
HPLC	high pressure liquid chromatography
HOAc	acetic acid
HRMS	high resolution mass spectrometry
Me	methyl
mRNA	messenger ribonucleic acid
MS	mass spectroscopy
Ms	methanesulfonyl
MsCl	methanesulfonyl chloride
NMR	nuclear magnetic resonance
NTPs	nucleoside triphosphates
RNA	ribonucleic acid
Se	selenium
T	thymine
TBAF	tetrabutylammonium fluoride
TBDMS	*tert*-butyldimethylsilyl
TEA	triethylamine
THF	tetrahydrofuran
TLC	thin layer chromatography
TTP	thymidine triphosphate
TTPαSe	TTP with selenium at the α-phosphate group
U	uracil
UTP	uridine triphosphate
UTPαSe	UTP with selenium at the α-phosphate group

Chapter 1

BACKGROUND AND INTRODUCTION

1.1 Biological Roles of Nucleic Acids

There are two types of nucleic acids known in nature today: deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). Nucleic acids play a variety of important roles in biological systems. The primary roles of DNA and RNA are found in the pathways leading to the flow of genetic information in cells (**Figure 1**). The flow of genetic information has long been known as the "Central Dogma" of molecular cell biology. Accordingly, DNA directs its own replication during cell division. The genetic information stored in DNA is then transcribed into RNA in the form of messenger RNA (mRNA) and subsequently translated into proteins. Under this scope, ribosomal RNA (rRNA) fulfills a structural role and transfer RNA (tRNA) an adapter role in the conversion of message to protein.

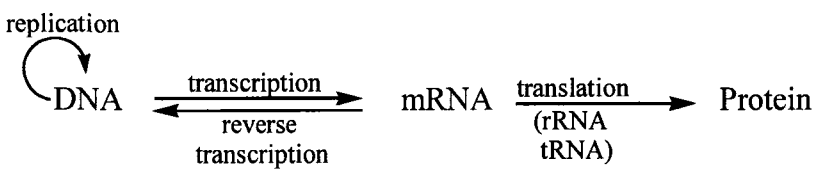


Figure 1. The flow of genetic information in cells.

In contrast to DNA, RNA has a much larger variety of biological functions. Hence this introduction will primarily focus on RNA. In addition to the secondary role that RNA plays during transcription and translation, it is the carrier of genetic information in some viruses. Furthermore, the discovery of catalytic RNAs (ribozymes) by Thomas Cech in 1982 proved that RNA could catalyze biochemical reactions (*1*). To date, several examples of naturally existing ribozymes include hammerhead, hairpin, hepatitis delta virus (HDV), *Neurospora* Varkud Satellite (VS), group I and II introns, ribonuclease P (RNAase P), and the ribosomal RNA. Additionally, recent biochemical analyses strongly suggest that the RNA component of the spliceosome is a ribozyme. These ribozymes catalyze the cleavage and ligation of the RNA backbone by tranesterification or hydrolysis of the phosphate groups by 10^5- to 10^{11}-fold over uncatalyzed reactions (*2-4*).

The discovery of catalytic RNAs led scientists to propose that RNA possesses the intrinsic capacity of the information storage function of DNA. This realization that RNA combines features of both proteins and DNA has constituted the basis for the formulation of the "RNA world" hypothesis for the origin of life (*5, 6*). This

hypothesis has received much attention in recent years after the discovery by *in vitro* selection of a ribozyme that utilizes nucleoside triphosphates and coding information of an RNA template to catalyze the polymerization of RNAs up to 14 nucleotides in a sequence-independent manner and with high fidelity (*7*). Further support to the RNA world hypothesis has emerged from the *in vitro* selection of ribozymes that have the ability to catalyze a variety of unique biochemical reactions, such as peptide bond formation (*8-11*), carbon-carbon bond formation (*12*), and RNA-protein ligation via the formation of a phosphoamide bond (*13*).

More recently, the biological function of RNA has increased significantly after the discovery of an expanding universe of noncoding RNAs (ncRNAs) in organism as diverse as the worm *Caenorhabditis elegans* and humans *(14-17)*. Noncoding RNAs have been found to play important roles in a great variety of processes in cells, including transcriptional regulation, chromosomal replication, RNA processing and modification, mRNA stability and translation, and protein degradation and translocation (*18*). These ncRNAs range in size from 21 to 25 nucleotides for the family of microRNAs (miRNAs) that modulate developmental processes in *C. elegans,*

Drosophila, and mammals; 100 to 200 nucleotides for small RNAs (sRNAs) commonly found as translational regulators in bacterial cells; and to >10,000 nucleotides for RNAs involved in gene silencing in higher eukaryotes. The recent discovery of these RNA functions adds up to the well-known roles of RNA transcriptional regulation by attenuation and *rho*-dependent transcription termination (*19*).

1.2 DNA Helical Structures

DNA and RNA are linear polymers of nucleotides linked together by the phosphate group at the 3' and 5' positions (**Figure 2**). The chemical structures of DNA and RNA differ from one another in terms of some characteristic structural components. In terms of nucleotide units, in DNA the ribose sugar has a hydrogen atom at the 2' position instead of the normal hydroxyl group; hence the name of DNA (see **Figure 5**).

The structure of DNA was elucidated by James Watson and Francis Crick in 1953 based on X-ray diffraction studies of DNA fibers in solution (*20*). The structure of DNA is depicted as a double helix of two complementary strands running in different directions. Accordingly, the base-pair interactions on the double strands are

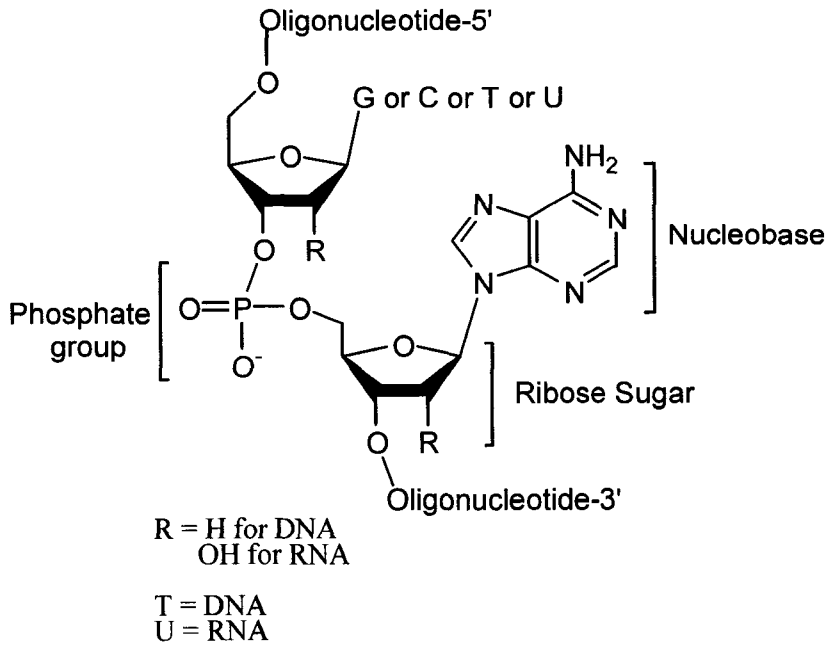


Figure 2. Schematic representation of the basic unit of DNA or RNA.

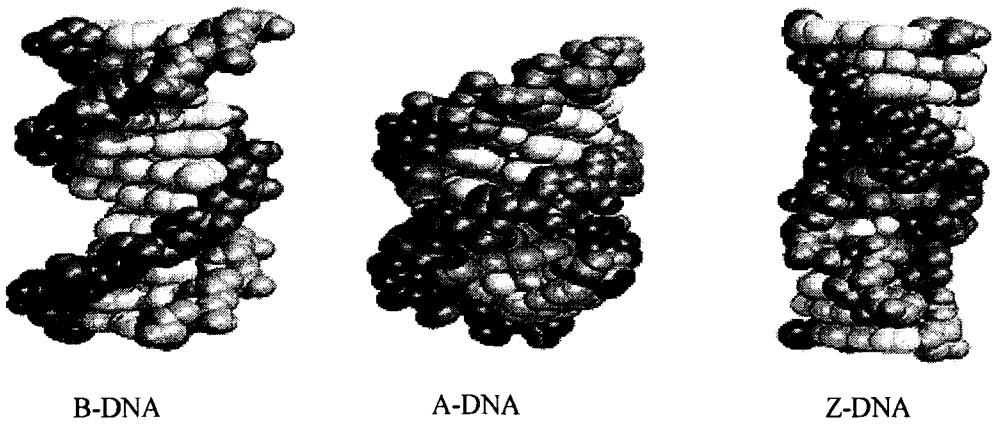


Figure 3. Space-filling models of the conformations of B-DNA, A-DNA and Z-DNA (*21*).

formed almost exclusively between G and C, and T and A. Double-stranded DNA is a conformationally variable molecule. The DNA double helix adopts three different conformations depending on the counterion and the relative humidity of the solution (**Figure 3**). These conformations have been elucidated using X-ray diffraction. The B-DNA is know as the native conformation because its X-ray diffraction pattern resembles that of the DNA found in cells. This conformation is also known as the Watson-Crick structure because it was the form that they originally proposed, and has a relative humidity of about 92% and a $C_{2'}$-*endo* sugar puckering conformation (**Figure 4**). The two strands are also wound about an axis with a right-handed twist.

When the relative humidity of B-DNA is reduced to about 75%, it adopts a reversible conformation change to A-DNA. A-DNA forms a wider and flatter right-handed helix than B-DNA. As opposed to B-DNA, the deoxyribose sugar in A-DNA adopts a $C_{3'}$-*endo* conformation (**Figure 4**). The other type of DNA conformation is Z-DNA. Z-DNA adopts left-handed double helix (**Figure 3**). Z-DNA has a dinucleotide repeating unit, as opposed to a single nucleotide, as in the cases of B-DNA and A-DNA. This repeating unit is a d(XpYp), where X represent a pyrimidine residue and Y is a purine residue.

The purine nucleotides are all C_3'-*endo* while the pyrimidine nucleotides are C_2'-*endo*. Z-DNA is stabilized under high salt concentration. The biological function of A-DNA and Z-DNA remains largely unknown.

1.3 RNA Structure Determination

RNA does not form double helical secondary structures to the same degree as DNA. It cannot form the B-DNA like conformation because of the gauche effect involving the 2'-hydroxyl group.

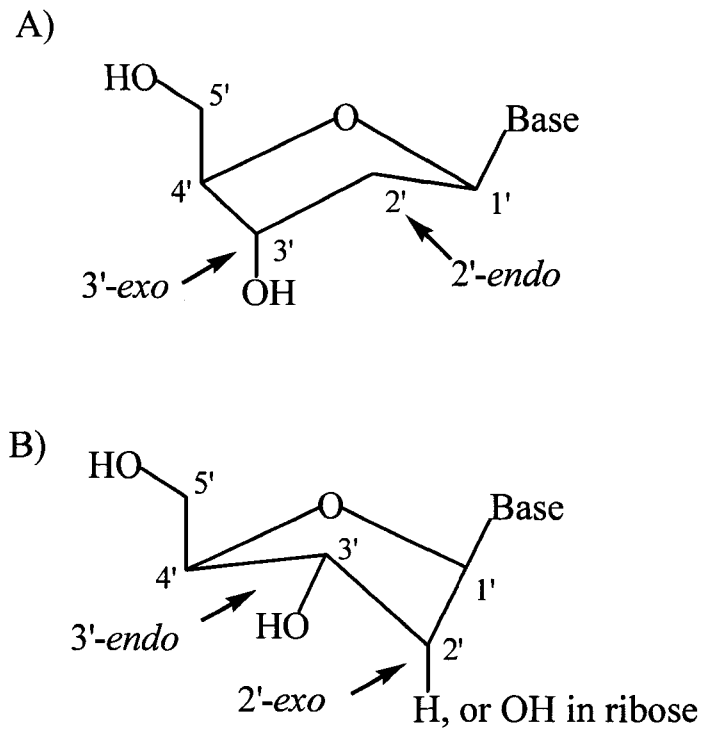


Figure 4. Sugar puckering of the furanose ring. A) The C_2'-*endo* conformation is found in B-DNA; B) The C_3'-*endo* is adopted by A-DNA and RNA.

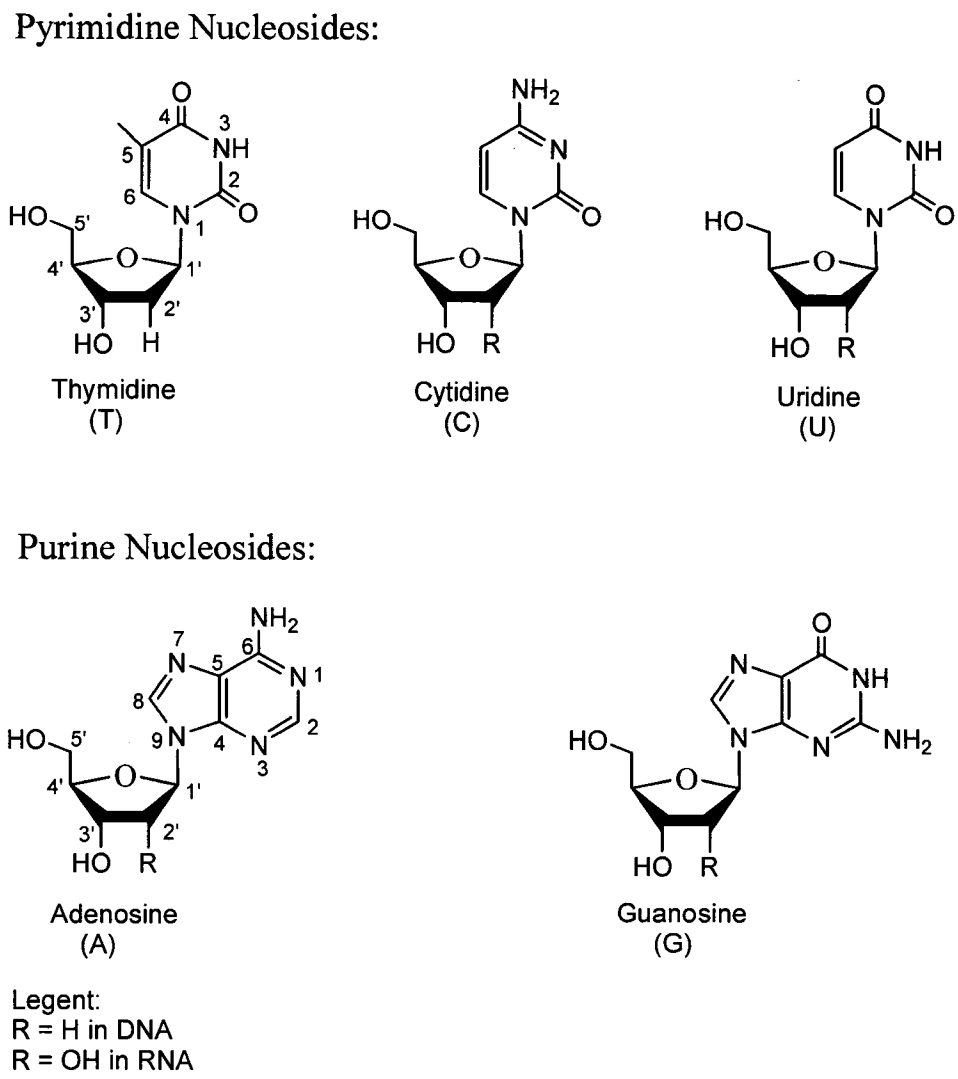


Figure 5. Chemical structures of the basic units composing DNA and RNA. DNA is composed of A, G, T, and C; U replaces T in RNA. In DNA, the purine bases A and G form base pair with the pyrimidine bases T and C, respectively.

Rather, it usually adopts a conformation resembling A-DNA. This conformation is known as A-RNA, and as in the case of A-DNA, the sugar pucker has a $C_{3'}$-*endo* conformation (**Figure 4**). Many RNA

molecules, for example, tRNA and rRNA, contain complementary sequences that form double helical stems. These stem regions are separated from each other by single-stranded segments or loops (**Figure 6**).

A unique characteristic of RNA is its ability to form tertiary structures. The first tertiary structure of RNA was elucidated in 1974 with the X-ray crystal structure of $tRNA^{Phe}$ by Alexander Rich (*22*). The main features holding the tertiary structure are bridging interactions between far distant regions in the molecule (**Figure 6**). These tertiary interactions are mostly non-Watson--Crick associations. Following the structure of $tRNA^{Phe}$, other crystal structures of tRNAs were solved. This started the era of RNA X-ray crystallography.

Over the years, there has been a tremendous amount of efforts from the scientific community to study the structure and function relationships of RNA. Basic insights into this area of RNA research have also come from investigation of the RNA structure by comparative sequence analysis of phylogenitically distinct family members of organisms (*23-26*), site-specific mutagenesis (*27, 28*), chemical protection, photo cross-linking (*29*), and nucleotide analog interference mapping (*30, 31*). Although these approaches have

provided valuable information in elucidating specific functions of especially large RNA molecules, as in the case of rRNA and group I intron (*24, 31*), the data that can be obtained pertain mostly to secondary structural features and very limited tertiary interactions. Electron microscopy and neutron scattering have also been very useful in the study of large RNA molecules, such as the ribosome (*32, 33*).

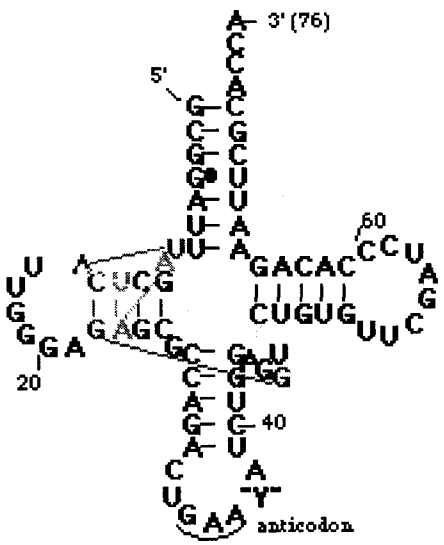


Figure 6. Secondary structure of a tRNA showing some tertiary interactions (*22*).

In the last few years, thee-dimensional (3D) electron microscopic images of the ribosome have been produced to fit a 7.5 Å resolution electron microscopic map of the 50S ribosomal subunit from

Escherichia coli (*34*). This resolution allowed the visualization of some of the proteins and nucleic acids bound to the ribosome that assist in protein synthesis. Detailed information at atomic resolution about the 3D structures of RNA can only be achieved through solution nuclear magnetic resonance (NMR) and X-ray crystallographic studies. The 3D structures of a few small RNA molecules have been elucidated by NMR to date (*35-38*) and a steady increase is expected in the next few years, as higher field NMR instruments are built. To date, X-ray crystallography has become a widely used and powerful technique to determine the 3D structure of small and large RNAs. Yet, full appreciation of the functions and mechanisms of actions of RNAs is still limited due to the lack of enough X-ray crystallographic data because of some inherent problems in X-ray crystallography. This dissertation addresses the development of new techniques for making X-ray crystallography more accessible to scientists working in the RNA and DNA fields.

1.4 RNA Crystallography

The era of RNA crystallography began in the search for the structure of tRNA in the late 1960s with the crystallization and structure

determination of yeast phenylalanine tRNA by several research groups (*39-46*). These crystal structures constituted the basis for the structural principles of not only all tRNA, but also nucleic acids in general. The data demonstrated the complexity of RNA folding, loop formation, tertiary interactions utilizing backbone-backbone, base-backbone, and base-base associations. It also illustrated the capacity of RNA to form base triplets, intercalation of base in single-strand regions, and metal-binding sites (*47, 48*).

In addition, the discovery of catalytic RNAs fueled great interest in the field of RNA structure analysis by X-ray crystallography. Indeed, the past few years have brought great advances as a number of biologically important RNA molecules (catalytic and non-catalytic) have been determined at atomic resolution and their mechanisms of action significantly elucidated; crystal structures include those of several ribozymes such as hammerhead (*49, 50*), hepatitis delta virus (*51*), hairpin (*52*), the core of the group I intron (*53*), and the 2.4 Å structure of the 50S subunit of ribosomal RNA (*54*). Moreover, the crystal structure of an in vitro-selected self-cleaving ribozyme, called the leadzyme, has been determined (*55*). This wealth of RNA crystallographic data has provided further detailed views of the

tertiary folds and motifs of these RNA molecules such as pseudoknot, kissing hairpin, triple helix, and loop-helix contact (*56*).

This period of enlightenment in structure determination of RNAs has been possible mainly to the introduction of innovative technologies in the synthesis and purification of RNAs (*57-59*), methods of RNA crystallization (*60-62*), and advances in methods of data acquisition and structure determination. The development of chemical and enzymatic methods of RNA synthesis has facilitated the preparation of RNA in great amount and good quality for structure analysis. In this respect, the automated large-scale synthesis of RNA has become routine for sequences up to about 30 nucleotides in length and for the incorporation of specific modifications. The preparation of longer RNAs containing homogeneous 5'- and -3'ends has been achieved by transcription utilizing T7 RNA polymerase and a DNA template, followed by specific ribozyme treatment to process the transcript ends to the correct length (*63, 64*). Likewise, methods for the purification and crystallization of RNA have been improved with the introduction of reverse phase HPLC, ion exchange chromatography, gel electrophoresis, and the exploitation of a variety of screening crystallization conditions (*65-67*) .

Additionally, a revolution has also occurred in the field of X-ray diffraction data collection and structure determination. To date, the data collection system consists of a powerful rotating anode generator, image plate detector systems with integrated computer, and a cryogenic device to freeze the crystals, thus increasing their diffraction lifetimes. Likewise, the increasing availability of synchrotron radiation sources has accelerated the data acquisition process; the high intensity of the X-ray bean allows data collection on very small crystals at high diffracting resolution (*68, 69*). Moreover, synchrotron radiation facilitates tunable wavelengths, which permit the resolution of very large molecules; provides phasing information by optimizing the anomalous diffraction effects of a few selected heavy atoms, and may also minimize X-ray damage to RNA crystals (*47*).

1.5 The Phase Problem in X-ray Crystallography

Despite the dramatic advances experienced in the field of RNA structure analysis, the area of RNA crystallography is still quite challenging. Two major problems often encountered are crystallization and heavy atom derivatization for phase determination.

Crystallization is a major bottleneck in X-ray crystallography because it relies on trial-and-error experiments involving the screening of a variety of different crystallization conditions. Similarly, heavy atom derivatization involves the search for a suitable scattering center by soacking and co-crystallization with metal ions or by the use of a chemical approach. Heavy atom derivatization is required to solve the phases of the diffracted X-rays measured from a crystal of a molecule whose structure is completely unknown.

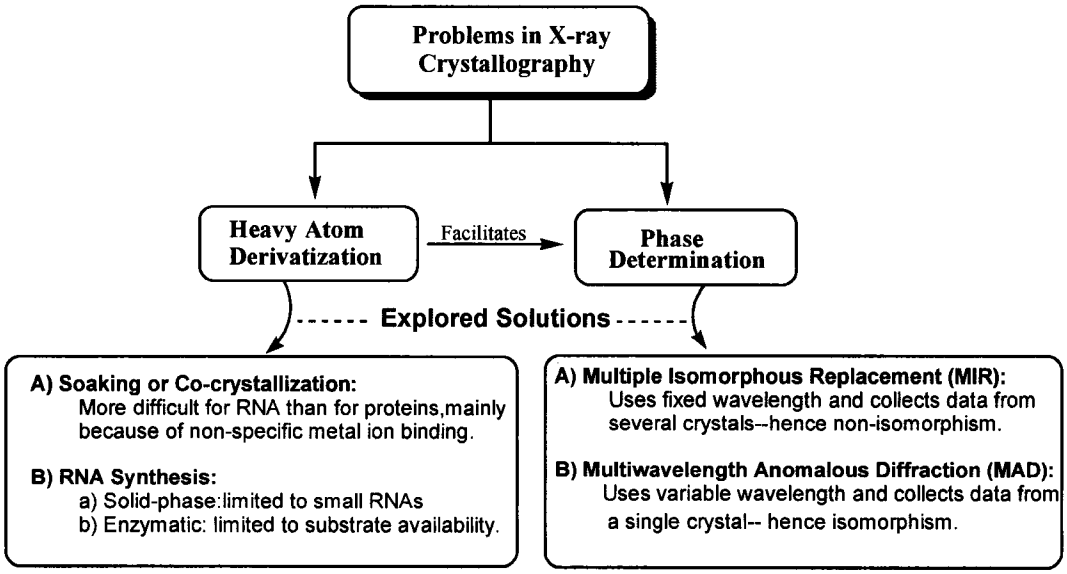


Figure 7: Outline of the problems usually encountered in the field of X-ray crystallography.

These diffracted X-rays are the Fourier transform of the electron density of the molecule; since the Fourier transform can be reversed

algebraically, an image of the molecule from the diffracted X-ray waves can be built. Unfortunately, only the intensities of the diffracted waves can be measured, which provide their amplitudes but not their phases. This challenge is known as the classical phase problem in X-ray crystallography and has been the subject of intense scientific studies over the years (*70-73*).

Currently, there are three general biophysical methods for solving the phase problem (**Figure 7**). These include molecular replacement, multiple isomorphous replacement (MIR), and multiple-wavelength anomalous diffraction (MAD). In molecular replacement, a known 3D structure of a homologous molecule is used as the initial model for the unsolved structure. Consequently, this method will not work for solving a completely unknown structure, but as the database of known structures increases, so should the number of successful applications of molecular replacement. MIR and MAD are methods that do not exploit a suitable model structure, but they do required heavy-atom derivatives. MIR has been used traditionally for phase evaluation of new macromolecular structures, mostly protein structures. This method, however, suffers from some inherent difficulties in obtaining a suitable heavy-atom derivative. Since the phase information is

obtained from several crystal derivatives of the molecule at a single wavelength, the crystals must remain isomorphous with the native crystal. At the molecular level, this means that the heavy atom must not disturb crystal packing or the conformation of the RNA molecule. Thus, MIR entails the trial-and-error preparation of heavy-atom derivatives and constitutes a bottleneck in X-ray crystallography.

The MAD method was introduced as an alternative to MIR (*74-77*). It is becoming the method of choice in X-ray structure analysis; about two thirds of all new protein structures reported are solved using MAD. This method exploits the unique property of the anomalous scattering effect of a suitable heavy atom to solve the phase problem and determine the structure. Significantly, with this technique the data from one single crystal suffice to solve the phase problem.

In addition, using the innovative advances in synchrotron radiation, which provides continuous tunability to the incident X-ray bean, variation of the heavy atom scattering strength can be accomplished by changing the wavelength. Indeed, it is the difference in scattering strength among each set of chosen wavelengths that provide the

information to solving the phase problem. The advances of the MAD method have been the subjects of a few recent review papers (*78-82*).

At this stage it becomes clear that the crystal structure of a molecule can be solved once a suitable heavy-atom derivative is found, using either MAD or MIR or a combination of both. Hence heavy-atom derivatization is the key step in X-ray crystallography. Therefore, the implementation of new methodologies to incorporate suitable heavy atoms into macromolecules will significantly facilitate the elucidation of new crystal structures. This dissertation is directed toward the development of a new technique to derivatize RNA and DNA with heavy atoms for MAD phasing.

1.6 Heavy Atom Derivatization

The field of protein crystallography has been revolutionized with the development of an exquisite technique to biosynthetically incorporate selenium into proteins. This technique utilized the well-known expression system of *E. coli* to overexpress proteins containing selenomethionine instead of the natural methionine amino acid (*75, 83*). The use of selenomethionyl proteins proved to be excellent for MAD phasing by exploiting the unique scattering properties of

selenium, which contains an absorption K edge of 0.979 Å and is very close to the energy used in synchrotron radiation. This technique has been further extended to the use of eukaryotic expression systems to biosynthetically prepare selenomethionyl proteins that cannot be done in bacteria (*84, 85*). In addition, other different amino acid analogs harboring selenium, tellurium and other heavy atoms have been synthesized to offer different choices of producing protein heavy-atom derivatives (*86, 87*).

More significantly, the success of selenomethionyl proteins has not been only limited to the use of protein structure analysis but it has been extended to the field of RNA X-ray crystallography using MAD phasing. Several RNA structures have been determined *de novo* using selenomethionyl proteins. Three of these structures include the hepatitis delta virus and the hairpin ribozymes, and the RNA-protein core of the signal recognition particle (*51, 52, 88, 89*). The strategy employed in these studies was to design the RNA molecules to specifically bind the selenomethionyl protein, U1A. Similarly, the utilization of selenomethionyl proteins that bind intrinsically to RNAs has facilitated the elucidation of a few other RNA crystal structures (*90*). These RNA-protein complexes were co-crystallized using

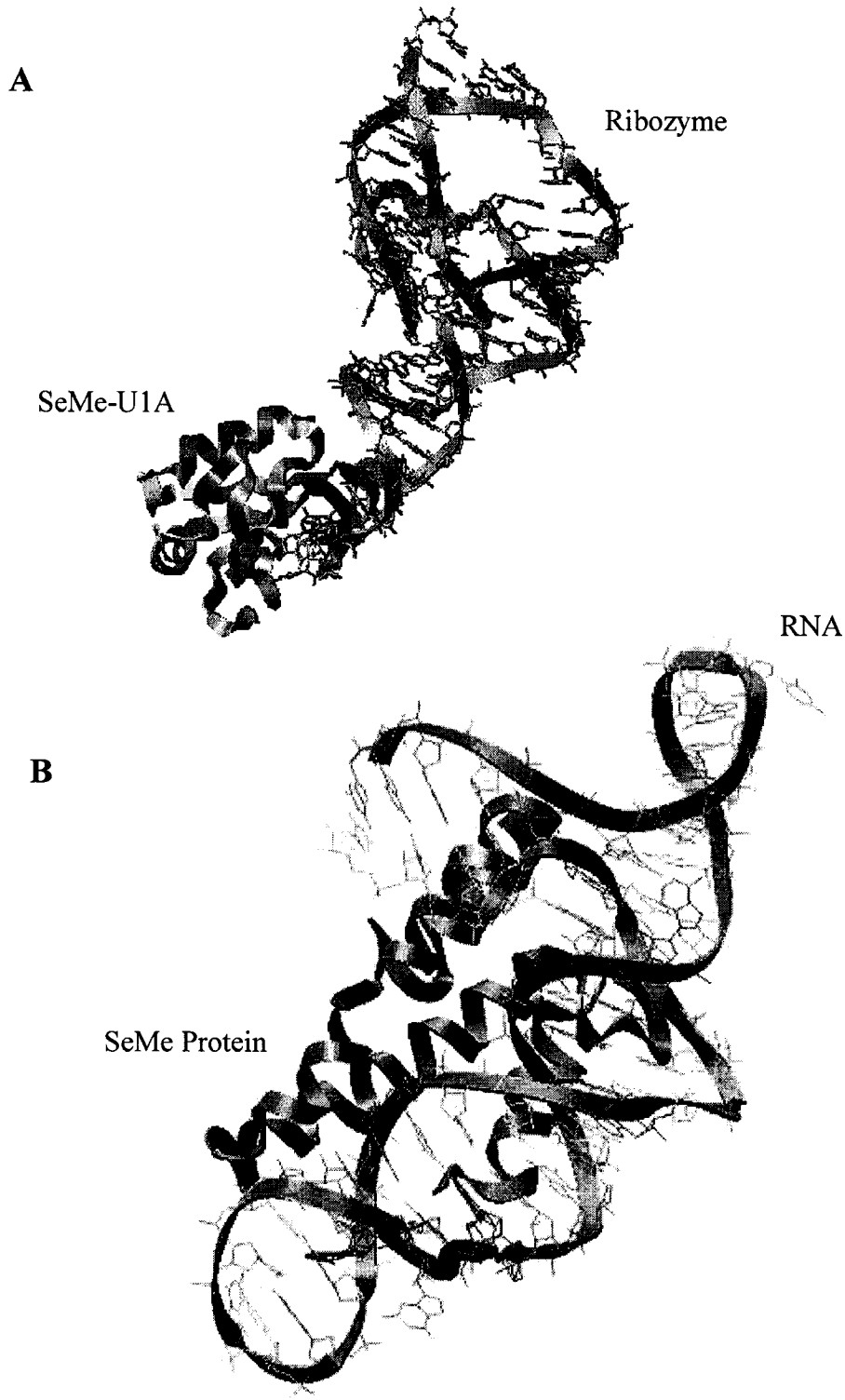


Figure 8. RNA derivatization *via* selenomethionyl proteins. A) HDV ribozyme-U1A complex: U1A extrinsically binds to ribozyme; B) S15-rRNA complex: protein bind intrinsically to RNA (*90*).

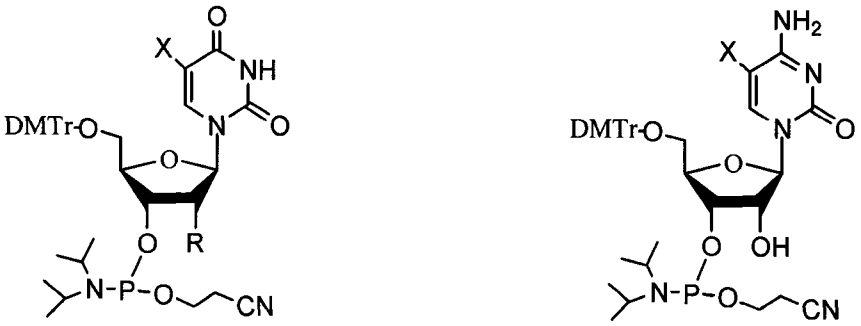
protein screening crystallization conditions (**Figure 8**).

Another method that has been exploited successfully in the preparation of protein heavy-atom derivatives, which has also been extended to RNA and DNA crystallography, is soaking. To prepare the heavy-atom derivative in this way, the crystal of the molecule under study is soaked in a buffer solution of the heavy atom ions. This method, however, has not been very successful in nucleic acids crystallography. A possible reason is that RNA and DNA molecules do not form many specific binding sites as opposed to proteins. In addition, the highly negative charged phosphate backbone in nucleic acid molecules might promote non-specific ligand binding properties.

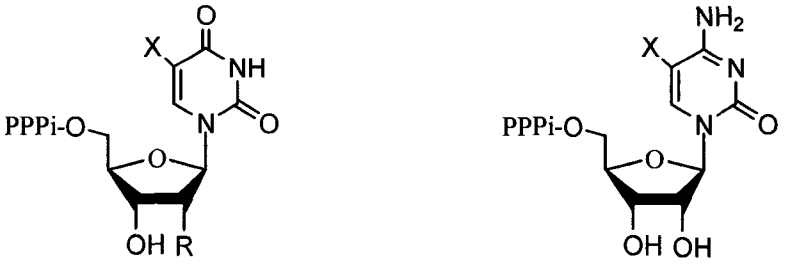
This difficulty in RNA and DNA crystallography has encouraged scientists to explore alternative ways to prepare other types of heavy-atom derivatives. In this context, the syntheses of nucleoside phosphoramidite and nucleoside triphosphates containing covalently bound halogens (bromine or iodine) have been pursued as potential X-ray scattering centers in nucleic acids (**Figure 9**). Accordingly, 5-halouridine, 5-halo-2'-deoxyuridine (a mimic of thymidine) and 5-halo-2'-deoxycytidine have been synthesized to be compatible for the chemical and the enzymatic synthesis of RNA and DNA (*91*). Since

bromine exhibits an absorption K edge very close to that of selenium (0.920 Å), it has been used in MAD phasing structure analyses (*92-94*).

A) 5-halonucleoside phosphoramidite



B) 5-halonucleoside triphosphate



Legend:

a) X = Br or I; R = OH in uridine or H in 2'-deoxyuridine
c) X = Br or I in cytidine
d) PPPi = triphosphate functional group
e) DMTr = dimethoxy trityl protecting group

Figure 9. Nucleoside phosphoramidites and nucleotide triphosphates.

In addition, a few RNA and DNA structures have been elucidated in the past two decades with MIR by exploiting properties of halogens alone or in combination with other types of derivatives (*54*).

Although halogen derivatives have facilitated the elucidation of the crystal structures of a few RNA and DNA molecules, sometimes their applications are not practical. A major problem is that halogen derivatives often cause base-stacking disruptions and other structural perturbations. In addition, it has been found that these derivatives cannot be crystallized under native conditions; and that sometimes they do not diffract as well as the native crystals. These problems might be associated with the fact that halogen derivatives only offer the limited choice of incorporation to the 5-postion of pyrimidines. As a result, these challenges have limited the usefulness of these derivatives to oligonucleotides that can be synthesized chemically with the option of site-specific incorporations; hence halogen derivatization of a long RNA or DNA molecule will not be practical since there is no means to control site-selective incorporations by enzymatic syntheses. Moreover, it has been reported that oligonucleotides modified with halogens are light sensitive; as a

result, long time exposure to X-ray may cause their decompositions (*95, 96*).

1.7 Selenium Chemistry in Nucleic Acids

Although the chemistry of selenium in nucleic acids in not well developed, the presence of 2-selenouridine in the anticodons of several bacterial tRNAs has been the center of investigations during the last two decades. The most abundant seleno-tRNAs are lysine, glutamate, and glutamine (*97, 98*).

The first step of the biological reaction involves the modification of the natural tRNA transcript with sulfur to form 2-thiouridine. In the second step, sulfur is replaced with selenium leading to the formation of 2-selenouridine. It has been demonstrated that the source of selenium comes from selenophosphate. The synthesis of selenophosphate is dependent on the expression of selenophosphate synthetase in the presence of ATP and selenide (*99*).

Chapter 2

OBJECTIVE AND EXPERIMENTAL DESIGN

Over the years, the field of DNA and RNA X-ray crystallography has been quite challenging due to inherent difficulties in creating suitable heavy atoms derivatives for MAD phasing. In order to approach this problem, a novel study is presented in this dissertation. This study exploits a synthetic approach to prepare nucleoside and nucleotide triphosphates containing selenium covalently at desired selective positions for the synthesis of DNA and RNA heavy-atom derivatives.

The utilization of selenium for phasing in DNA and RNA X-ray crystallographic is based on several criteria: 1) Selenium produces a suitable anomalous scattering signal for crystallographic phasing via MAD. It has an absorption K edge of 12.6578 keV, 0.9795 Å that is readily accessible with synchrotron radiation (**Figure 10**). 2) The determination of numerous protein crystal structures has demonstrated that selenium derivatives lead to the formation of isomorphous and highly ordered crystals. To date, protein X-ray structures account for about two-thirds of all new structures determined by MAD phasing.

3) Selenium-heavy atom derivatives can sustain long exposure of high intensity X-ray without decomposition. 4) Finally, selenium lies in the group VI of the periodic table and contains a van der Waals radius of 2.00 Å as compare to 1.85 Å for sulfur and 1.4 Å for oxygen.

These chemical properties suggest that selenium in principle can replace any oxygen atom in a nucleotide. Accordingly, selenium offers many choices of incorporation sites per nucleotide monomer by replacing oxygen atoms located in different chemical and geometrical environments (**Figure 11**). This leads to the aim of this study: the selective replacement of oxygen atoms with selenium to avoid structural perturbations to the modified DNA or RNA molecule.

The present study is divided into two major components. These components are: I) the synthesis of nucleoside phosphoramidites containing selenium for the solid phase synthesis of DNA and RNA oligonucleotides; and II) the synthesis of nucleoside triphosphates containing selenium for the enzymatic synthesis of DNA and RNA (**Figure 12**). Hence the syntheses of three types of nucleoside phosphoramidites compatible for the solid phase synthesis of RNA and A-form DNA are presented. In the first one, the selenium functional group is introduced at the 5'-position of nucleosides; while

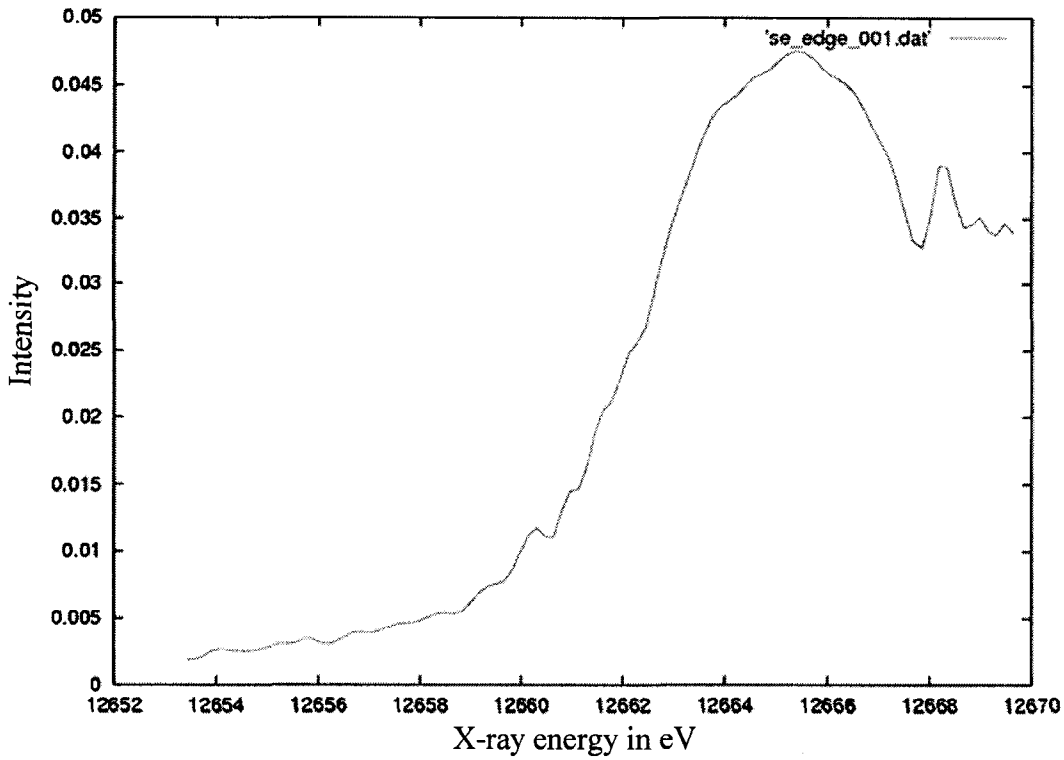


Figure 10. X-ray fluorescence absorption spectrum of a DNA modified with selenium. Se K-edge is 12.6578 keV, 0.9795 Å (*100*).

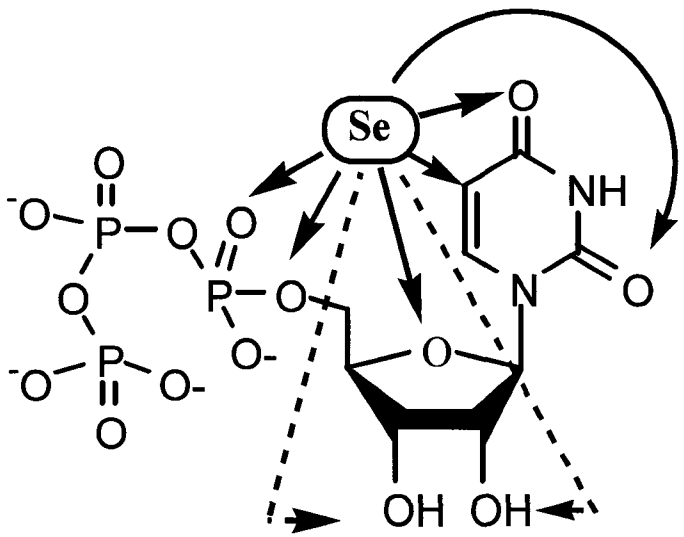


Figure 11. Positions where selenium can be introduced in a nucleoside triphosphate. In principle, selenium can replace every oxygen atom. It also can be incorporated at the 5-position of uridine, as indicated in this drawing.

in the second one, the selenium functionality replaces the 2'-hydroxyl group of uridine and cytidine. Likewise, the synthesis of three types of nucleoside triphosphates for the enzymatic synthesis of DNA and RNA is investigated. These nucleoside triphosphates contain the selenium functionality at one of the non-bridging positions in the α-phosphate group. The syntheses and the enzymatic incorporation of the thymidine, uridine, and adenosine analogs are also presented to make selenium-derivatized DNAs and RNAs.

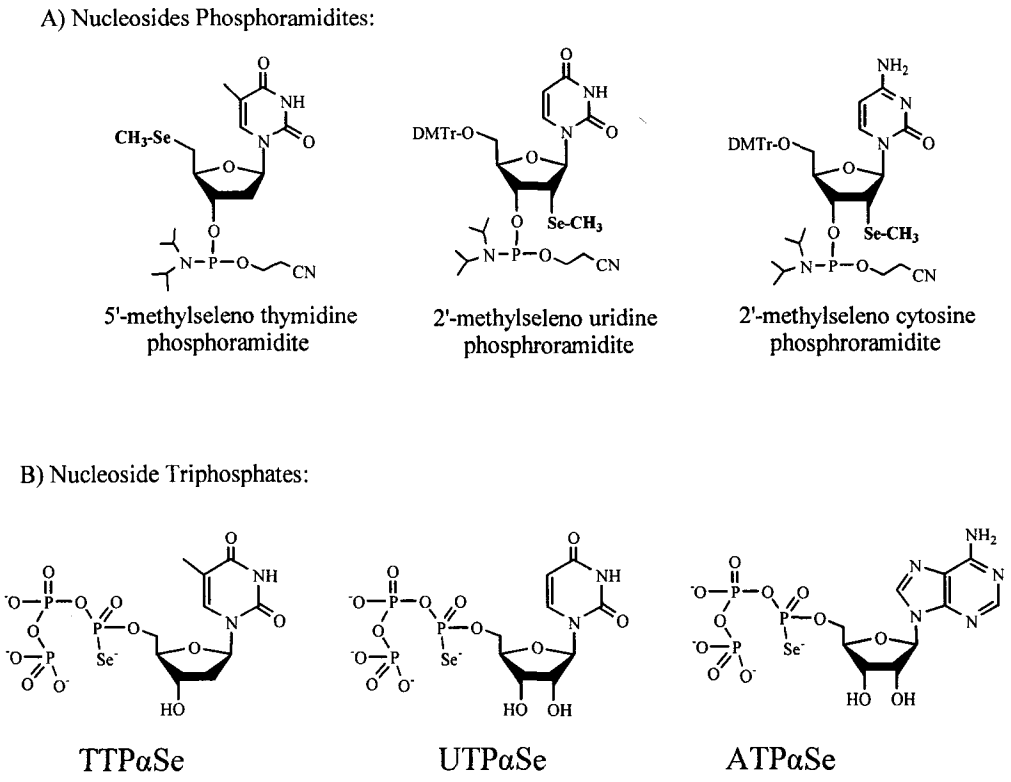


Figure 12. Chemical structures of nucleoside phosphoramidites and nucleoside triphosphates containing selenium.

Chapter 3

SYNTHESIS OF NUCLEOSIDE PHOSPHORAMIDITES CONTAINING SELENIUM AT THE 5' POSITION AND THEIR INCORPORATION INTO DNA

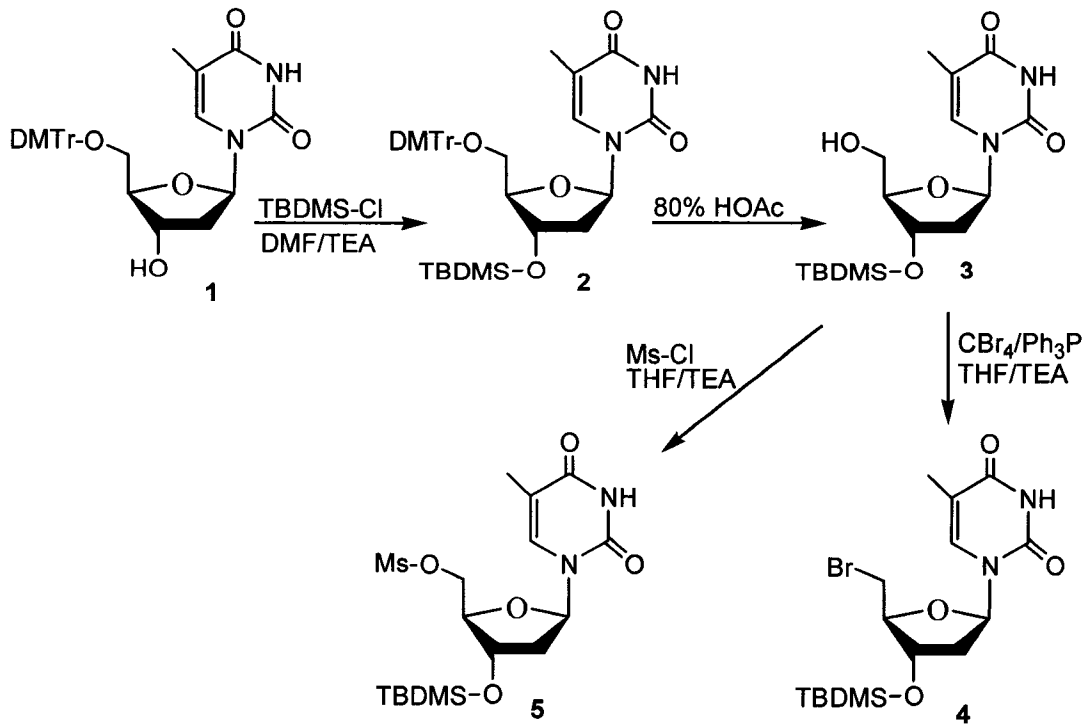
3.1 Introduction

The 5' position of nucleosides was studied for the incorporation of selenium. The objective of this study was to investigate the compatibility of the selenium functionality with the solid phase synthesis of DNA or RNA oligonucleotides using phosphoramidite chemistry. The strategy involved the replacement of the 5'-hydroxyl group with various selenium functionalities by nucleophilic substitution. Nucleophilic substitution was achieved *via* activation of the 5' position with either the bromo or the mesyl derivative and displacement with nucleophilic selenium reagents. Several nucleophilic selenium reagents were studied in combination with a one-phase or a two-phase system. They included sodium methyl selenide ($NaSeCH_3$), sodium hydrogen selenide (NaHSe), and sodium diselenide (Na_2Se_2). Activation of the 3' position with the phosphoramidite functionality was carried out in good yield. A thymidine-thymidine DNA dinucleotide (^{Se}TpT) and a hexamer DNA oligonucleotide (^{Se}TpGpCpGpCpA) were prepared to confirm the

compatibility of the selenium functionality with the solid phase synthesis following standard procedures of the phosphoramidite chemistry (*101*). Characterization of the oligonucleotides was performed spectroscopically using ^{1}H-NMR, ^{31}P-NMR, ^{77}Se-NMR, and mass spectroscopy. The data demonstrated the suitability of selenium for the synthesis of selenium-derivatized DNA and RNA using the solid phase approach *via* the phosphoramidite chemistry.

3.2 Preparation of the Nucleoside Precursors

The preparation of the building blocks for the synthesis of nucleoside phosphoramidites containing selenium at the 5' position was achieved using standard procedures. The thymidine nucleoside was used as a model to develop the chemistry because it did not require any protecting group on the base. The reaction started with 5'-DMTr protected thymidine followed by protection at the 3' position with TBDMS, and subsequent deprotection of the DMTr moiety to produce **3** (**Scheme 1**). The 5' position of compound **3** was initially activated *via* the Mitsunobu reaction (*102*) by replacement of the hydroxyl moiety through bromo substitution to produce compound **4**.



Scheme 1. Synthesis of the 5'-mesyl and 5'-Bromo nucleoside precursors.

Some difficulties were encountered during the course of the Mitsunobu reaction. Preliminary experiments demonstrated *in situ* generation of acids that resulted in the deprotection of the DMTr functionality and the generation of unidentified products. The generation of the acidic environment was attributed to trace of water in the reaction mixture that promoted the production of hydrobromic acid. The inclusion of some preparatory steps to run the reaction under anhydrous condition in the presence of dry argon and the base triethylamine (TEA) afforded compound **4** in good yield.

Additionally, after evaporation of all the solvents, chromatographic purification of **4** from the byproduct triphenylphosphine oxide became a major challenge. As indicated by analytical thin layer chromatography (TLC), triphenylphosphine oxide co-migrated with the desired product. This challenge was later circumvented by precipitating out triphenylphosphine oxide in a 30% solution of ethyl acetate (EtOAc) in hexane before chromatographic purification.

Table 1. Nucleophilic selenium reagents used in nucleoside chemistry

Se Reagents	Structures	Solubility	References
Sodium hydrogen selenide	NaHSe	poor	(*103*)
Sodium diselenide	Na_2Se_2	poor	(*103*)
Sodium methyl selenide	$NaSeCH_3$	good	(*104*)
Sodium phenyl selenide	NaSePh	good	(*105*)

A more practical approach to activate the 5' position of compound **3** was later developed by direct mesylation of the hydroxyl group. Mesylation was carried out with 1.5 molar equivalents of methanesulfonyl chloride under argon in an ice-water bath. Chromatographic purification of the crude product gave compound **5** quantitatively. Other nucleoside precursors, such as adenosine, guanosine, and cytosine, were also prepared with little modifications to these procedures.

3.3 Choosing a Selenium Reagent

The introduction of selenium at the 5' position by nucleophilic substitution of either the bromo or the mesyl of the nucleoside derivative was studied using several selenium reagents. The availability of an ideal nucleophilic selenium reagent that would be compatible with the chemistry of nucleic acids in general became a limited option. Chemically, an ideal reagent for this purpose would be one that is easy to prepare, highly reactive, stable, neutral, and soluble in most organic solvents; sterically, the selenium functionality would need to be as small as possible in order to minimize structural perturbations to the nucleosides. Based on these criteria, three nucleophilic selenium reagents were chosen for preliminary studies. These reagents were NaHSe, Na_2Se_2, and $NaSeCH_3$.

The preparation of Na_2Se_2 was first attempted following the procedures reported by Ludwig Syper (*104*). In this case, elemental powdered selenium was reacted with hydrazine hydrate in strong basic aqueous conditions (5 M NaOH). This condition turned out to be impractical since it promoted the partial hydrolysis of the nucleosides. A more practical approach was later explored in which elemental selenium was reacted with sodium borohydride in water

(*103*). This method allowed the synthesis of both NaHSe and Na_2Se_2 by varying the molar ratio of selenium metal and sodium borohydride. For example, sodium selenide was easily prepared by reducing selenium metal with one molar equivalent of $NaBH_4$ to afford a clear aqueous solution. Another equivalent of selenium metal was added to the solution of sodium selenide to prepare sodium diselenide. The appearance of a brown-reddish color indicated the formation of the reagent. The pH of an aqueous solution of these selenide salts was determined to be around 11. In addition, a two-phase system was utilized to carry out the reaction under mild condition because of the poor solubility of Na_2Se and Na_2Se_2 in organic solvents. The preparation of $NaSeCH_3$ is described in a later section.

3.4 Introduction of Selenium *via* a Two-Phase System

The incorporation of selenium at the 5' position was first investigated using a two-phase system. This system was composed of a toluene and an aqueous layer **(Figure 13)**. The reaction was facilitated by the phase-transfer catalyst tetrahexyl ammonium hydrogen sulfate. The rationale for using a two-phase system was that it provided a mild environment where selenium nucleophilic substitution would take

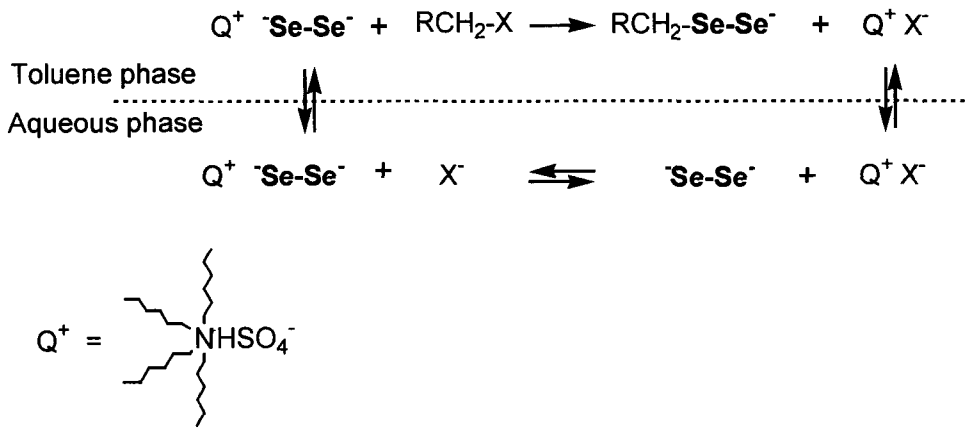
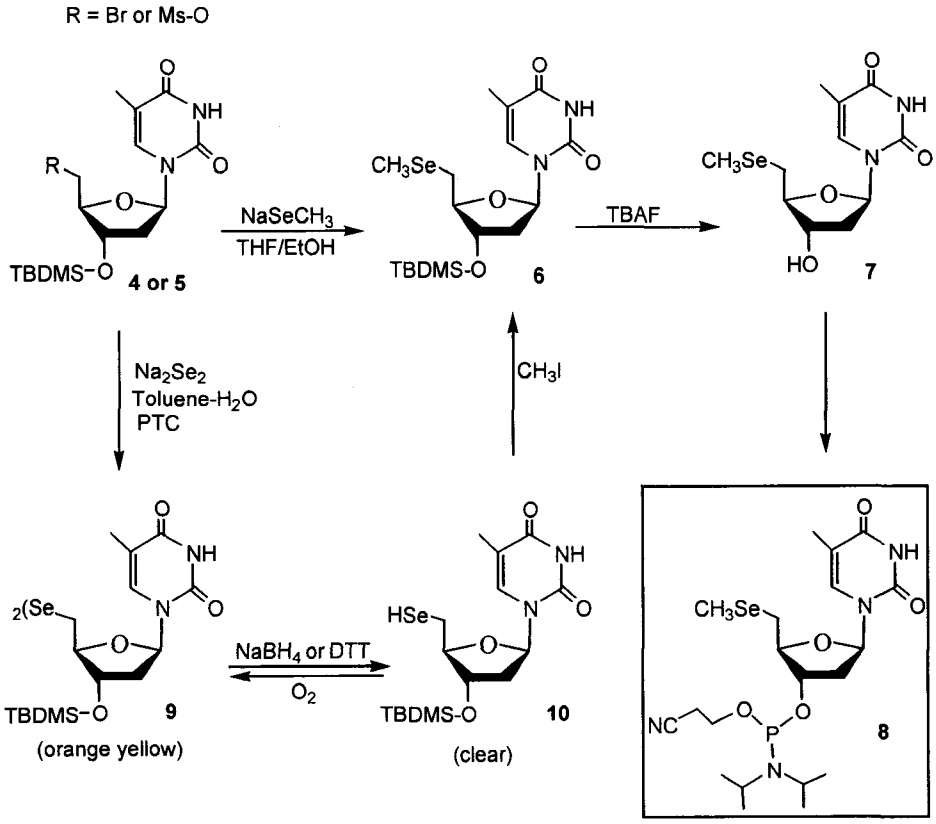


Figure 13. Schematic representation of a two-phase system reaction catalyzed by a phase-transfer catalyst.



Scheme 2. Incorporation of the selenium functionality at the 5' position of thymidine and the synthesis of the corresponding phosphoramidite.

place under ideal conditions. For example, in the toluene phase, the selenide anions (HSe^- or Se_2^{2-}) would be less solvated and, thus, more reactive toward the electrophiles. Furthermore, since the nucleoside would be dissolved in the toluene phase, it would be less susceptible toward the strong basicity present in the aqueous medium.

The convenience of this two-phase system was first demonstrated in the use of NaHSe and compound **4** or **5** at room temperature. The reaction was carried out by injecting the aqueous solution of NaHSe into the toluene solution containing either compound **4** or **5**. In this case, the nucleophilic reaction did not stop at the selenol intermediate, but it proceeded and reacted with another molecule of **4** or **5** to produce a dialkylated and symmetrical product. The formation of the dialkylation product was observed even when a large amount of the catalyst was used. Experiments demonstrated that DTT or $NaBH_4$ could not reduce this product.

The formation of the dialkylated species was later overcome by replacing NaHSe with Na_2Se_2 as the nucleophile in the aqueous medium in the two-phase system. In this case, the desirable dialkylated diselenide was obtained as an orange crystalline product in high yield after chromatographic purification. This product was

characterized by electrospray mass spectroscopy. As illustrated in **Figure 14,** the mass spectrum shows a unique isotopic distribution of corresponding to selenium. Selenium has four isotopes of most natural abundance; they are ^{80}Se (49.6%), ^{78}Se (23.5%), ^{82}Se (9.4%), and ^{76}Se (9.0%). The mass for ^{80}Se plus a proton was 839.2, which is consistent with the expected results.

Reduction of the diselenide bond with DTT of $NaBH_4$ produced the corresponding selenol product. Preliminary experiments indicated that this selenol was very unstable in the presence of O_2 as it was oxidized to the corresponding diselenide functionality. As a result, to minimize air oxidation, the synthesis of the selenol was always performed under the presence of dry argon. A permanent protection was then performed on the selenol with methyl iodide to form the corresponding 5'-methylseleno derivative. The synthesis of the selenium triphosphate analog was also attempted without success from the selenol using standard procedures for the preparation of ordinary nucleoside triphosphates.

3.5 Introduction of Selenium *via* a One-Phase System

The instability of the selenol species prompted us to investigate a

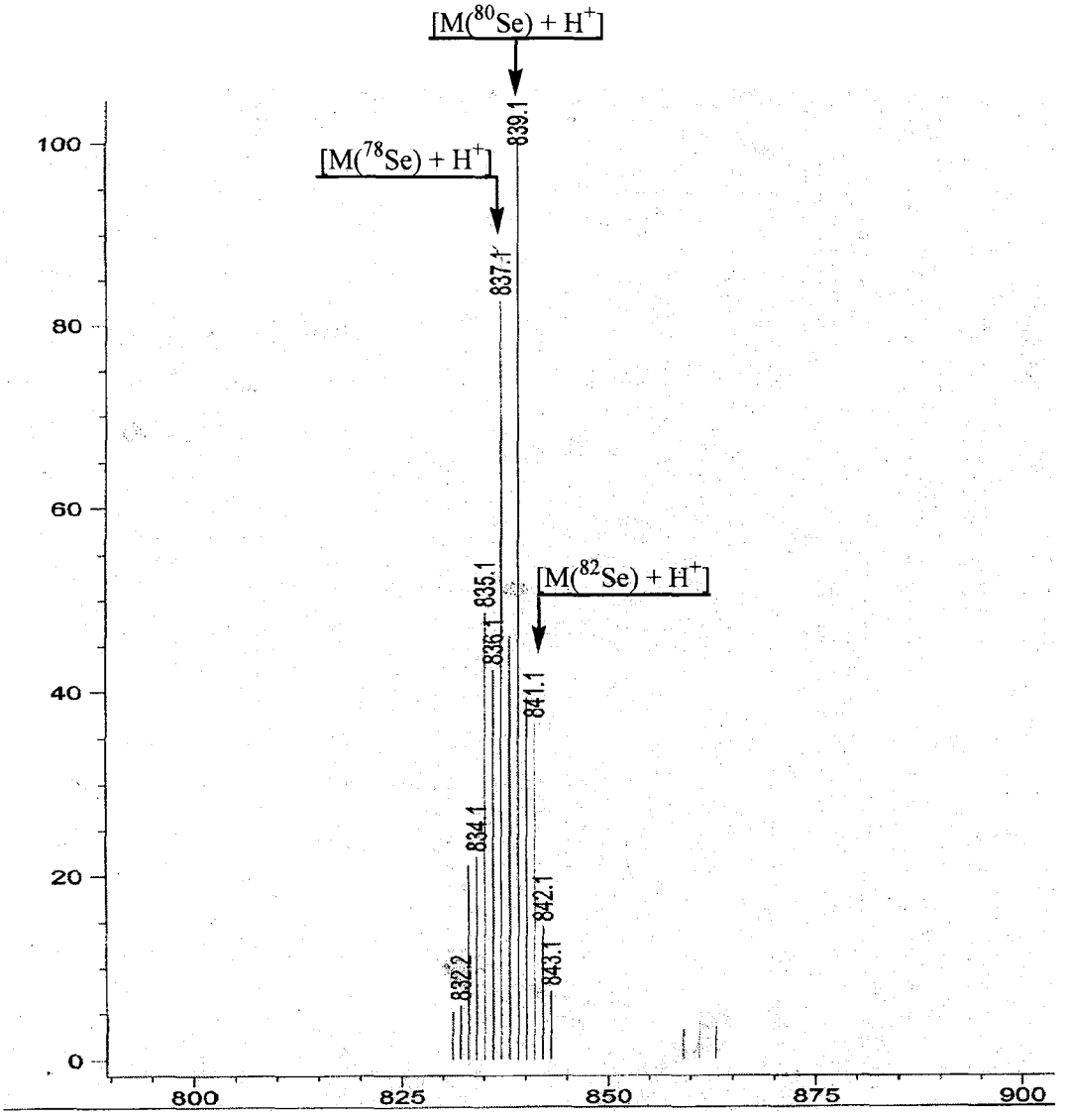


Figure 14. Electrospray mass analysis (positive mode) of dialkylated diselenide **9**. Calculated mass ($C_{32}H_{54}N_4O_8Se_2Si_2$) 839.2 $[M(^{80}Se) + H]^+$.

more convenient route for the synthesis of the 5'-methylseleno nucleoside derivatives. Thus, the use of sodium methyl selenide was explored in order to directly displace the bromo or the mesyl group at

the 5' position of nucleosides. $NaSeCH_3$ was easily prepared by reacting dimethyl diselenide with three molar equivalents of sodium borohydride in THF containing 25% ethanol. The disappearance of the yellowish color of dimethyl diselenide indicated the complete formation of $NaSeCH_3$. In a typical procedure, using the conventional approach of one-phase system, the incorporation of the methylseleno group was accomplished by injecting a freshly prepared solution of the sodium methylselenide into a THF solution of **5** (**Scheme 2**).

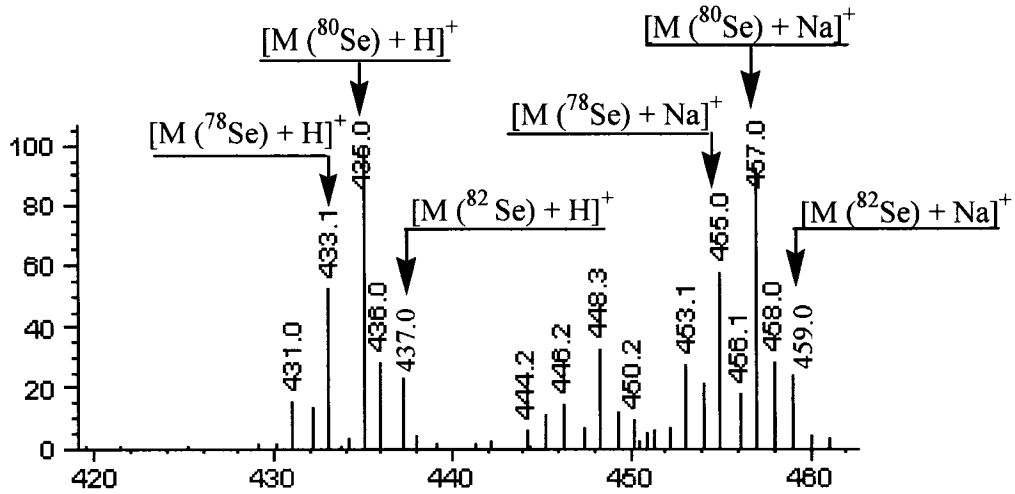


Figure 15. MS spectrum (positive mode) of 5'-methyleno thymidine with 3'-TBDMS. The molecular weight ($C17H30N2O4SiSe$) is 433 for the Se atomic mass of 80; MS peaks: 434 $[M(^{78}Se) + H]^+$, 436 $[M(^{80}Se) + H]^+$, 438 $[M(^{82}Se) + H]^+$, 456 $[M(^{78}Se) + Na]^+$, 458 $[M(^{80}Se) + Na]^+$, and 460 $[M(^{82}Se) + Na]^+$.

Although chromatographic purification of the crude mixture gave **6** in good yield, it let to some unsatisfactory results when performed using nucleosides containing base-labile protecting groups; partial deprotection of these protecting groups occurred. The use of the two-phase system greatly reduced the partial hydrolysis of the base-protecting groups. The results of mass spectroscopy and ^{77}Se-NMR analyses of compound **6** are illustrated in **Figure 15** and **Figure 16.**

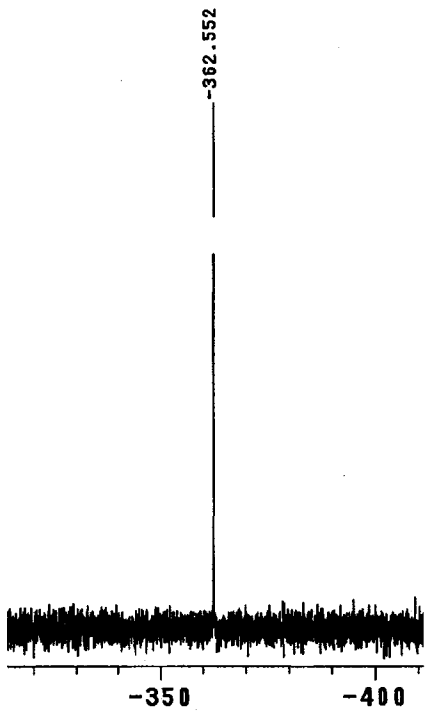


Figure 16. ^{77}Se-NMR spectrum of 5'-methylseleno thymidine. The analysis was done in $DCCl_3$ using dibenzyl diselenide as reference, 133.25 ppm.

3.6 Synthesis of the Phosphoramidite Analogs and Oligonucleotides Derivatized with Selenium

The compatibility of the 5'-methylseleno nucleosides with the phosphoramidite chemistry was studied using standard conditions. In order to free the 3' position, the 5'-methylseleno analogs (**6**) were treated with tetrabutylammonium fluoride to remove the 3'-TBDMS protecting group in a THF solution. After chromatographic purification, the product was reacted with 2-cyanoethyl N,N-diisopropylchloro phosphoramidite to activate the 3' position following standard procedures. To confirm the compatibility and synthesis of the phosphoramidite analog, 5'-$^{Se-}$TpT was prepared in one-pot reaction in acetonitrile. The coupling reaction was carried out by adding one molar equivalent of 5'-OH thymidine to the reaction mixture of the phosphoramidite analog in the presence of tetrazole. After the completion of the coupling reaction, the phosphite intermediate was oxidized to the corresponding phosphate intermediate with an iodine solution. Subsequent addition of aqueous ammonia to remove the 2-cyanoethyl functionality gave rise to the desired dinucleotide analog carrying selenium at the 5' position. The

structure of 5'-^{Se}-TpT was confirmed spectroscopically (**Figure 17** and **Figure 18**).

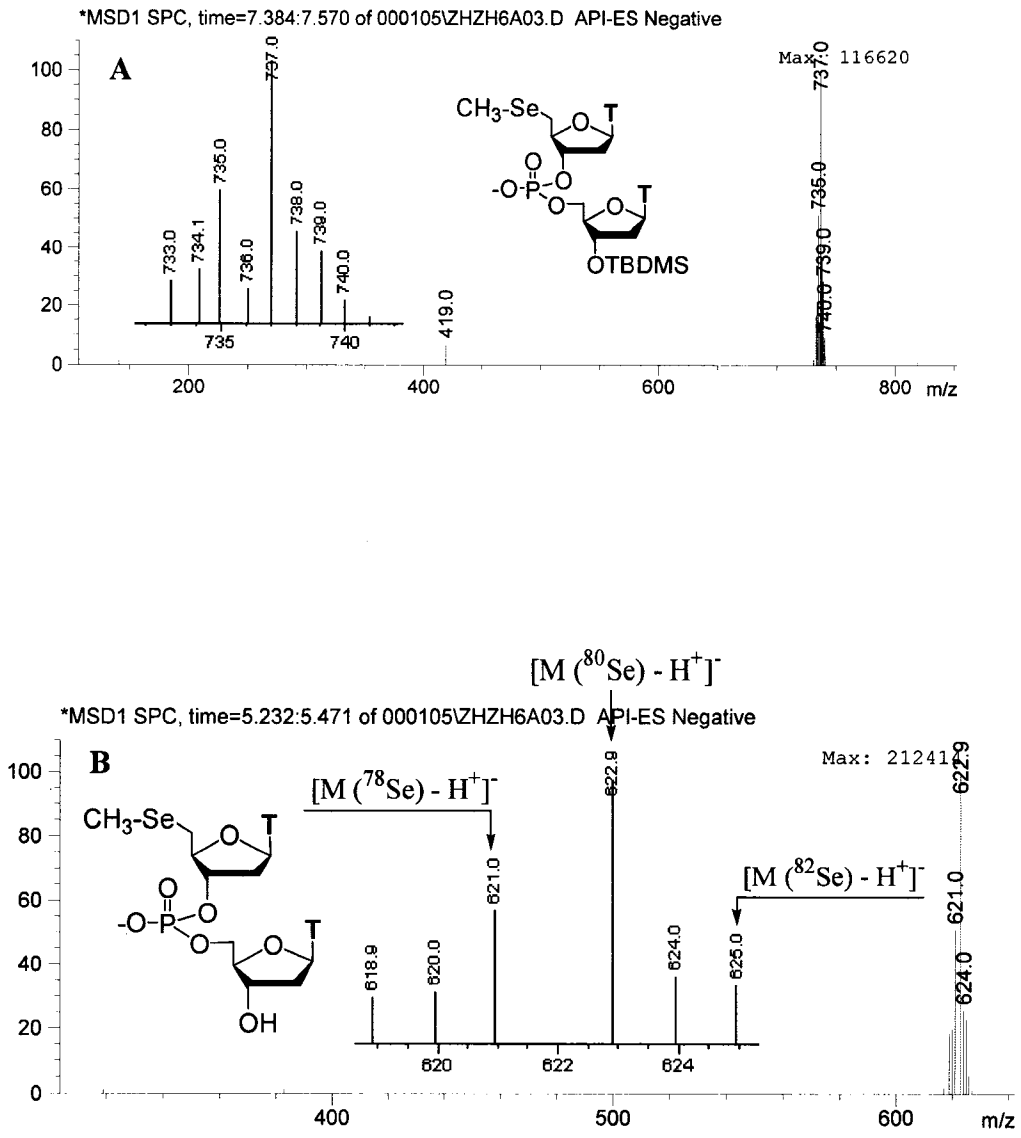


Figure 17. Electrospray negative ion MS analysis of 5'-Se-TpT dinucleotide. A) With 3'-TBDMS group; B) Without 3'-TBDMS group.

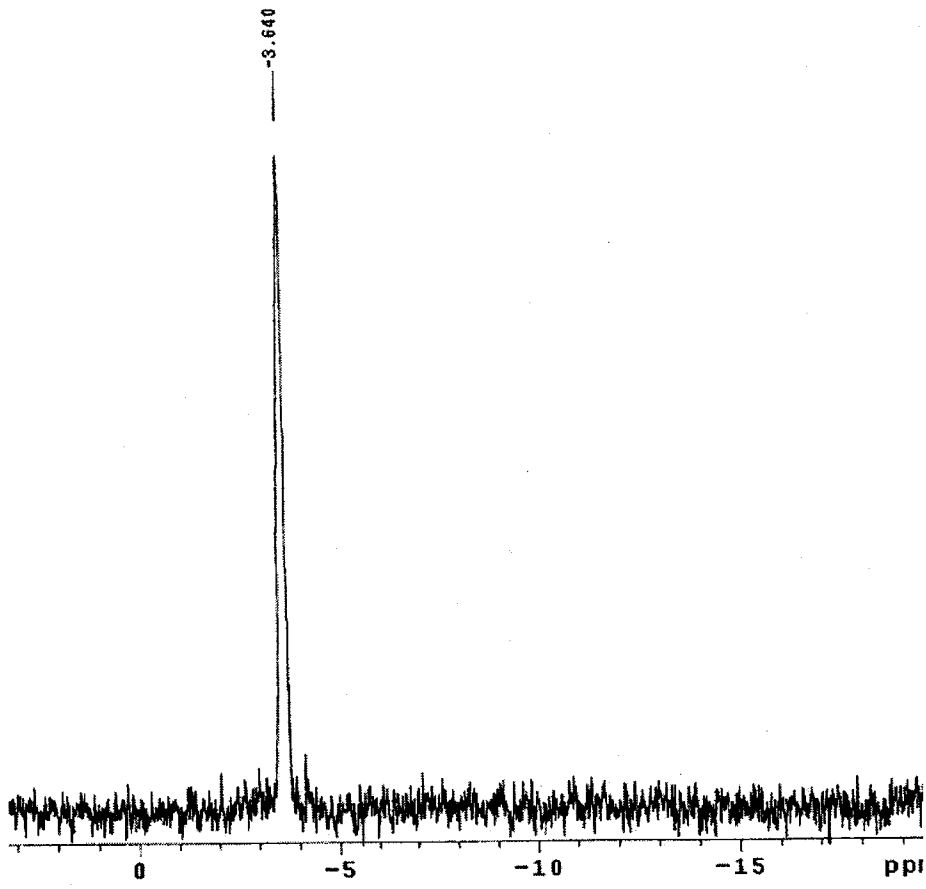


Figure 18. ^{31}P-NMR and ^{77}Se-NMR of 5'-Se-TpT. ^{31}P-NMR was carried out in D_2O and Na_3PO_4, pH 7.0.

These results prompted the synthesis of longer DNA oligonucleotides on solid support using the 5'-methylseleno phosphoramidite analog. Therefore, a hexamer DNA containing selenium at the 5'-position (5'-$^{Se-}$TpGpCpGpCpA) was synthesized under standard condition using a DNA/RNA synthesizer. The result of the incorporation of selenium

was confirmed by mass spectroscopy (**Figure 19**). Although the iodine oxidation step was a concern, the MS data did not show any trace of selenium oxidation in either case of both oligonucleotides under the experimental conditions.

3.7 Conclusions

Nucleoside derivatives were synthesized by the displacement of the 5'-hydroxyl group with the selenium functionality. These nucleoside derivatives were converted into the corresponding 3'-phosphoramidite analogs for the solid phase synthesis of DNA or RNA. The compatibility of the selenium functionality with the solid phase method was investigated in the synthesis of a $^{Se-}$TpT dinucleotide and a 5'-Se-hexamer DNA. The results presented here demonstrated the suitability of selenium for the solid phase synthesis of DNA or RNA using the phosphoramidite chemistry. It was observed that the selenium functionality was stable under the iodine oxidation step during the solid phase synthesis.

Although the 5'-methylseleno nucleosides allow derivatization of DNA or RNA *via* the solid phase method, this application has some limitations. During the solid phase synthesis, the polymerization of

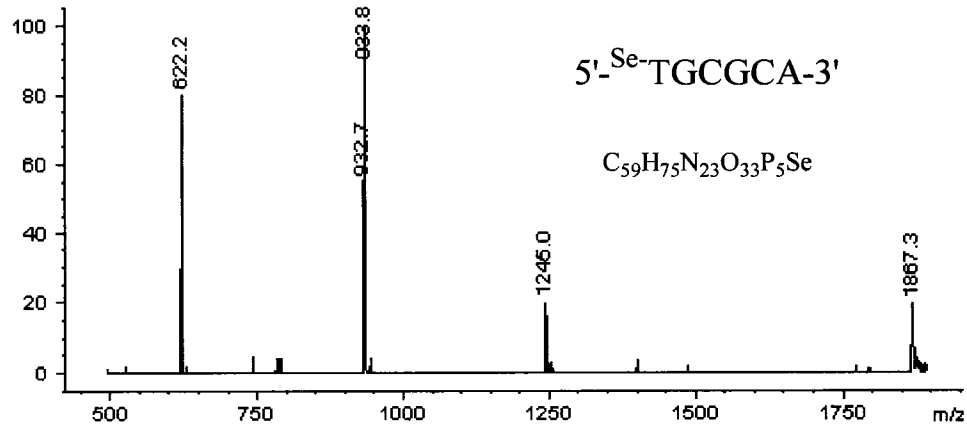


Figure 19. MS spectrum of 5'-^{Se}-TGCGCA-3' by electrospray (negative ion mode). The molecular weight ($C_{59}H_{75}N_{23}O_{33}P_5Se$) is 1868 for ^{80}Se. Measured (expected) m/z: $[M-H^+]^-$= 1867.3 (1867), $[M-2H^+]^{2-}$ =933.8 (933), $[M-3H^+]^{3-}$ =622.2 (621.7), $[2M-3H^+]^{3-}$ =1245.0 (1244.3).

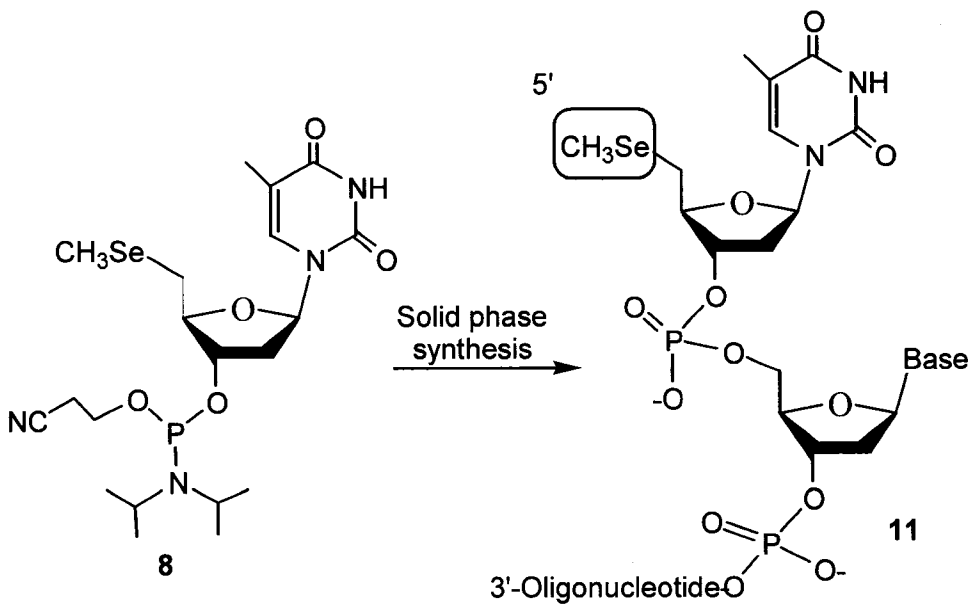


Figure 20. DNA polymerization on solid support using 5'-methylseleno thymidine. The 5'-methyl group prevents further polymerization.

DNA and RNA occurs in a 5' to 3' direction on solid support. Since the methylated selenium functionality at the 5' position interferes with the continuation of the polymerization reaction, it only offers the choice of incorporation at the 5' termini of oligonucleotides. It does not support the internal incorporation of selenium (**Figure 20**). The results of these findings were published in reference (*106*).

In order to provide a more versatile approach to make DNA and RNA derivatized with selenium at internal positions, a new approach was developed to introduce selenium at the 2' position of nucleoside phosphoramidites. The development of this method is examined in the next chapter.

Chapter 4

SYNTHESIS OF NUCLEOSIDES PHOSPHORAMIDITES CONTAINING SELENIUM AT THE 2' POSITION AND THEIR INCORPORATION INTO DNA AND RNA

4.1 Introduction

The incorporation of the selenium functionality at the 5' position of nucleoside phosphoramidites and DNA oligonucleotides provided some valuable insights about the compatibility of the selenium functionality with the phosphoramidite chemistry. It was demonstrated that one cycle of iodine oxidation, during the standard solid phase synthesis of DNA, showed no evidence of selenium oxidation. These preliminary results provided impetus for the development of a new type of selenium nucleoside analogs that would facilitate the synthesis of DNA and RNA oligonucleotides containing heavy atom labels either internally or at a terminal end on solid support using the phosphoramidite chemistry. This idea was tested with uridine and cytidine.

Uridine, as in the case of thymidine, can serve as a synthetic model since it requires no base-protecting groups and is relatively easy to work with. In addition, uridine will provide a window to, not only investigate the selenium derivatization strategy with RNA, but also

with A-form DNA oligonucleotides, since the $C_{3'}$-*endo* conformation of the sugar pucker in RNA duplexes are also adopted by A-form DNA oligonucleotide duplexes. More significantly, uridine provides an easy route for its transformation into cytidine.

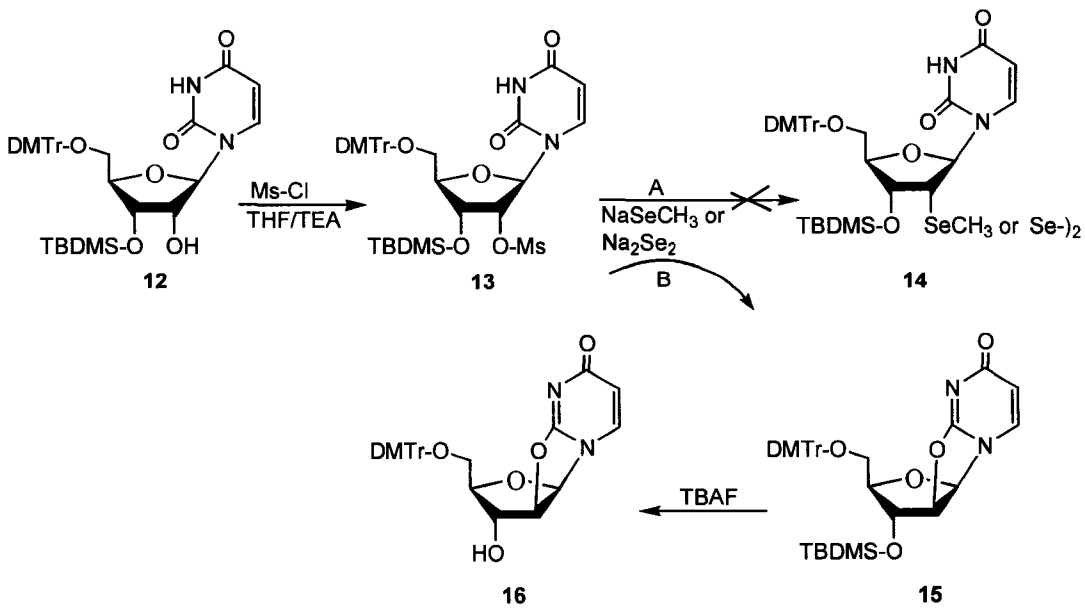
The 2' position was selected among the different positions in uridine to perform oxygen replacement by selenium. The design involved the replacement of the 2'-hydroxyl group with selenium by nucleophilic substitution. In this way, the introduction of selenium at this position was achieved through activation of the 2' carbon center by the formation of a 2,2'-*O*-anhydro cyclic nucleoside intermediate. This intermediate was synthesized *via* two different routes using either unprotected or partially protected uridine nucleosides as immediate precursors.

The incorporation of selenium into the 2' position of cytidine was also investigated. The synthesis of the cytidine analog was first studied using the 2,2'-*O*-anhydro cyclic intermediate without satisfactory results. A more practical approach was later explored, in which 2'-methylseleno uridine was directly converted into the corresponding cytidine analog via a triazolide intermediate using known procedures (*107*). The compatibility of 2'-methylseleno

uridine and 2'-methylseleno cytidine was then demonstrated in the synthesis of DNA and RNA oligonucleotides on solid support using the phosphoramidite chemistry. The results presented in this chapter provided further insights toward the development of the selenium derivatization strategy for nucleic acid X-ray crystallography.

4.2 Preparation of the Uridine Precursors

The incorporation of selenium at the 2' position of uridine was first investigated using partially protected nucleoside precursors. The synthesis started by mesylation of **12** at the 2' position to produce the mesylate **13** in high isolated yields (**Scheme 3**). The incorporation of the selenium functionality was first approached by direct nucleophilic displacement of the mesyl group of **13** by either Na_2Se_2 or $NaSeCH_3$ using the two-phase system method described previously. In the course of theses experiments, spectroscopic analyses failed to confirm the presence of selenium. What these data showed was the formation of the same product from the two types of selenium displacement reactions. Careful analyses of these data later revealed the identity of this product as the cyclic intermediate: 2,2'-*O*-anhydro uridine **15**.



Scheme 3. Synthesis of 2,2'-*O*-anhydro uridine using partially protected uridine precursors.

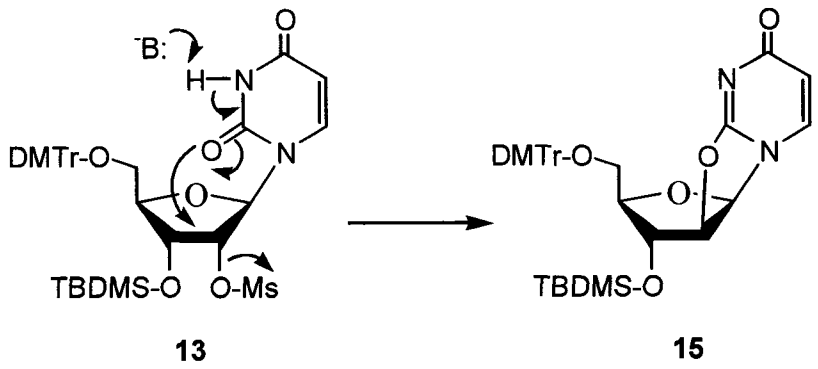


Figure 21. Base promoted formation of 2,2'-*O*-anhydro uridine.

The displacement of the mesyl leaving group by the exo-2-oxygen moiety gave rise to the cyclic intermediate. The formation of this intermediate was promoted by the basic medium, pH about 11 (**Figure 21**). Under this condition, the exo-2-oxygen in uridine was activated and displaced the mesyl group at the 2'-position. Literature reports revealed the identification and the utilization of **15** in nucleophic reactions (*107*).

Intermediate **15** was essential for the development of a synthetic approach to incorporate the selenium moiety at the 2'-position, and became a key precursor to control the stereochemistry of the reaction. It provided a conformational environment that allowed the selenium nucleophile to attack from the α-face and displace the exo-2-oxygen. Based on these preliminary results, **15** was prepared in large scale using toluene and sodium carbonate (Na_2CO_3) in the presence of the phase transfer catalyst. After partial purification of **15** by extraction, it was deprotected at the 3'-position with tetrabutylammonium fluoride to afford **16** in good yield.

The synthesis of 2-2'-*O*-anhydro uridine was later performed following the procedures developed by McGee (*107*). This method was very convenient since it allowed its preparation at a 50-gram

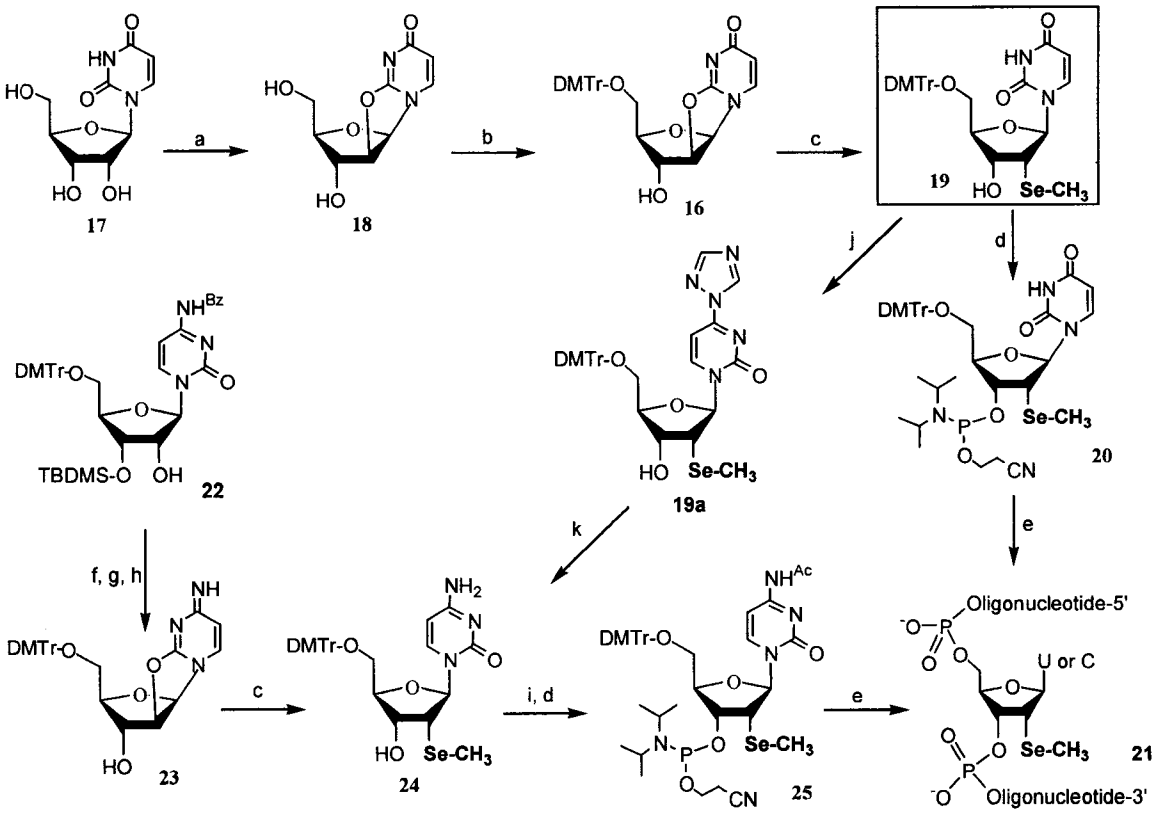
scale starting from cheap and unprotected uridine. Briefly, uridine (**17**) is reacted with diphenyl carbonate and sodium bicarbonate in a minimum amount of DMF (**Scheme 4**). After chromatographic purification, 2-2'-*O*-anhydro uridine was protected at the 5' position with the DMTr group, followed by another purification step, to afford **16** over 80% yield.

4.3 Synthesis of 2'-Methyseleno Uridine Phosphoramidite

The selenium functionality was introduced at the 2' position of uridine by nucleophilic substitution. The incorporation was achieved by direct displacement of the exo-2-oxygen in **16** with sodium methyl selenide in a THF solution containing about 10-15% ethanol. Sodium methyl selenide was prepared as indicated in the previous chapter. Spectroscopic analyses were used to confirm the synthesis of 2'-methylseleno uridine. The stereochemistry of methylseleno group was determined to adopt the α-conformation by 2D-NMR and NOE experiments (*108*).

The 3' position of 2'-methylseleno uridine was then activated with the phosphoramidite functionality using similar procedures for the

synthesis of the corresponding 5'-methylseleno thymidine analog examined in the previous chapter.



Scheme 4. Synthesis of 2'-methylseleno uridine and 2'-methylseleno cytidine phosphoramidite derivatives. a) $(Ph)_2CO_3$, Na_2CO_3 and DMF; b) DMTr-Cl in Pyridine; c) $NaSeCH_3$ and EtOH-THF; d) 2-Cyanoethyl N,N-diisopropyl-chlorophosphoramidite and N,N-diisopropylethylamine in CH_2Cl_2; e) Synthesis of oligonucleotides on solid phase. (f) Ms-Cl, TEA and THF; g) Toluene/tetrahexylammonium hydrogen sulfate and Na_2CO_3 (sat.); h) $(Bu)_4N^+ F^-$ in THF; i) TMS-Im, then Ac_2O, TEA and DMAP in THF; j) TMS-Im, then $POCl_3$-triazole-TEA in CH_3CN; k) NH_4OH.

4.4 Conversion of 2'-Methylseleno Uridine into 2'-Methylseleno Cytidine Phosphoramidite

In order to increase the repertoire of nucleosides analogs containing the selenium functionality for the preparation of DNA and RNA heavy atom derivatives, the synthesis of 2'-methylseleno cytidine was undertaken. The original design was based on the synthesis of the aforementioned 2'-methylseleno uridine using partially protected nucleoside precursors. The reaction steps involved mesylation at the 2'-position of **22**, followed by the displacement of the 2'-mesyl group by the cytidine exo-2-oxygen under a toluene-aqueous methylamine two-phase system that was catalyzed by a phase transfer catalyst (**Scheme 4**). Treatment with tetrabutylammonium fluoride to remove the TBDMS group at the 3' position afforded compound **23**. Attempt to introduce the methylseleno group using a one-phase system gave rise to unsatisfactory yields. The poor substitution by methyl selenide was attributed to the fact that the exo-2-oxygen in cytosine is a much poor leaving group in comparison to its uridine counterpart (**16**).

An alternative strategy to introduce the methylseleno functionality into cytidine was later investigated. This approach consisted in the conversion of uridine into cytidine. Similar conversion reactions have

been reported in the literature in the past few years involving the synthesis of other types of cytidine analogs from uridine (*107, 109, 110*). The key step in this transformation is the activation of position 4 of uridine using a triazolide intermediate. The chemical transformation started from transient silylation of compound **19** at its 3' position with TMS-imidazole, followed by the formation of the triazolide intermediate at position 4 (**19a**). The formation of this intermediate was carried out using *in situs* generated phosphorus oxytriazolide from 1,2,4-triazole and phosphorus oxychloride in acetonitrile in the presence of TEA. The triazolide intermediate **19a** was then converted to the corresponding cytidine derivative by aqueous ammonia treatment in dioxane. After purification, the cytidine product **24** was re-silylated at the 3' position, followed by acetylation of the amino group and 3' deprotection. The isolated product was later activated at the 3' position with the phosphoramidite functional group for the synthesis of DNA or RNA oligonucleotides containing selenium on solid support (**25**).

4.5 Design and Incorporation of Methylseleno Uridine and Cytidine into DNA and RNA

The strategy to chemically incorporate 2'-methylseleno uridine and 2'-methylseleno cytidine into DNA and RNA oligonucleotides using the phosphoramidite chemistry involved several criteria. A series of studies were performed to obtain insights about the importance of selenium internal modification as a facilitator for DNA and RNA structural analyses. Firstly, synthetic experiments were performed in order to probe the coupling efficiency of these selenium phosphoramidite analogs using standard protocols (*101*). Secondly, the stability of selenium upon iodine oxidation during the solid phase synthesis was evaluated in connection to its relative position in the oligonucleotide sequence. Thirdly, different HPLC conditions were investigated to isolate highly pure samples for crystallization screening. Next, melting temperature experiments were performed to investigate the relative stability of both DNA and RNA duplexes. Finally, preliminary structural studies were carried out in order to determine if the pucker of the selenium-derivatized furanose was consistent with the unmodified ring.

In order to perform these experiments, a number of biologically and structurally important RNA and DNA oligonucleotides were derivatized with either 2'-methylseleno uridine or 2'-methylseleno cytidine. For decades, biological and structural studies in the field of RNA have attracted lots of attention. Many studies have focused on their structural domains. Two important secondary structures of most RNA domains are the stem and the loop (**See Figure 6, Chapter 1**). Stems and loops are relatively easy to predict in RNA using available software applications. Many loop regions have been found to be of biological functional importance. In order to avoid major structural perturbations that might interfere with RNA function, internal derivatization of RNA oligonucleotides was performed in the stem regions.

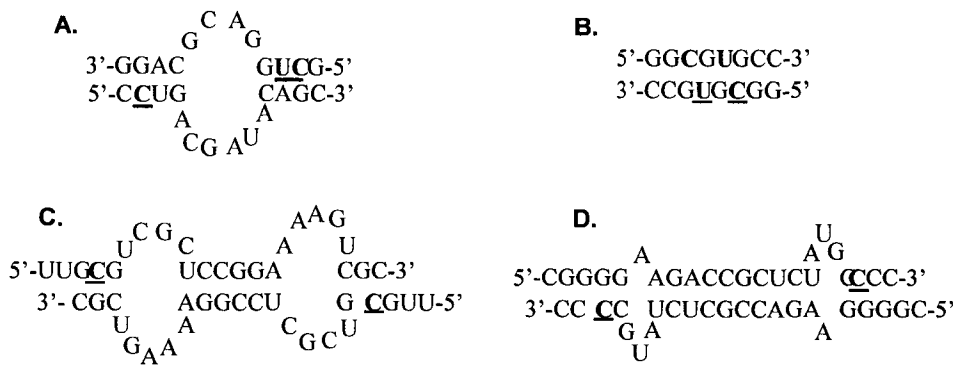


Figure 22. Schematic representation of the secondary structures of selected RNA oligonucleotides for internal selenium derivatization. The nucleotides to be modified with the methyl seleno group are underlined.

To demonstrate this method, several RNA oligonucleotides varying from 8-23 nucleotides were chosen and derivatized at selected positions by the solid phase synthesis. For example, a 12mer RNA fragment of the HIV-1 Rev binding element was modified with both 2'-methylseleno uridine and 2'-methylseleno cytidine at three different positions (**Figure 22A**). Although the crystallization condition of this RNA has been determined, it has not been possible to determine its 3D structure using bromine or other types of derivatives. In addition, a self-complementary RNA oligonucleotide $(GGCGUGCC)_2$ was chosen for selenium derivatization at both U and C sites (**Figure 22B**). This RNA oligomer has been studied as a model to understand G-U wobble base pairing in RNA. Derivatization of this RNA has been attempted previously using other types of heavy atom derivatives without success in the structure determination. Since this is an RNA duplex, it could serve as a model to investigate the thermodynamic stability caused by the 2'-methylseleno modification. Similarly, other RNA oligonucleotides were selected for selenium derivatization at their stem region (**Figure 22C** and **22D**).

The compatibility of 2'-methylseleno uridine and 2'-methylseleno cytidine was also investigated in relation to both structural and functional important DNA oligonucleotides. For example, a decamer DNA duplex, whose crystal structure is known, $(GCGTAU_{Se}ACGC)_2$ was functionalized with selenium (*111*). This DNA also served as a model for structural studies to compare the effect of the selenium modification on A-form DNA duplexes.

Furthermore, other DNA oligonucleotides containing short repeats that are implicated in human genomic diseases were also investigated for selenium derivatization. Two of these DNA containing purine repeats included the oligonucleotides U_{Se}GGAGGAGGAT and U_{Se}AGGAGGAGGAT, and their structures are unknown. The selenium modification was introduced at the last nucleotide at the 5' end to minimize perturbations to the structures. In addition, a purine-rich DNA oligonucleotide ($GGAAGGTU_{Se}TGGGAT$), a Z-DNA ($U_{Se}GCGCA$), an A-form DNA ($GU_{Se}GTACAC$) and a transcription promoter 32mer DNA, containing seven selenium atoms per molecule, were analogously modified with selenium. The chemical structures of these oligonucleotides were confirmed by mass spectroscopy.

4.6 Probing the Coupling Efficiency of Methylseleno Uridine and Cytidine into DNA and RNA

The compatibility of 2'-methylseleno uridine and 2'-methylseleno cytidine with the solid phase synthesis was studied using phosphoramidite chemistry. Several DNA and RNA oligonucleotides were synthesized under standard phosphoramidite conditions containing the selenium functionality internally (or at the 5' terminal end) at different positions. During the early chemical synthesis, it was noticed that the efficiency of incorporation of the selenium phosphoramidite analogs was much lower than that of unmodified DNA and RNA oligonucleotides when 1H-tetrazole was used as the coupling reagent. These preliminary results were first attributed to some possible intramolecular interactions between methylselenide and the phosphorus center of the phosphoramidite moiety.

However, further studies in which a chromatographic step was implemented during the preparation of the phosphoramidite analogs, led to a substantial increase in the coupling yield of selenium incorporation. This result suggested that some technical limitations were responsible for the unsatisfactory yields. The coupling efficiency was further improved by the replacement of 1H-tetrazole

with 5-benzylmercapto-1H-tetrazole (5-BMT). Recent studies have shown that 5-BMT (pK_a = 4.08) is a superior coupling reagent over 1H-tetrazole (pK_a = 4.89) in the chemical synthesis of RNA (*112*).

These new conditions were tested in the synthesis of a self-complementary dodecamer DNA oligonucleotide containing two consecutive thymidines (*113*). We replaced these Ts with 2'-methylseleno uridines or 2'-methylthio uridine and examined the yield of incorporation by analyzing the UV absorption of the DMTr group eluted.

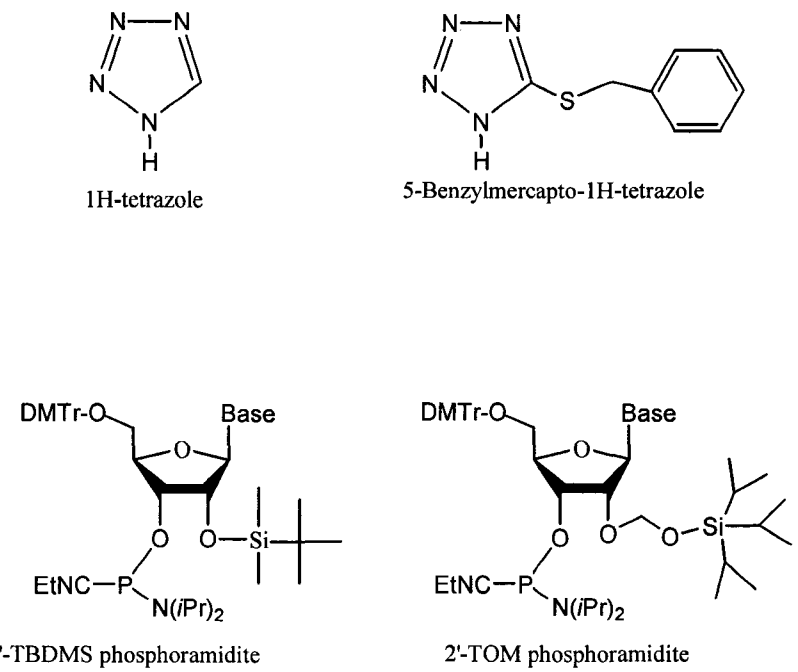
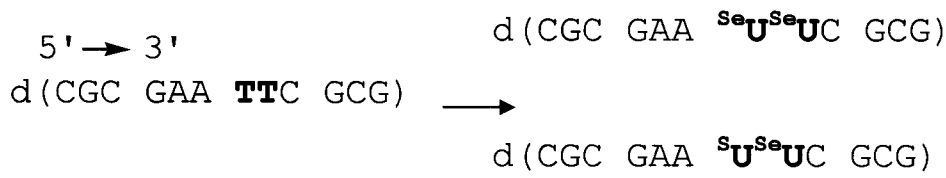


Figure 23. Chemical structures of reagents used in the solid phase synthesis of nucleic acids.



The data indicated no obvious differences in the incorporation efficiency between both analogs. In fact it was observed that the overall yield of each coupling cycle was more than 99% at a coupling time of 25 seconds.

Similarly, the efficiency of the chemical incorporation of 2'-methylseleno uridine was investigated using the aforementioned conditions in the synthesis of the dodecamer Rev binding element of HIV-1 and an octomer RNA of unknown function.

a) Dodecamer RNA: 5′-GC^{Se}U GGA CGC AGG-3′

a) Octomer RNA: 5′-GGC G^{Se}UG CC-3′

Since the synthesis of RNA *via* the phosphoramidite requires 2' protection, the incorporation efficiency was first attempted using 2'-*O*-TBDMS phosphoramidite building blocks in combination with 1H-tetrazole. During the course of preliminary experiments, it was

noticed that the overall yield for the synthesis of the ordinary RNA oligonucleotide was unsatisfactory. As an alternative to 2'-*O*-TBDMS protection, 2'-*O*-TOM phosphoramidite building blocks in conjunction with 5-BMT were studied. The results demonstrated that, under these conditions, as in the case of DNA, the efficiency of each coupling cycle was over 99%. More significantly, it was found that the 2'-methylseleno uridine or the cytosine analog was as reactive as ordinary deoxynucleotide phosphoramidites under the same coupling time, 25 second.

4.7 Effect of Iodine Oxidation on the Methylseleno Functionality

The stability of the methylseleno functionality in DNA and RNA oligonucleotides during the solid phase synthesis was investigated by LC-MS. The oxidation step was originally conducted under the conventional condition of 20 mM of I_2 for 20 seconds. Interestingly, no measurable selenium oxidation was observed in most of the cases. In some situations, however, especially when the methylseleno functionality was close to the 3'-termini, about 2-5% of selenoxide byproducts were observed. These byproducts were identified as compounds displaying an extra 16 Da of mass over the corresponding

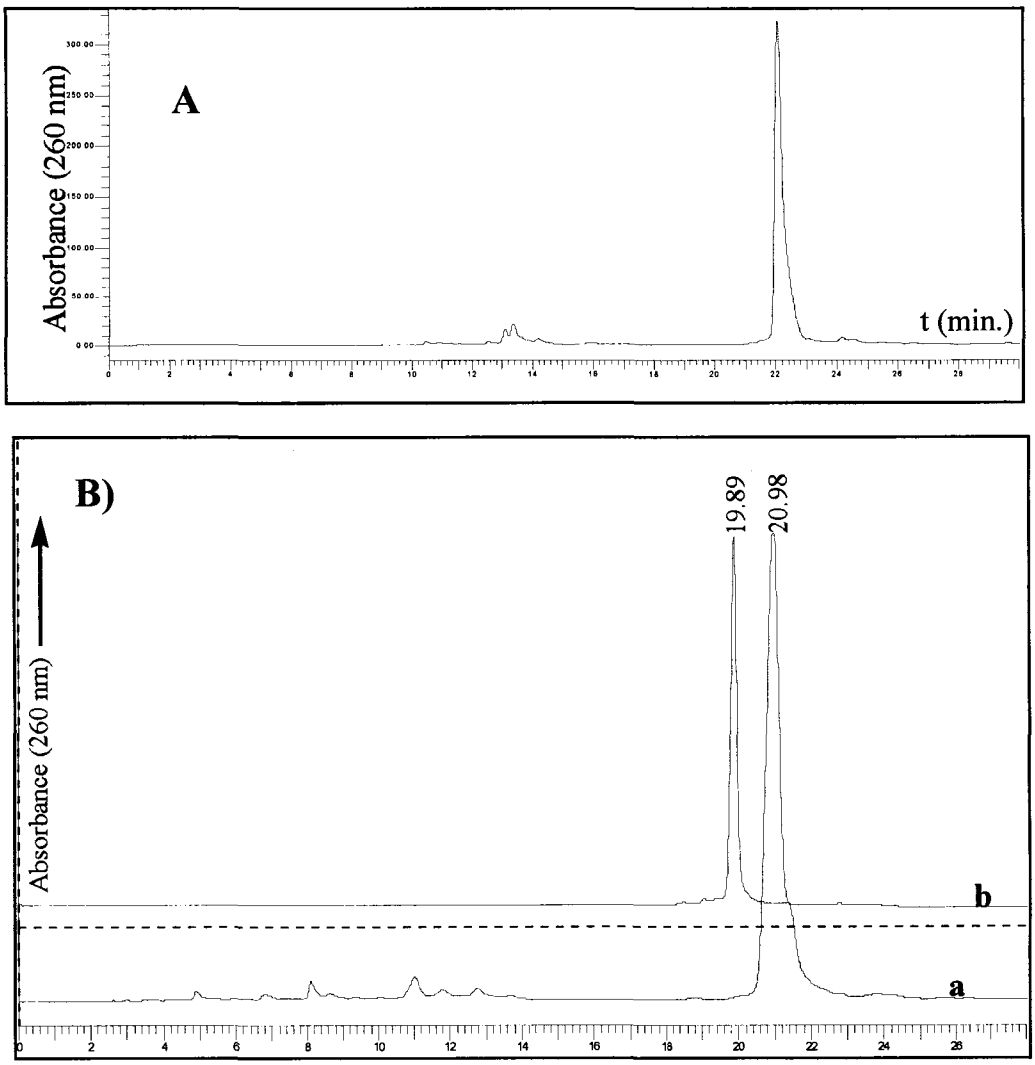


Figure 23. RP-HPLC analyses of RNA oligonucleotides synthesized using 5-BMT and 2'-*O*-TOM protecting groups. A) Se-C RNA 8mer (5'-DMTr-GGC$_{Se}$GUGCC-3'): crude after ammonia treatment and 2'-TOM deprotection, Retention time, 221 min. B) Se-U RNA 8mer (5'-GGCGU$_{Se}$GCC-3'): a) Crude, DMTr-on and 2'-TOM off product; b) Pure, DMTr-off product. Detritylation was carried out for 1.5 min using a 2% solution of trichloroacetic acid. The HPLC condition is described under Material and Method.

selenide on mass spectroscopy. Although we did not investigate the effect of using a lower concentration of I_2 during synthesis, it was

observed that oligonucleotides containing the 2'-seleno functionality were quite stable under air without detectable oxidation or degradation for months.

4.8 RP-HPLC Purification of the Methylseleno DNA and RNA

The DNA and RNA oligonucleotides were analyzed and purified by reverse phase high performance liquid chromatography (RP-HPLC) both DMTr-on and DMTr-off. A two-step purification process was undertaken in order to obtain highly pure oligonucleotides suitable for crystallization screening. The DMTr-on full-length product had a longer retention time on reverse phase columns compared to the truncated DMTr-off fragments generated during the solid phase synthesis. After this purification step the samples were lyophilized to remove all the solvents. It was noticed that partial detritylation of the full-length oligonucleotides usually occurred during lyophilization. RP-HPLC and TLC analyses indicated that about 50% of the oligonucleotides had already lost the DMTr group. The cause of this detritylation was later determined to be due to the generation of acid during lyophilization. It was suggested that this acidic environment

(pH 4) originated from the *in situs* production of acetic acid from the triethylammonium acetate buffer (50 mM).

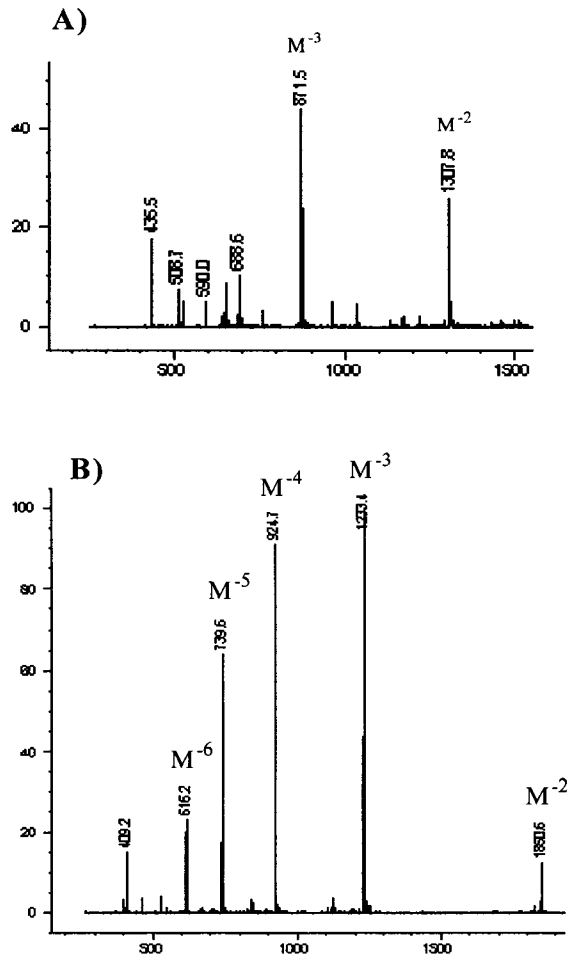


Figure 24. Electrospray MS analysis of 2'-Se-C oligonucleotides. (A). DMTr-off 2'-Se-C RNA8mer ($C_{77}H_{98}N_{31}O_{54}P_7Se$), Isotopic Mass: 2617.3; M^{-2}: 1307.8 (1307.7); M^{-3}: 871.5 (871.4); (B). DMTr-off 2'-Se-C DNA12mer ($C_{119}H_{153}N_{38}O_{73}P_{11}Se$), Isotopic Mass: 3702.6; M^{-2}: 1850.6 (1850.3); M^{-3}: 1233.4 (1233.2); M^{-4}: 924.7 (924.6); M^{-5}: 739.6 (739.5); M^{-6}: 616.2 (616.1).

In addition, it was observed that this acidic environment had a direct effect on the stability of the oligonucleotides during the removal of the trityl group. Severe degradation of both methylseleno DNA

and RNA oligonucleotides occurred when detritylation was carried out using the conventional condition of 2% aqueous trichloroacetic acid for 5 minutes. As a result, RP-HPLC analyses were carried out to find a new optimum condition to perform detritylation. A time course detritylation experiment determined that 1.5 minutes was sufficient to completely remove the DMTr group without causing any oligonucleotide degradation.

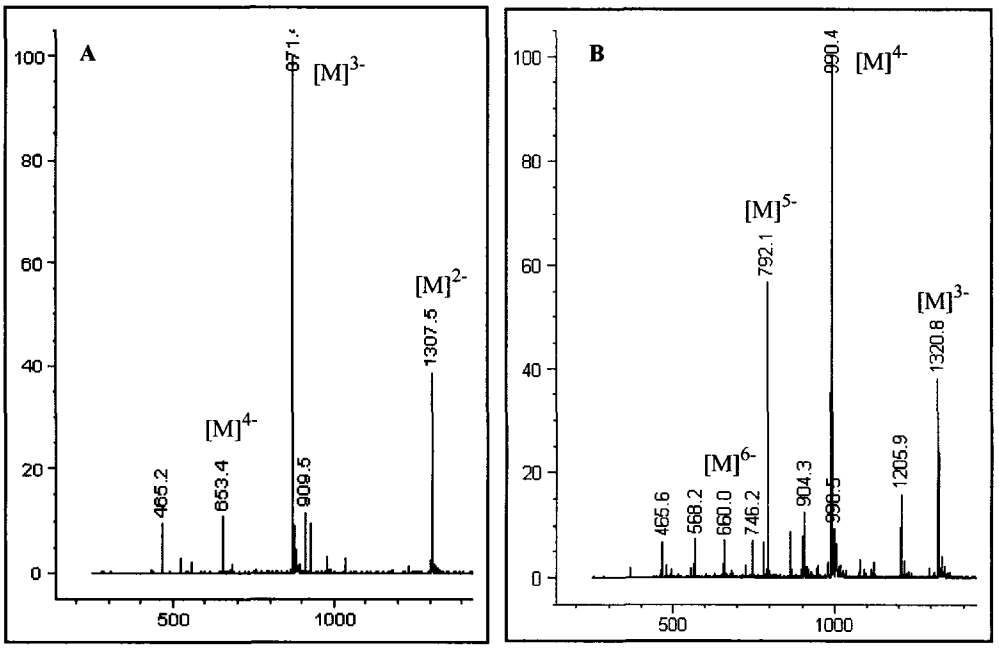


Figure 25. Electrospray MS spectra of the RNAs derivatized with the 2'-selenium Uridine. A) The RNA8mer (GGCGUSeGCC), $C_{77}H_{98}N_{31}O_{54}P_7Se$, Isotopic Mass: 2617.3, measured (calculated) m/z: $[M]^{2-}$: 1307.5 (1307.7); $[M]^{3-}$:871.4 (871.4); $[M]^{4-}$: 653.4 (653.3); B) The RNA12mer (GCUSeGGACGCAGG), $C_{117}H_{146}N_{51}O_{80}P_{11}Se$, Isotopic Mass: 3965.5, measured (calculated) m/z: $[M]^{3-}$: 1320.8 (1320.8); $[M]^{4-}$: 990.4 (990.4); $[M]^{5-}$: 792.1 (792.1); $[M]^{6-}$: 660.0 (659.9).

Table 2. Electrospray analysis of DNA and RNA oligonucleotides containing the 2'-methylseleno funtionality.

RNA & DNA Oligonucleotides	Molecular Formula	LC-MS Measured (Calculated) m/z
5'-GC$\mathbf{U_{Se}}$GGACGCAGG-3' (RNA12mer, REV RNA motif)	$C_{117}H_{146}N_{51}O_{80}P_{11}Se$ Isotopic Mass: 3965.5	M^{3-}: 1320.8 (1320.8); M^{4-}: 990.4 (990.4); M^{5-}: 792.1 (792.1); M^{6-}: 660.0 (659.9).
5'-G$\mathbf{C_{Se}}$UGGACGCAGG-3' (RNA12mer, REV RNA motif)	$C_{117}H_{146}N_{51}O_{80}P_{11}Se$ Isotopic Mass: 3965.5	M^{3-}: 1321.0 (1320.8); M^{4-}: 990.6 (990.4)
5'-C$\mathbf{C_{Se}}$UGACGAUACAGC-3' (RNA14mer, REV RNA motif)	$C_{134}H_{169}N_{54}O_{93}P_{13}Se$ Isotopic Mass: 4504.6	M^{3-}: 1500.8 (1500.5); M^{4-}: 1125.3 (1125.2)
5'-GGCG$\mathbf{U_{Se}}$GCC-3' (RNA8mer)	$C_{77}H_{98}N_{31}O_{54}P_{7}Se$ Isotopic Mass: 2617.3	M^{2-}: 1307.5 (1307.7); M^{3-}: 871.4 (871.4); M^{4-}: 653.4 (653.3)
(5'-GG$\mathbf{C_{Se}}$GUGCC-3') (RNA8mer)	$C_{77}H_{98}N_{31}O_{54}P_{7}Se$ Isotopic Mass: 2617.3	M^{2-}: 1307.8 (1307.7); M^{3-}: 871.5 (871.4)
5'-CGGGGAAGACCGCUCUAUG C$\mathbf{C_{Se}}$C-3'	$C_{219}H_{276}N_{88}O_{157}P_{22}Se$ Isotopic Mass: 7411.0	$[M+NH_3]^{5-}$: 1484.9 (1484.6); M^{6-}: 1234.2 (1234.2)
5'-UUG$\mathbf{C_{Se}}$GUCGCUCCGGAAA AGUCGC-3'	$C_{219}H_{274}N_{86}O_{159}P_{22}Se$ Isotopic Mass: 7412.9	M^{3-}: 2470.2 (2470.0)
5'-$\mathbf{U_{Se}}$GGAGGAGGAT-3' (DNA GGA repeat)	$C_{110}H_{135}N_{49}O_{63}P_{10}Se$ Isotopic Mass: 3539.5	M^{2-}: 1768.7 (1768.8); M^{3-}: 1179.0 (1178.8); M^{4-}: 883.8 (883.9); M^{5-}: 707.0 (706.9); M^{6-}: 588.9 (588.9).
5'-$\mathbf{U_{Se}}$AGGAGGAGGAT-3' (DNA AGG repeat)	$C_{120}H_{147}N_{54}O_{68}P_{11}Se$ Isotopic Mass: 3852.6	M^{2-}: 1925.3 (1925.3); M^{3-}: 1283.3 (1283.2); M^{4-}: 962.2 (962.2); M^{5-}: 769.5 (769.5); M^{6-}: 641.0 (641.1).
5'-GGAAGGTU$\mathbf{_{Se}}$TGGGAT-3'	$C_{140}H_{173}N_{58}O_{83}P_{13}Se$ Isotopic Mass: 4476.7	M^{3-}: 1491.4 (1491.2); M^{4-}: 1118.2 (1118.2); M^{5-}: 894.5 (894.3); M^{6-}: 745.2 (745.1).
5'-G$\mathbf{U_{Se}}$GTACAC-3' (A-form, 2'-MeSe-DNA)	$C_{78}H_{99}N_{30}O_{46}P_{7}Se$ Isotopic Mass: 2488.4	M^{2-}: 1243.2 (1243.2); M^{3-}: 828.5 (828.5); M^{4-}: 621.3 (621.1).
5'-$\mathbf{U_{Se}}$GCGCA-3' (Z-DNA)	$C_{58}H_{74}N_{23}O_{34}P_{5}Se$ Isotopic Mass: 1871.3	M^{-}: 1870.3 (1870.3); M^{2-}: 934.6 (934.7)
5'-ATTCAGC$\mathbf{_{Se}}$G-3' (DNA 8mer)	$C_{79}H_{101}N_{30}O_{46}P_{7}Se$ Isotopic Mass: 2502.4	M^{2-}: 1250.2 (1250.2); M^{3-}: 833.2 (833.1); M^{4-}: 624.7 (624.6); M^{5-}: 499.5 (499.5)

4.9 Effect of the Methylseleno group on the Sugar Conformation

The effect of the 2'-methylselno group on the sugar pucker was studied using an A-form DNA duplex of known structure (*111*). This DNA duplex was a decamer with the nucleotide sequence of $d(GCGTATACGC)_2$. The design involved the replacement of one of the thymidine nucleotides with methylseleno uridine. Hence, $d(GCGTAU_{Se}ACGC)_2$ was synthesized on solid support and crystallized.

Several important pieces of information were obtained. First of all, it was found that the crystals of the selenium derivatized DNA duplex was isomorphous to that of the ordinary DNA duplex dodecamer. When the X-ray crystal structure of the selenium DNA derivative was compared to that of the ordinary duplex DNA, it was revealed that the methylseleno functionality at the 2' position did not cause any distortion to the geometry of the A-form duplex DNA. In fact, the furanose containing the 2'-methylseleno group adopted a C_3'-*endo* conformation (**Figure 26**). This geometry was consistent with that adopted by the ordinary A-form DNA duplex (*100*). These results conclusively stablished the suitability of the 2'-methylseno functionality as a facilitator for X-ray crystal structure determination

of RNA and A-form DNA duplexes. Since the sugar puckering of B-DNA duplexes has a $C_{2'}$-*endo* conformation, this derivatization strategy would not be suitable for this type of DNA, though.

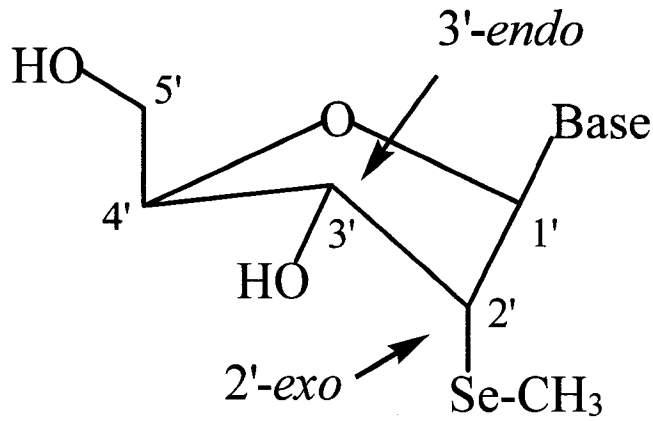


Figure 26. Effect of the 2'-methylseleno functionality on the sugar pucker of an A-form DNA. The 2'-selenium functionality did not alter the conformation of the sugar ring (also see Figure 4 in Chapter 1).

4.10 Thermodenaturation and Stability of Oligonucleotide Duplexes Containing the 2'-Methylseleno Functionality

Thermodenaturation experiments were also designed to study the effect of the 2'-methylseleno functionality on both duplex DNA and RNA oligonucleotides. The melting temperatures (T_m's) of a series of DNA and RNA oligonucleotide duplexes containing the selenium functionality at various positions were measured and their T_m's were

compared to those of the corresponding native duplexes and other analogs that contained the 2'-methoxy modification.

The results of these experiments are listed in **Table 3** and **4.** It was found that the incorporation of either the 2'-methylseleno uridine or 2'-methylselno cytidine has little effect on the thermodynamic stability of the DNA and RNA duplexes. In fact, in the case of the DNA duplexes, the T_m values of the selenium derivatives were slightly higher than those of the native duplexes. This increase in thermodynamic stability has also been observed in the case of 2'-methoxy DNA duplexes (*114*). The RNA duplexes showed a slight decrease in their T_m compared to the corresponding native duplexes. But when compared to the 2'-methoxy analogs, the T_m's of the 2'-methyseleno RNA duplexes were within the same values. This decrease in the T_m in RNA containing modifications at the 2' position suggested that the 2'-hydroxyl group in the RNAs might be involved in some type of stabilizing interactions. It is concluded, therefore, that this type of selenium derivatization strategy should be suitable for X-ray crystallographic studies of A-form DNA and RNA oligonucleotides.

Table 3. Melting temperatures of 2'-Se-RNA duplexes	
RNA 8mer	Melting temperatures (°C)
Native RNA 8mer (5'-GGCGUGCC-3')	36.6
2'-MeO RNA (5'-GGC$_{OMe}$GUGCC-3')	32.9
2'-MeSe RNA (5'-GGC$_{SeMe}$GUGCC-3')	31.7
2'-MeSe RNA (5'-GGCGU$_{SeMe}$GCC-3')	32.4

Table 4. Melting temperatures of 2'-Se-DNA duplexes	
DNA duplexes	Melting temperatures (°C)
Native octamer (5'-GTGTACAC-3')	21.2
2'-MeO-octamer (5'-GU$_{OMe}$GTACAC-3')	24.8
2'-MeSe-octamer (5'-GU$_{SeMe}$GTACAC-3')	21.5

4.11 Conclusions

A synthetic strategy was developed to incorporate the selenium functionality at internal and selective positions of DNA and RNA oligonucleotides using phosphoramidite chemistry. The method was based on the synthesis of 2'-methylseleno uridine and 2'-methyseleno cytidine. The synthesis of 2'-methylseleno uridine was accomplished

via nucleophilic displacement of the base-promoted 2,2'-*O*-anhydro nucleoside intermediate with sodium methylselenide. Similarly, the synthesis of 2'-methyseleno cytidine was carried out by conversion of 2'-methyseleno uridine into the corresponding cytidine analog *via* a triazolide intermediate.

The suitability of the phosphoramidite derivatives of 2'-methyseleno uridine and 2'-methyseleno cytidine was investigated in the solid phase synthesis of several DNA and RNA oligonucleotides. The coupling efficiency of these nucleoside derivatives was studied with the coupling reagents 1H-tetrazole and 5-benzylmercapto 1H-tetrazole in connection with 2'-*O*-TMDMS and 2'-*O*-TOM protected nucleoside building blocks. The results demonstrated that 5-benzylmercapto 1H-tetrazole and 2'-*O*-TOM protected nucleoside building blocks provided optimum conditions similar to that of the corresponding unmodified nucleosides. It was also illustrated that the methylseleno functionality can resist multiple cycles of the conventional iodine treatment with only minor oxidation.

More importantly, studies designed to investigate the effect of the methylseleno functionality on the conformation of the furanose ring concluded that the selenium modification does not cause any major

perturbation to A-form DNA and RNA oligonucleotides. The furanose derivatized with the 2'-methylseleno group adopted the C_3'-*endo* conformation, consistent with the stereochemistry of A-form DNA and RNA. This finding was also in agreement with the fact that the selenium modification only caused minor changes to the thermodynamic stability of both A-form and RNA oligonucleotide duplexes.

The results presented in this chapter provide strong evidences in support of the internal derivatization of DNA and RNA oligonucleotides with selenium using phosphoramidite chemistry. This chemical approach, however, only facilitates the synthesis of DNA and RNA oligonucleotides. The limit for large-scale DNA synthesis is about 100 nucleotides, and that for RNA is about 40 nucleotides. This limit is dictated by the coupling efficiency of each step in the synthesis of nucleic acid oligonucleotides on solid support using the phosphoramidite chemistry.

The above results has been published in the following references (*100, 108, 115, 116*).

In order to further explore the potential for selenium as a facilitator of nucleic acid X-ray crystallography, it would be ideal to develop a

new synthetic strategy that would allow the preparation of long DNA and RNA molecules. The next chapter addresses the development of a new series of nucleotide analogs containing selenium for the enzymatic synthesis of DNA and RNA.

Chapter 5

SYNTHESIS OF NUCLEOSIDE TRIPHOSPHATE CONTAINING SELENIUM AT THE α-PHOSPHATE

5.1 Introduction

The results presented in the previous chapters illustrated the suitability of selenium to prepare DNA and RNA oligonucleotides on solid support *via* the phosphoramidite chemistry. Although the chemical synthesis of DNA and RNA containing selenium has great benefits, this method has limited applications for structural studies. The limitation is that this method only allows large-scale synthesis of DNA and RNA oligonucleotides, because the synthetic yield is exponentially dependent on the coupling efficiency of each coupling cycle. In the case of RNA, for example, the limit is about 40 nucleotides long for a satisfactory synthetic yield.

Since the length of most biological important DNA and RNA molecules exceeds this size, new approaches need to be developed to prepare long nucleic acid molecules harboring the selenium functionality. To accomplish this objective, an enzymatic approach to introduce the selenium moiety was undertaken. This approach was based on the synthesis of nucleoside triphosphates containing

selenium at specific positions to avoid major local structural perturbations that might prevent enzymatic recognition. Hence, the replacement of one of the non-bridging oxygen atoms at the α-phosphate group was investigated (**Figure 27**). Selenium replacement at this position was studied using both a nucleoside H-phosphonate and a nucleoside phosphite as the immediate intermediates. The structures of the nucleoside triphosphates containing selenium were confirmed by spectroscopic data, including ^{31}P-NMR and mass spectroscopy.

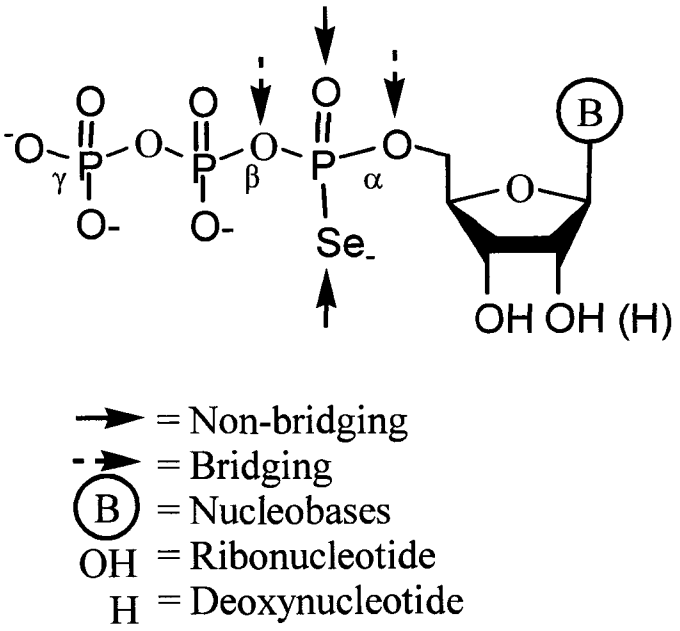


Figure 27. Replacement of a non-bridging oxygen atom by selenium at the α-phosphate position in nucleoside triphosphates.

5.2 Selecting a Suitable Selenium Reagent

The synthesis of nucleoside triphosphate containing selenium at the α-phosphate was first attempted using an H-phosphonate intermediate. The use of this intermediate has been previously reported to introduce the thiol and the borane groups into the phosphate moiety in nucleotides by the replacement of one of the non-bridging oxygen atoms (*117-119*). In our situation, the design involved the synthesis of the nucleoside H-phosphonate at the 5' position, oxidation of the trivalent phosphorus center with an electrophilic source of selenium, followed by a coupling reaction with pyrophosphate. The thymidine and the uridine nucleosides were selected as synthetic models for their convenience to develop the synthesis.

The search for a suitable electrophilic selenium reagent became the main focus of the investigation. First of all, the use of the conventional procedure to make the ordinary nucleoside triphosphates and the corresponding thiol analogs was impractical because the corresponding selenium reagent (**28**) used in these reactions is not readily available (**Figure 28**). To achieve the synthesis of natural nucleoside triphosphates and the thiol analogs of **39**, **26** or **27** is reacted with nucleoside **35**, respectively (*120-123*).

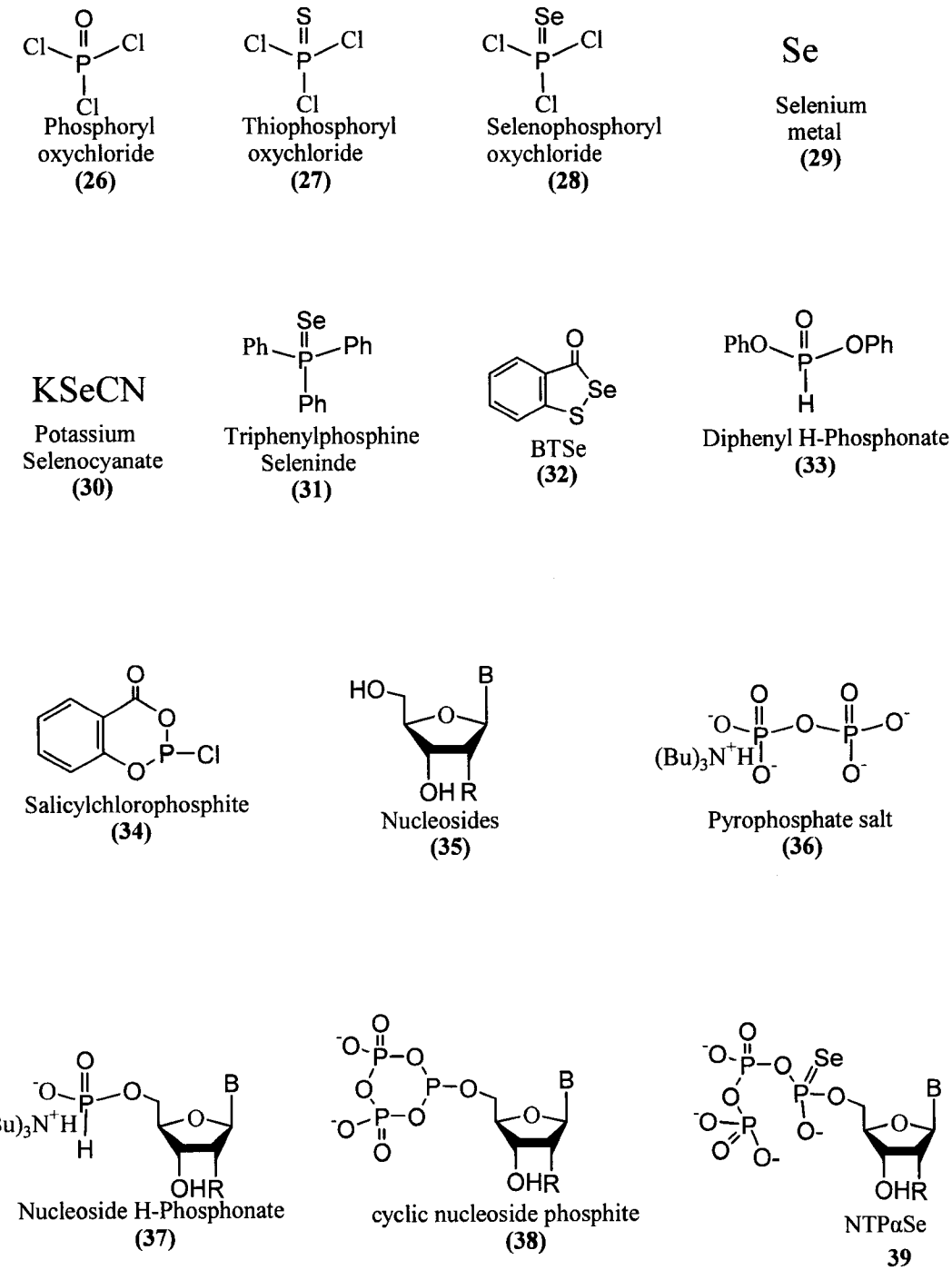


Figure 28. Important reagents and intermediates considered for the incorporation of selenium at the α-phosphate position of nucleoside triphosphates.

Among the electrophilic selenizing reagents that were initially utilized to introduce selenium to the H-phosphonate intermediate **37** included selenium metal (**29**) and potassium selenocyanate (**30**). Although the utilization of these reagents has been reported in similar chemistry (*124, 125*), they turned out to be impractical under our experimental conditions; heterogeneous reaction mixtures and long refluxing time above 100°C usually led to the production of side reactions and very little of the desired compounds.

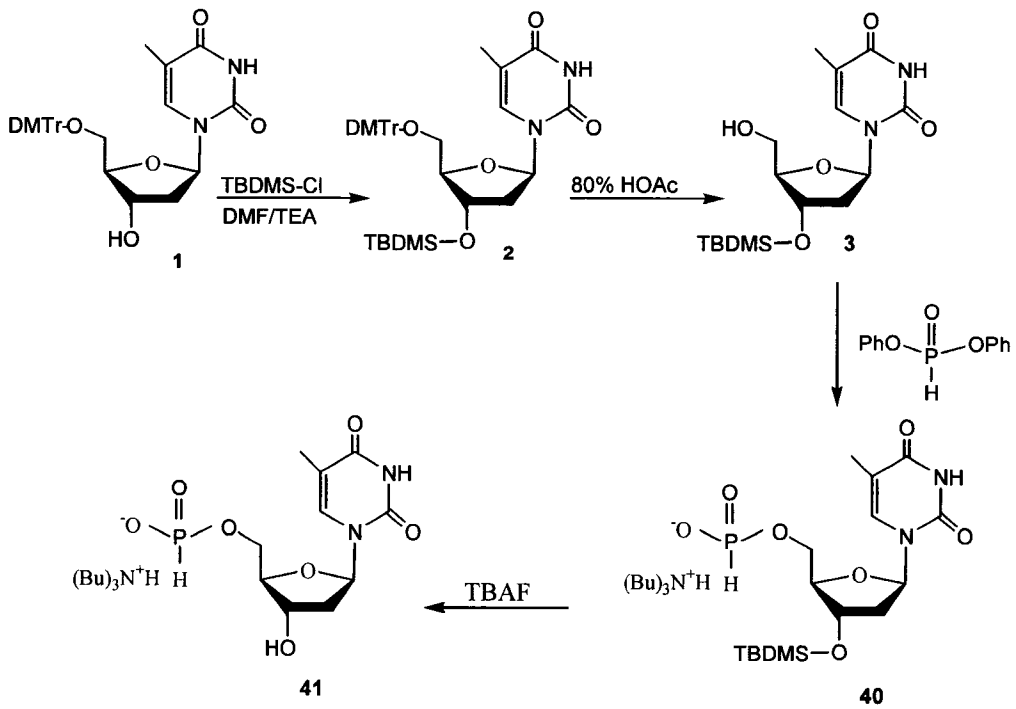
These preliminary results provided some insights about the search for a more versatile selenium reagent. A suitable reagent would have to be soluble in most organic solvents, stable, and most important, reactive with the H-phosphonate intermediate (**Table 5**). In this way, 3H-1,2-benzothiaselenol-3-one (BTSe, **32**) became a leading candidate. BTSe is the selenium analog of the Beaucage reagent. It was developed by Jacek Stawinski for the conversion of nucleoside H-phosphonates and nucleoside H-phosphonothioates diesters into the corresponding phosphoroselenoate and phosphorothioselenoate diesters containing selenium at the non-bridging position (*126, 127*). The selenization reaction was found to be stereospecific and occurred with retention of configuration at the phosphorus center.

Table 5. Relative solubility of electrophilic selenium reagents in organic solvents			
Se Reagents	Structures	Solubility	References
Se metal	Se_8	poor	(*128*)
Potassium selenocyanide	KSeCN	poor	(*125*)
BTSe	Fig. 27	good	(*126, 127*)
Triphenylphosphine selenide	Fig. 27	good	(*129*)

5.3 Preparation of Nucleoside Precursors

The synthesis of the H-phosphonate intermediate was carried out using the procedures developed by Stawinski (*118*). In a typical procedure, the synthesis of the thymidine H-phosphonate intermediate **41** was accomplished by reacting the partially protected **1** in DMF with four molar equivalents of *tert*-butyldimethylsilyl chloride (TBMS-Cl) in the presence of triethylamine as a base (**Scheme 5**). Without purification, the stable intermediate **2** was treated with 80% acetic acid to selectively remove the 5'-DMTr protection group. Chromatographic purification of the crude product recovered from the acid-promoted hydrolysis afforded **3** in nearly quantitative yield (95%).

Compound **3** was subsequently reacted with three molar excess of diphenyl phosphite (**33**) in dry pyridine following published



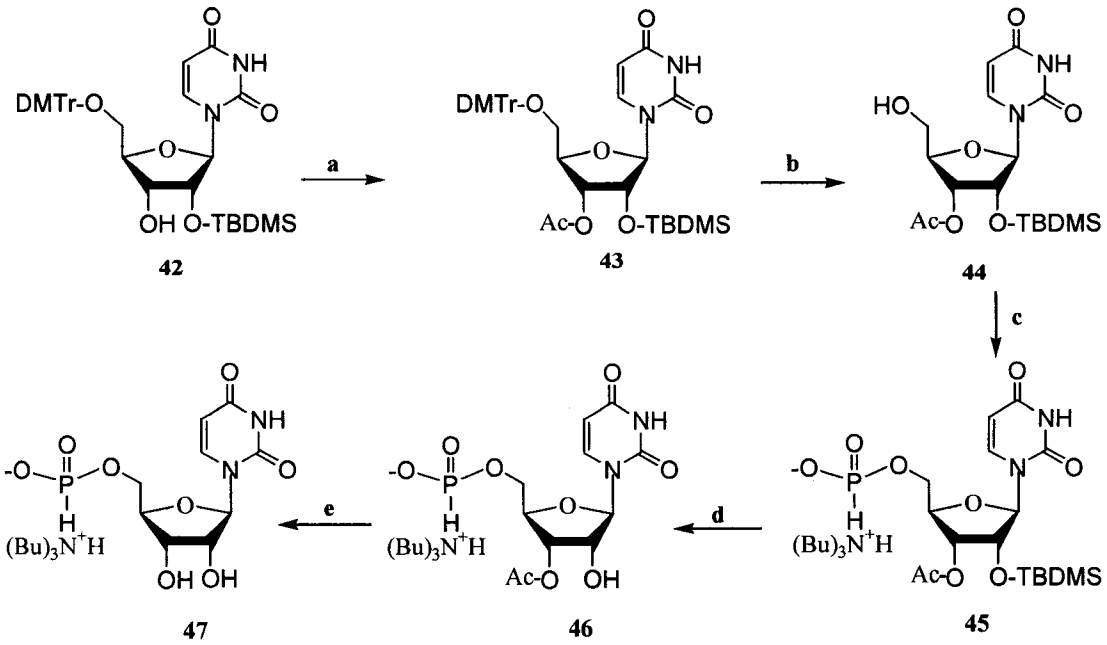
Scheme 5. Synthesis of thymidine 5'-H-phosphonate.

procedures (*118*). After quenching the reaction with a 1:1 solution of tributylamine and water, the tributylammonium salt of the crude product was purified on silica gel column chromatography to obtain the desirable H-phosphonate intermediate **40** in good yield (87%). Pure **40** was characterized spectrocospically by ^{1}H-NMR and ^{13}C-NMR. The compound produced two typical symmetrical peak characteristic of the ^{31}P-^{1}H coupling, which resonated at 5.72 ppm and 8.12 ppm, respectively, with a J_{PH} coupling constant of 600 Hz.

The corresponding uridine H-phosphonate analog was also successfully prepared with the inclusion of some extra synthetic steps to account for the additional hydroxyl group present at the 2'-postion of the uridine nucleoside. **Scheme 6** shows a general and convenient route to synthesize this compound. The synthesis of UTPαSe started from partially protected **42,** which is commercially available (ChemGenes Corporation), followed by acetylation of the 3'-moiety with acetic anhydride in the presence of 4-dimethylaminopyridine as a catalyst. Experiments indicated that the presence of the catalyst was required for acetylation to occur. The removal of the 5'-DMTr group, followed by the subsequent incorporation of the H-phosphonate functionality to give **45**, was carried out as previously mentioned in the preparation of the corresponding thymidine analog.

5.4 Incorporation of Selenium *via* an H-Phosphonate Intermediate

Initial attempts to replace one of the non-bridging oxygen atoms at the α-phosphate group by selenium *via* an H-phosphonate intermediate utilized partially protected thymidine **40** and uridine **45** (**Scheme 6** and **Scheme 5**). When the synthesis was carried out before acetyl or TBDMS deprotection, partial degradation of the product usually



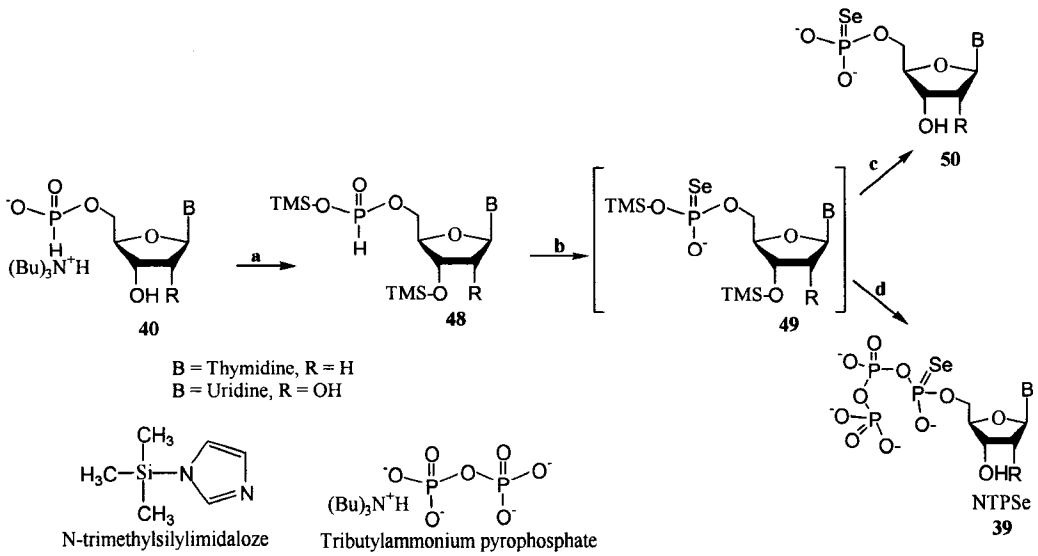
Scheme 6. Preparation of the uridine H-phosphonate intermediate. Reagents: a) acetic anhydride, 4-dimethylaminopyridine, and THF; b) 80 % HOAc; c) *i:* diphenylphosphite and pyridine; *ii:* $(Bu)_3N$: H_2O; d) $(Bu)_4N^+F^-$, THF; e) NH_3 and methanol.

occurred during NH_3 or $(Bu)_4N^+F^-$ treatment. To alleviate this problem, deprotection was performed before the introduction of the selenium functionality. In the case of the uridine analog **45**, the deprotection step proved to be very slow and further interventions were necessary: 90 hr of NH_3 treatment in methanol failed to remove the acetyl protecting group at the 3'-position (**Scheme 6**). This difficulty was overcome by first removing the TBDMS functionality

before carrying out the de-acetylation step. Under this condition the acetyl group was removed in less than three hours.

In addition, these deprotection procedures to afford either the corresponding uridine or thymidine derivative **40** created additional concerns in the design of the subsequent steps to introduce selenium. Since a key step in this reaction involved the activation of the phosphorus center with a reactive condensing reagent, the presence of the free and reactive hydroxyl groups could lead to the production of side reactions. The selection of a good condensing agent to couple **40** with pyrophosphate was based on whether it could serve the dual purposes of activation and transient protection, simultaneously. In this way, N-trimethylsilylimidazole was selected. As illustrated in **Scheme 7**, compound **48** was prepared with a sixfold molar excess of TMS-imidazole in a 0.05 M of freshly distilled 1:4 pyridine-methylene chloride solution. The reaction was completed in 30 minutes as evidenced by comparative TLC analysis (30% methanol/methylene chloride).

In one-pot reaction, the selenium functionality was introduced by adding three molar equivalents of BTSe and 10 molar equivalents of triethylamine.



Scheme 7: A general route for the introduction of selenium at the α-phosphate of nucleoside; shown here is only the route for the synthesis of the thymidine analog. Reagents: a) trimethylsilane imidaloze and pyridine; b) BTSe, TEA, and pyridine; c) H_2O; d) tributylammonium pyrophosphate and pyridine, then H_2O.

The reaction was run under argon for 2 hours at room temperature. Half of the reaction mixture was then added to a separate solution containing a three molar excess of tributylammonium pyrophosphate dissolved in pyridine to offer compound **39** in very low yield after water quenching (evidenced by comparative TLC analysis); the other half was quenched with water to give compound **50** also in poor moderate yield after anion exchanger chromatography.

5.5 Mass Spectroscopic Analysis

Compound **50** was analyzed by electrospray mass spectroscopy. Selenium-containing compounds are usually recognized in mass spectra studies from the characteristic groups of peaks resulting from the typical isotopic distribution of naturally occurring selenium. There are six isotopes of most natural abundance; ^{80}Se (49.6%), ^{78}Se (23.5%), ^{82}Se (9.4%), and ^{76}Se (9.0%) are the major ones. **Figure 29** illustrates the mass spectroscopic analysis after the synthesis of the thymidine analog of **50** harboring the 3'-TBDM protecting group. This spectrum clearly indicates the pattern of two selenium atoms being incorporated into the nucleotide. The cluster of peaks from *m/z* 597-608 matches a molecule with an empirical formula ($C_{16}H_{29}O_7N_2PSe_2SiNa$). There is also a series of peaks that differ from the first one and from each other by the *m/z* of a sodium atom.

In order to understand this data, it is important to outline the mechanism of selenium incorporation. This mechanism is shown in **Figure 30**. As proposed by Stawinski, in this base-promoted reaction, the selenium atom in BTSe acts as the electrophilic center, while the phosphorus atom in the nucleoside H-phosphonate acts as the

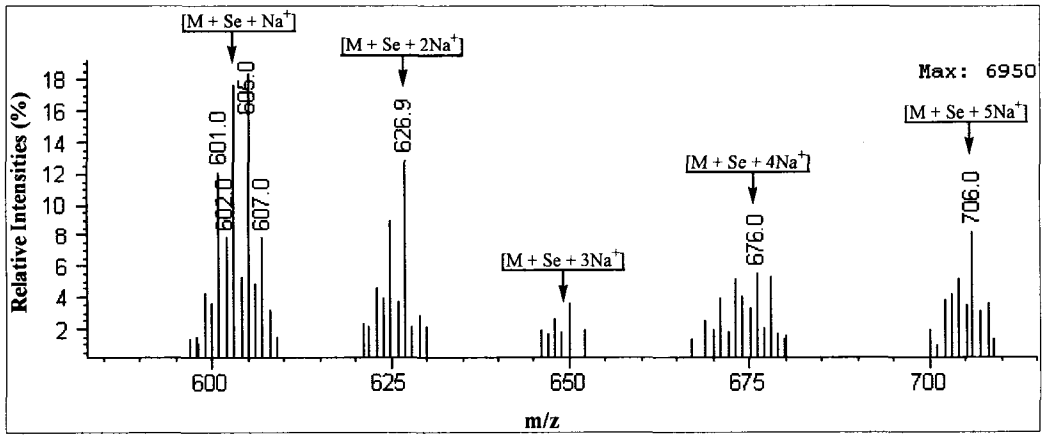


Figure 29: Electrospray mass spectrum of **50** carrying the TBDMS protecting group at the 3'-position, with an empirical formula of $C_{16}H_{29}O_7N_2PSeSi$. The HPLC retention time is 12.15 min. The molecular weight (M) for the adjustment of the isotope of ^{80}Se is 500.

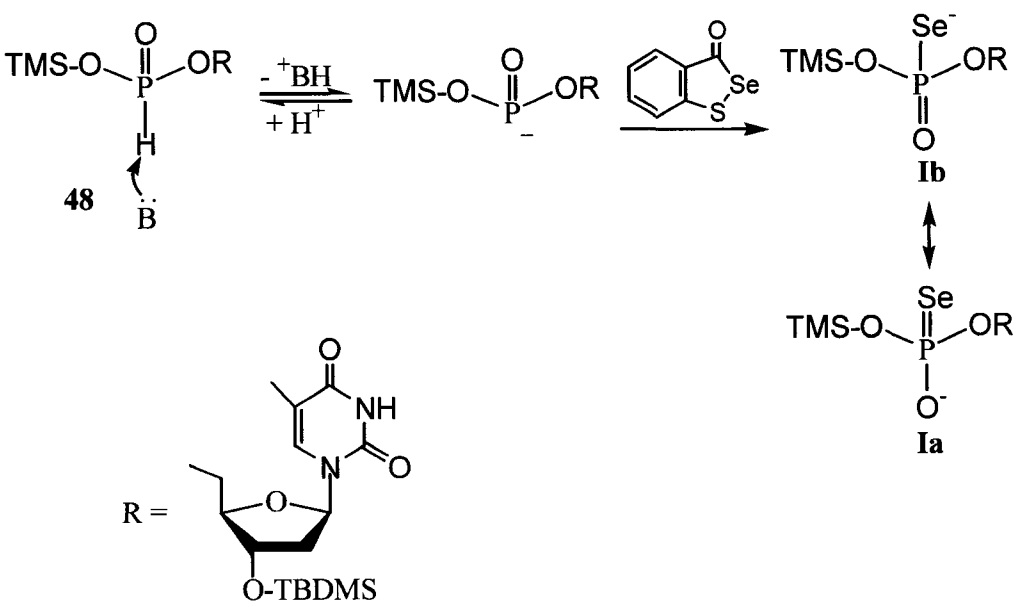


Figure 30. General mechanism of selenium incorporation *via* an H-phosphonate intermediate using BTSe as the selenium donor agent (*127, 130*).

nucleophilic center, to afford the resonance stabilized intermediate **Ia** and **Ib** (*127, 130*). **Ib** may be the dominant resonance contributor. The formation of this intermediate is supported by recent X-ray crystallographic data of a DNA oligonucleotide containing selenium at the non-bridging position in the phosphate backbone, in which selenium appears to be singly bounded to phosphorus (*131*). The reason why **Ib** is more stable than **Ia** may be due to the superior stability of P=O over P=Se. In addition, it has been shown that the phosphorus-selenium bond may forms a resonance hybrid of –P=Se and $-P^{+}-Se^{-}$, when the phosphorus atom carries three others single-bonded substituents (*132*).

As illustrated in **Figure 31**, **Ib** can lead toward the formation of the nucleoside monophosphate containing two selenium atoms *via* the formation of a diselenide intermediate in the presence of molecular oxygen. This diselenide intermediate gives a distinctive orange color that disappears in the presence of DTT. Furthermore, previous literature reports have found that selenium at the non-bridging position in the phosphate backbone gets oxidized with a half-life of 30 days (*125*). This phenomenon has been known as "selenium washout." Although the mechanism for "selenium washout" is

unknown, the data presented here suggest that formation of **Ic** might be the key step for selenium oxidation. **Ic** can further get oxidized to produce compound **50**, which in turn can get oxidized to nucleoside monophosphate **51** and selenium metal.

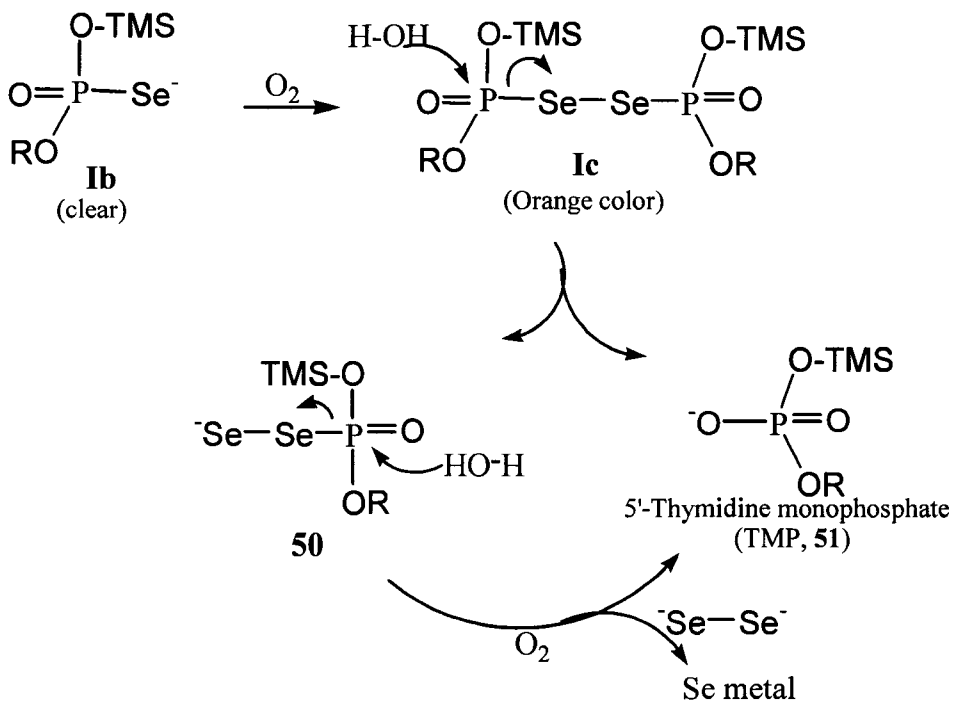


Figure 31: Possible mechanism for "selenium washout" in the presence of oxygen *via* a P-Se-Se-P specie.

5.6 Introduction of Selenium *via* a Phosphite Intermediate

Initial attempt to synthesize nucleoside triphosphates carrying selenium at the α-phosphate *via* the nucleoside H-phosphonate intermediate was permeated with certain limitations and disadvantages. Although this method can be used to synthesize NTPαSes, it involves two ion-exchange separation steps and a condensation reaction that may not be specific. These limitations suggested the exploration of an alternative approach to prepare the NTPαSe analogs. This approach exploits a nucleoside phosphite intermediate that can be readily synthesized following the procedures developed by Ludwig and Eckstein for the synthesis of NTPαS (*133*). In addition, this method has been successfully extended to the synthesis of NTPαB (*134-136*). The chemistry is centered on the utilization of 2-chloro-4H-1,3,2-benzodioxaphosphorin-4-one, commonly known as salicylchlorophosphite (**Figure 32**). Salicylchlorophosphite is capable of undergoing three nucleophilic displacement reactions. This reagent is not selective, however, and the synthesis requires 2' and 3' protected nucleosides; it reacts very rapidly with any hydroxyl group or water. Protection on the base of the nucleosides is not necessary, however (*133*). To develop a

synthetic procedure with this reagent, 3'-acetylated thymidine was used **(Scheme 8)**.

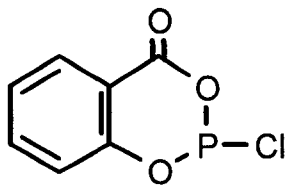
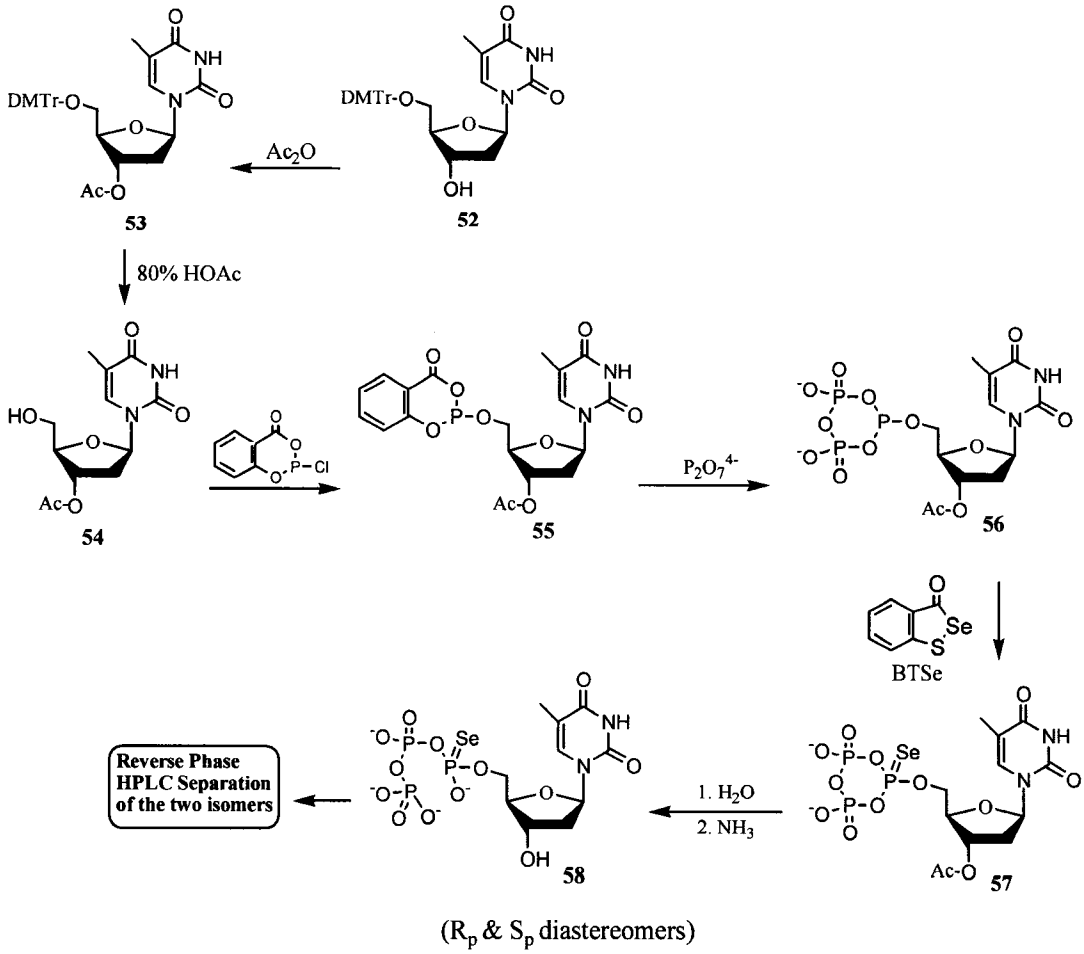


Figure 32. Chemical structure of salicylchlorophosphite.

The synthesis of the selenium analog of thymidine triphosphate was accomplished in one-pot involving 5 synthetic steps. The series of reactions are illustrated in **Scheme 8**. In a typical procedure, the synthesis started with partially protected **52**, which is commercially available (ChemGenes Corporation). Briefly, acetylation of the 3'-moiety was performed using a twofold excess of acetic anhydride in a 0.1 M solution of dry THF in the presence of triethylamine and a catalytic amount of 4-dimethylaminopyridine. Without purification, **53** was treated with 80% acetic acid to remove the 5'-DMTr protecting group.



Scheme 8. Synthesis of TTPαSe *via* a phosphite intermediate.

Chromatographic purification afforded **54** in 87% yield. Compound **54** was subsequently phosphitylated with 1.2 molar equivalents of salicylchlorophosphite in a mixture of dry dioxane and pyridine (3:1 ratio) under dry nitrogen at room temperature. Investigation of the reaction progress with ^{31}P-NMR indicated the formation of a doublet centered around 127 ppm that corresponded with the formation of

intermediate **55**. In one-pot reaction, 1.5 molar equivalents of tributylammonium pyrophosphate dissolved in a 0.5 M of dry DMF containing 3 equivalents of tributylamine were added to the reaction mixture to form the key cyclic nucleoside [α-P]phosphite intermediate **56**. Formation of intermediate **56** was indicated by the appearance of a triplet centered at 106 ppm in the ^{31}P-NMR spectra and a doublet at −18 ppm (**Figure 33**).

Oxidation of the trivalent phosphorus center of intermediate **56** with selenium became the critical step in the pursuit of the synthesis of TTPαSe. Among the electrophilic selenium reagents (**Table 5** and **Figure 28**), 3H-1,2-benzothiaselenol-3-one was investigated based on our previous experience with the H-phosphonate approach. In addition, knowledge of the reaction mechanism of BTSe (*127, 130*) prompted the study of trimethylsilyl chloride (TMS-Cl) and N-trimethylsilyl imidazole (TMS-Im) as possible activating reagents. Although the mechanism of how these reagents activate the phosphite intermediate is not very clear, it is possible that activation might occur by silylation of the non-bridging, negatively charged oxygen atoms of the phosphorus groups.

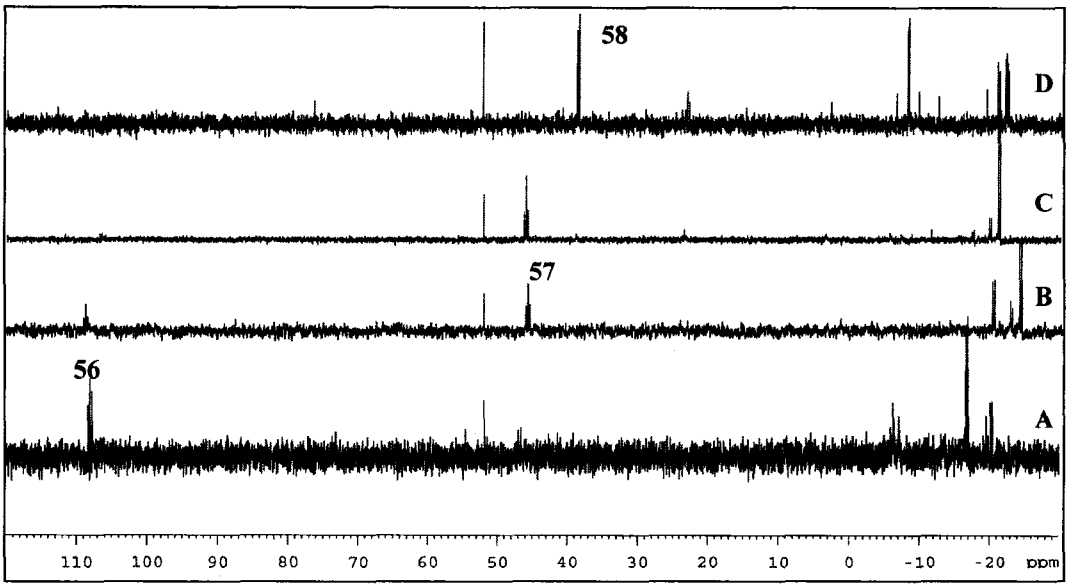


Figure 33. ^{31}P-NMR analysis of the formation of TTPαSe. A) 15 min after adding BTSe; B) 30 min after adding BTSe; C) Reaction quenched with water; D) 1 hr after adding water. The experiment was carried out in the presence of TMS-Cl.

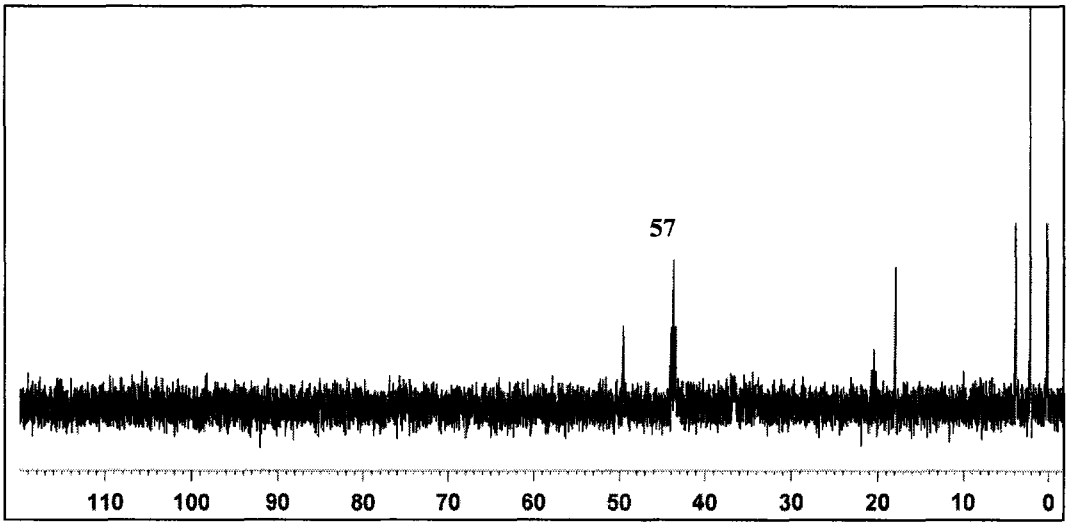


Figure 34. ^{31}P-NMR analysis of BTSe selenization after 15 min of adding water. TMS-Cl was included in the reaction mixture

^{31}P-NMR investigation of the course of selenization with BTSe alone indicated that the reaction was slow. The reaction was not completed after two hours. However, when TMS-Cl was included, the reaction was usually finished in 0.5-1.0 hr. TMS-Im was found to be a superior reagent over TMS-Cl. For example, 2 molar equivalents of TMS-Im promoted complete oxidation in less than 10 minutes (the time required to obtain the ^{31}P-NMR spectrum). However, the ^{31}P-NMR spectra also indicated a significant formation of a byproduct that resonated at 24 ppm. Although the formation of this byproduct was reduced when only 0.5 molar equivalents of TMS-Im were used, the yield was unsatisfactory.

In a practical approach to carry out the selenization reaction, a dioxane solution containing 1.5 molar equivalents of BTSe, 3 molar equivalent of TMS-Cl and 1.0 molar equivalent of triethylamine was added to the reaction mixture of **56 (Scheme 6)**. The complete disappearance of the triplet at 106 ppm and the concomitant appearance of a triplet at about 46 ppm was an indication of the formation of the cyclic intermediate **57**. Water was then added to produce **58** still containing the acetyl protection as a diastereomeric

mixture of Sp and Rp (**Figure 33D**). Ammonia was subsequently added to remove the acetyl functionality.

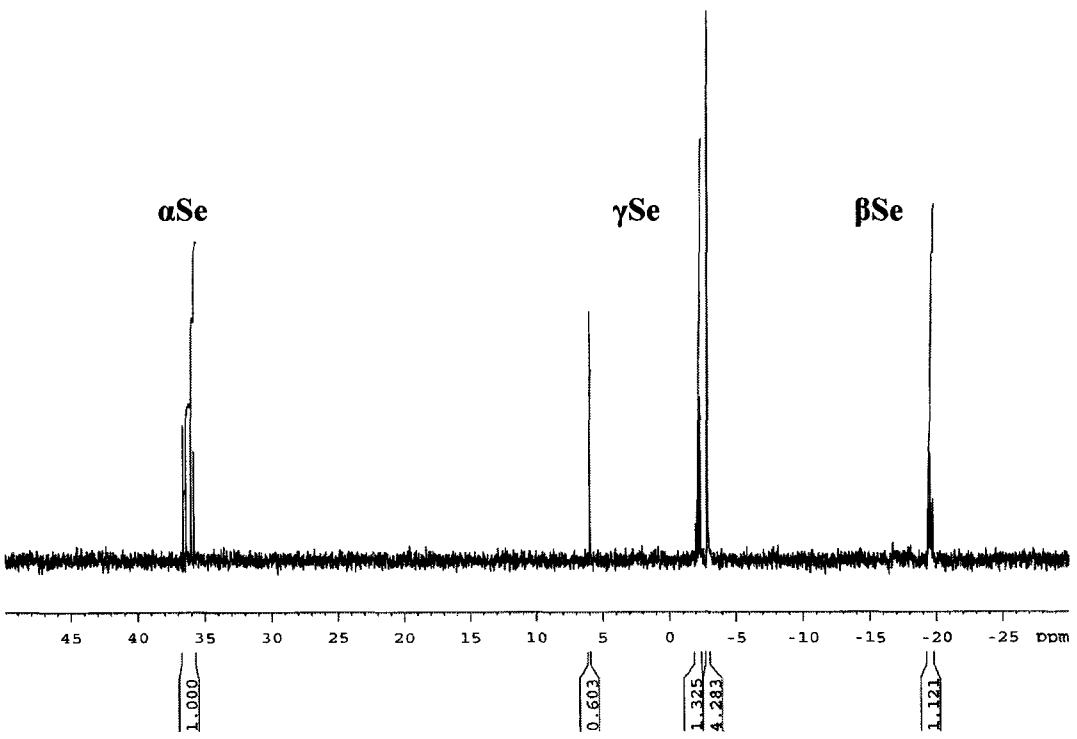


Figure 35. ^{31}P-NMR analysis of TTPαSe after Na/Ethanol precipitation.

It was observed that upon water hydrolysis, the triplet at 46 ppm split into two doublets centered at 39 ppm. In addition, these doublets shifted even further upfield upon acetyl deprotection to 36 ppm (**Figure 35**). Interestingly, the ^{1}H-undecoupled ^{31}P-NMR spectrum indicated that each set of the peaks at 36 ppm was a triplet; this result

was in support to the expected product, since the α-phosphate group containing selenium would couple with the hydrogen atoms at the 5'-position. Furthermore, the absence of the signal at –8.0 ppm, corresponding to the α-phosphate in the ordinary nucleoside triphosphate, was in further support to the incorporation of the selenium. Moreover, after the sample was treated with an iodine solution, which removed selenium oxidatively, ^{31}P-NMR analysis indicated the formation of the unmodified nucleoside triphosphate (data not included). **Table 6** shows the ^{31}P-NMR spectral data of partially purified TTPαSe.

After removal of the acetyl group with NH_3, a solution of the sodium salt of TTPαSe was precipitated in three volumes of ethanol. The residue was collected as a pellet after centrifugation. During the course of this precipitation, it was noticed that the solution and the pellet turned orange. This color was an indication that somehow α-Se-TTP was being oxidized to ordinary TTP. Further analyses concluded that ethanol was enhancing the oxidation process. Purging ethanol by bubbling argon through for 30 minute solved this issue. More important, experiments also demonstrated that 10 mM of DTT completely prevent the oxidation of TTPαSe during the Na/ethanol

precipitation. After DTT treatment, the solution turned completely clear. Some preliminary experiments indicated that oxidation could also be slowed down by the addition of a metal chelating agent, such as EDTA. ^{31}P-NMR studies showed that less than 10% oxidation occurred after the sample was incubated for more than a month at open air in the presence of 10 mM EDTA (**Figure 36).** After ethanol precipitation, the pellet was re-dissolved in a 10 mM Tris-Cl buffer containing 20 mM of DTT, and the sample was then purified by RP-HPLC.

The ^{31}P-NMR spectrum also showed the formation of some byproducts during the course of the reaction. A singlet resonating at 52 ppm was observed after the addition of BTSe. Ludwig and Eckstein also observed a similar signal during the synthesis of the corresponding thio triphosphate analogs (*133*). This signal was assigned to the formation of the symmetrical dinucleotide containing the selenium functionality **59** (**Figure 33**, and **Figure 37**). This compound is formed by the reaction of **54** and **55**. Using an excess of salicylchlorophosphite relative to **54** would reduce the formation of this byproduct. However, an excess of salicylchlorophosphite would also react with pyrophosphate to produce an inorganic cyclic

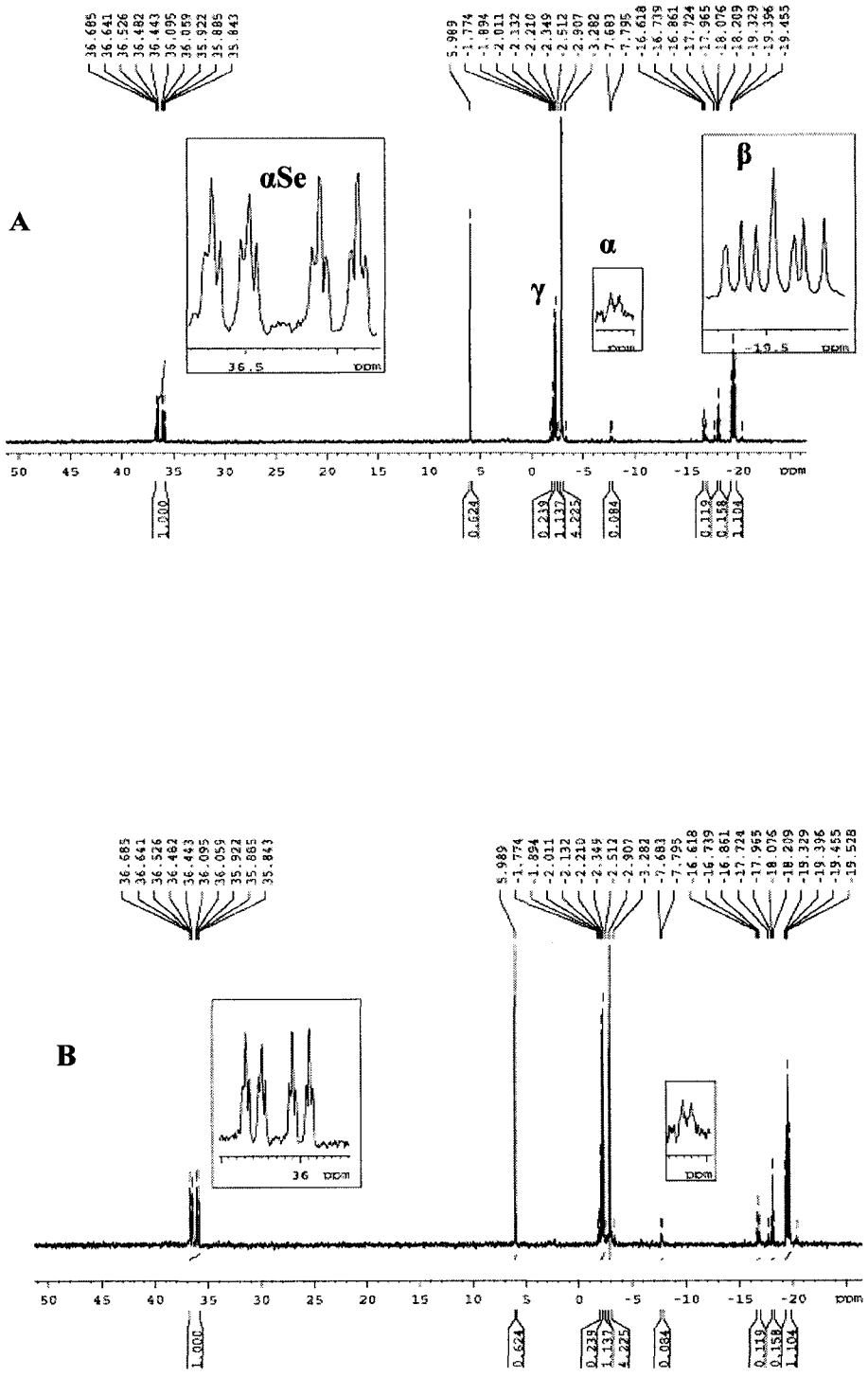


Figure 36. ^{31}P-NMR analysis of TTP. A) after 7 days open to air in NMR tube; B) after 33 days open to air.

Table 6: ^{31}P-NMR Spectral Data of TTPαSe

	Chemical shifts, ppm			Coupling constants, Hz	
	P_α	P_β	P_γ	$J_{\alpha\beta}$	$J_{\beta\gamma}$
TTPαSe I	36.58 (t)	-18.09 (d)	-2.24 (d)	32.56	20.24
TTPαSe II	35.97 (t)	-18.80 (d)	-2.36 (d)	33.53	20.89

phosphite intermediate, which upon BTSe treatment, will be oxidized by selenium (**69**). In fact, the formation of this product was observed when 1.2 molar equivalents of salicylchlorophosphite were used, as indicated by a triplet centered at 23 ppm and a doublet at –18 ppm in the ^{31}P-NMR spectrum of the reaction mixture (**Figure 33**). Importantly, adding **54** dropwise to the solution of salicylchlorophosphite over 5 minutes in a 1:1 molar ratio significantly reduced the formation of the symmetrical byproduct.

The synthetic strategy of the cyclic phosphite intermediate was also extended to the synthesis of UTPαSe and ATPαSe. The synthesis of these two ribose nucleoside analogs provided a window to investigate their compatibility with the enzymatic synthesis of RNA. The nucleoside precursors for their syntheses were obtained from commercial sources. Both the 2' and 3' positions were protected with the acetyl group.

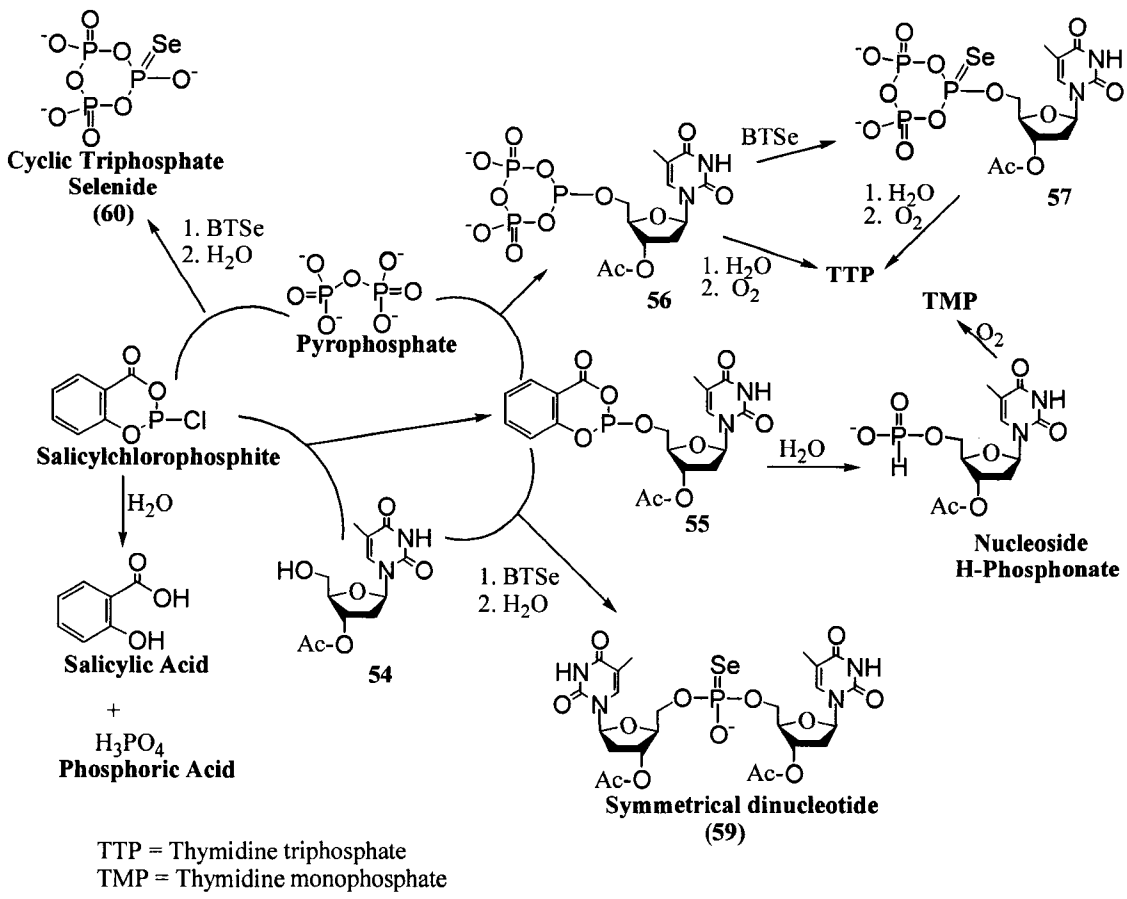


Figure 37. Formation of byproducts in the reaction mixture of TTPαSe.

5.7 HPLC Separation of NTPαSe Diastereomers

The mixture of the two diastereomers obtained after sodium ethanol precipitation was separated by reverse phase HPLC using a PE-LC-250B water pump, a LC-235 diode array UV detector operating at 260 nm, and a Zorbax C18 column (9.4 mm X 25 cm). The mobile phase consisted of 10 mM triethylammonium acetate (TEAAc) at pH 7 as buffer A; buffer B consisted of 30% acetonitrile and 10 mM TEAAc,

Ph 7. Elution was carried out using a linear gradient of buffer B from 0% to 20% in 25 min at a flow rate of 5 mL/min. The solution of the isolated isomers was concentrated by low-pressure rotary evaporation. After separation, the isomers were analyzed by RP-HPLC. The result of this analysis is shown in **Figure 38.**

The purified diastereomers were confirmed by high resolution mass spectroscopy. Two characteristic peaks were observed. One of these peaks corresponded to the unfragmented structure. The other peak, however, had the same 411.035 *m/z* among both deoxy and ribose nucleoside triphosphates. The fact that the same fragmentation pattern was observed in all the nucleoside triphosphate studied; it was proposed that the base domain, along with the 1' carbon center was cleaved off after ionization and fragmentation (**Figure 39, Figure 40 and Figure 41**).

5.7 Conclusions

The synthesis of both deoxy and ribose nucleoside triphosphates containing selenium at the α-phosphate position was investigated using nucleoside H-phosphonate and a phosphite intermediates. The selenium reagent used for the oxidation of the trivalent phosphorus

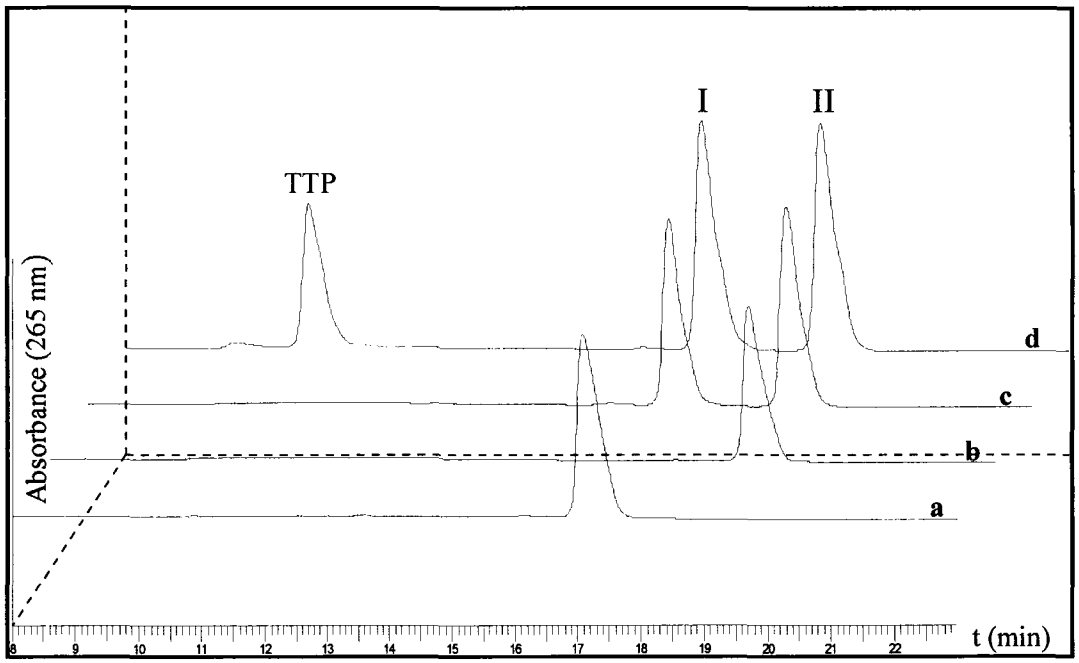


Figure 38. Assessment of the quality of the TTPαSe isolated isomers by RP-HPLC. Only a section of each spectrum is shown. The fast-moving isomer (I) and the slow- moving isomer (II) are shown in spectrum **a** and **b** respectively. Likewise, co-injections of isolated I and II alone, and together with TTP are depicted in **c** and **d**. The retention times for TTP, peak I, and peak II were 10.91, 17.18, and 19.05 min, respectively.

center in both cases was BTSe. BTSe was reactive, stable and soluble in most organic solvents.

The utilization of the H-phosphonate nucleoside intermediate produced unsatisfactory results. The reaction was not specific, the coupling efficiency was low, and undesirable byproducts were obtained. These results prompted for the study *via* nucleoside phosphite intermediate. The reaction of this intermediate had several

advantages. They included five synthetic steps in one-pot, appropriate reactivity, and selectivity toward BTSe.

The results presented here also provide new insights about the chemistry of the P—Se^- functionality. It is suggested that in the presence of oxygen, selenium gets oxidized to form a relatively stable diselenide P—Se-Se—P intermediate, which produces a typical orange color. The formation of this intermediate is promoted in the presence of ethanol, as well. Although no further experiments were performed to investigate the role of ethanol-promoted oxidation, it is possible that ethanol contains a relatively high concentration of molecular oxygen dissolved in it compared to H_2O, for example.

Further oxidation of the phosphorodiselenide intermediate gives rise to the ordinary phosphate group and selenium metal, which precipitates out in solution. This type of selenium oxidation has been observed previously (*117, 125*) and is known as "selenium washout," and probably has discouraged scientific research into the phosphoroselenoate chemistry. No mechanism has been suggested for this phenomenon, though.

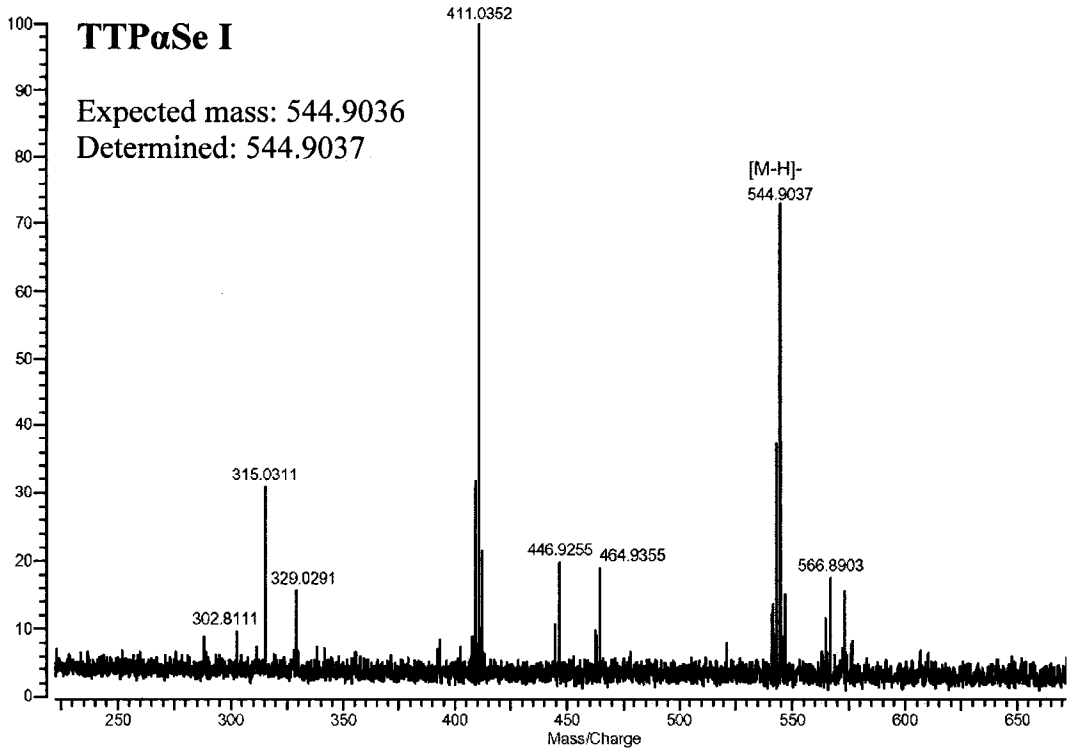


Figure 39. High resolution mass analysis of TTPαSe

Importantly, it was observed that "selenium washout" was easily prevented in the presence of DTT. The diselenide bond reacts rapidly with DTT to give rise to a clear solution of the reduced selenium product. The fact that EDTA somehow reduces the level of "selenium washout" also suggests that heavy metal ions might be involved in the oxidation process. Metal ions, such as magnesium, can coordinate

with the negative charges in the phosphate group to facilitate the elimination of selenium by water.

The results presented here only provide limited insights about the mechanism of "selenium washout." The complete elucidation of this phenomenon requires further studies.

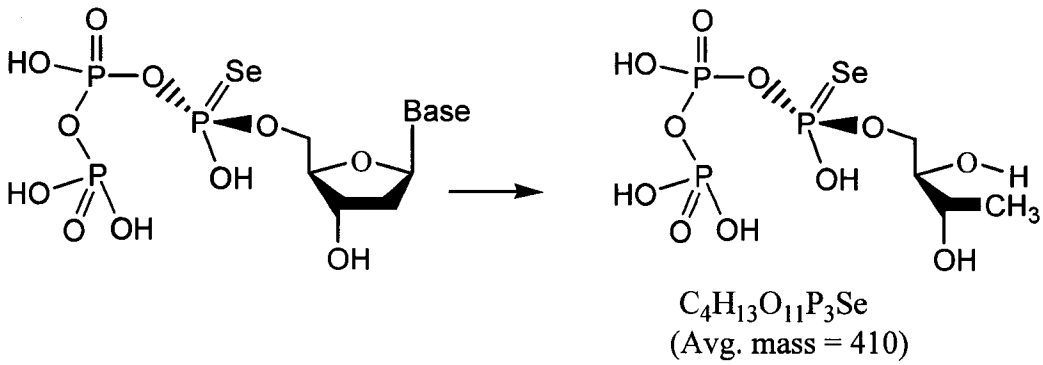


Figure 41. TTPαSe, UTPαSe, and ATPαSe produced a distinctive fragmentation pattern by ionization mass spectroscopy.

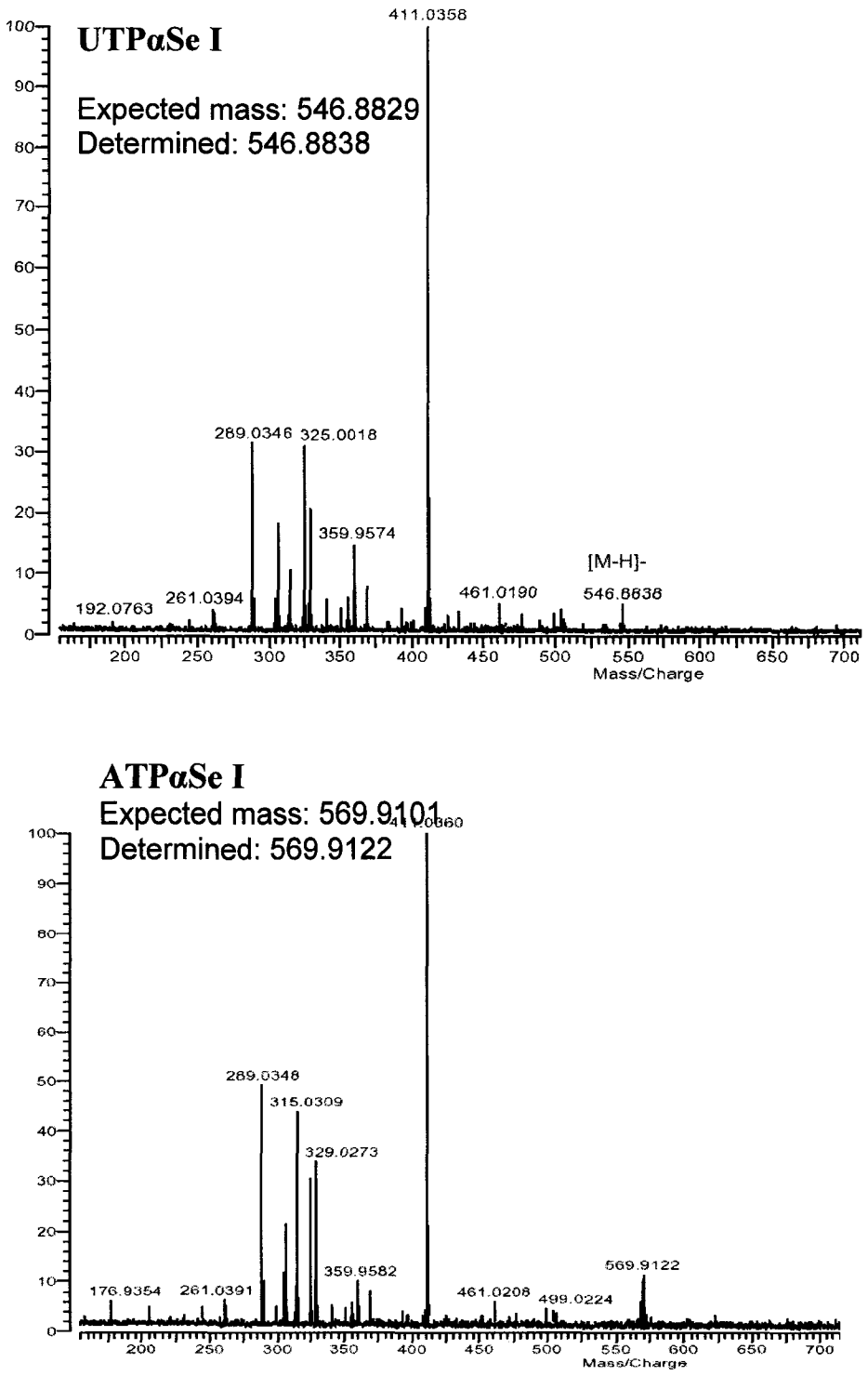


Figure 40. High resolution mass analysis of UTPαPI and ATPαSe I.

Chapter 6

ENZYMATIC SYNTHESIS OF PHOSPHOROSELENOATE DNA

6.1 Introduction

A system was developed to investigate the enzymatic synthesis of DNA containing selenium at the phosphodiester backbone, in which a non-bridging oxygen atom has been replaced by selenium. This type of DNA is called phosphoroselenoate (**Figure 42**). To date, the enzymatic synthesis of phosphorothioates and phosphoroboranoates nucleic acids have been reported (*137, 138*). Furthermore, it has been demonstrated that DNA and RNA oligonucleotides carrying such types of modifications exhibit quite different biological properties relative to their ordinary counterparts, such us resistance to enzymatic digestion and antisense activities (*137*).

6.2 Incorporation of TTPαSe into DNA

To study the compatibility of each diastereomerically pure TTPαSe in the synthesis of phosphoroselenoate DNA, the conventional method of primer extension was used. In primer extension analysis, a DNA polymerase elongates a primer DNA or RNA oligonucleotide on a DNA template sequence specifically. The primer used in the

experiment was a 20mer DNA oligonucleotide that was complementary to a region at the 3'-end of the DNA template.

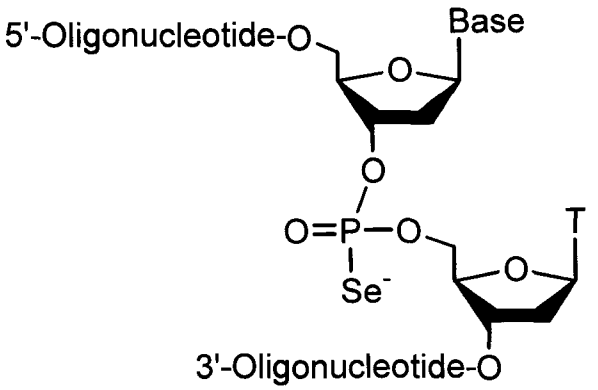


Figure 42. Chemical representation of a phosphoroselenoate DNA.

The 5'-terminus of this primer was radioactive labeled with ^{32}P through a kinase reaction. The template consisted of 55 nucleotides, of which 35 nucleotides were the overhang sequence. It is important to note that the primer is covalently linked to the growing new DNA strand during polymerization; and as a result, the full-length product will consist of 55 nucleotides. In order to assay the efficiency of selenium incorporation, the template sequence was designed to code

for three "T" nucleotides (see underlined "A". Two of these "Ts" were positioned close to each other in the extended 5' region. The other T was placed four nucleotides away from the 3' terminus. The rationale for this design had two basis: 1) if the DNA polymerase failed to recognize TTPαSe, a ladder pattern of fragments corresponding to abortive synthesis would be observed on gel electrophoresis analysis that matches the "T" sites; 2) these "T" sites would correspond to points of resistance to enzymatic digestion; as a result, they can serve to probe the presence of selenium in the DNA. The Klenow DNA polymerase was used in the experiments. The concentrations of the two selenium diastereomers and all the ordinary dNTPs were the same (0.4 mM), while the concentrations of the primer and template were 0.005 mM.

The results of a time-course experiment are shown in **Figure 43**. A few important conclusions can be drawn. First of all, the Klenow fragment of DNA polymerase I was able to recognize and incorporate both TTPαSe diastereomers into DNA. This result was unexpected since it has been known that nucleic acids polymerases are stereospecific.

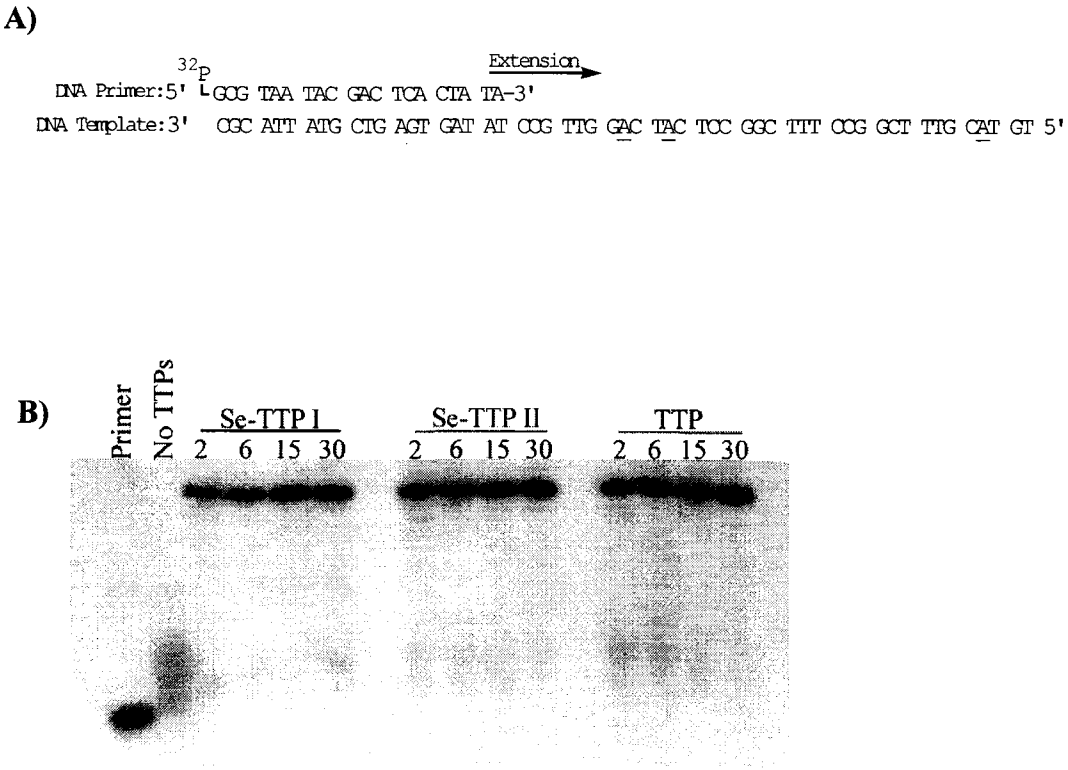


Figure 43. Synthesis of phosphoroselenoate DNA using the Klenow fragment of DNA polymerase I. A) Sequences of the primer and template used. B) Time-course primer extension using TTPαSe I and II. Gel electrophoresis was done in a 19% polyacrylamide gel.

This type of stereospecificity has been previously observed in the case thiotriphosphates (NTPαS) and boranotriphosphates (NTPαB). In the case of thiotriphosphates, for example, the Sp diastereomers was selectively recognized by T4 DNA polymerase. Thus, the polymerase recognition of both selenium diastereomers allowed the

individual synthesis of two types of diastereomerically pure DNA isomers. Determination of the absolute stereochemistry (Sp or Rp) of these TTPαSes requires further experiments. Secondly, it is observed that the Klenow fragment of DNA polymerase I recognized each of these TTPαSes diastereomers to the same extent as the ordinary TTP. This conclusion is supported by the fact that no abortive fragments were observed in any of the TTPαSe diastereomers. Complete full-length product was formed. The fact that phosphoroselenoate can be enzymatically synthesized in quantitative yield is significantly important in order to produce enough DNA for X-ray structure studies.

6.3 Resistance to Enzymatic Digestion of Phosphoroselenoate DNA

In order to exclude the possibility that the observed primer extension results of TTPαSe were not due to partial oxidation of the TTPαSe, experiments were designed to conduct exonuclease digestion to the full-length products. As mentioned earlier, phosphoroselenoate DNA should confer enhanced resistance toward enzymatic digestion relative to ordinary DNA. Based on the primer extension experimental

design, there are three points of resistance in each of the phosphoroselenoate DNA isomers: one of these points of resistance is close to the 3' terminus, while the other two are near the 5' terminus.

The enzyme used for the digestion experiments was exonuclease III. This enzyme removes mononucleotides of duplex DNA in a 3'-to-5' direction processively (*139*). It has been reported that modifications to the nucleotide backbone of DNA can affect its digestion function (*140*). A time-course experiment was performed using 5'-^{32}P labeled DNA. The results are illustrated in **Figure 44.** Several conclusions can be drawn from these data. First of all, both phosphoroselenoate isomers showed significant level of resistances relative to the ordinary DNA. These levels of resistances were present in the full-length DNA product and in the pattern reflected by the three points of resistances conferred by selenium. After 40 minutes of digestion, for example, the ordinary DNA was completely digested, while some appreciable quantity of phosphoroselenoate remained. More importantly, no points of resistances were observed in the digestion pattern of ordinary DNA. Secondly, it is particularly significant that phosphoroselenoate I and phosphoroselenoate II showed different levels of resistances. For example, the ratio of

resistance between the first point of resistance of phosphoroselenoate I after 40 minutes was much larger than the corresponding level of resistance in phosphoroselenoate II after the same amount of time. The fact that DNA made from TTPαSe I and II showed different digestion patterns was expected since these DNAs have different stereochemical centers (Sp or Rp).

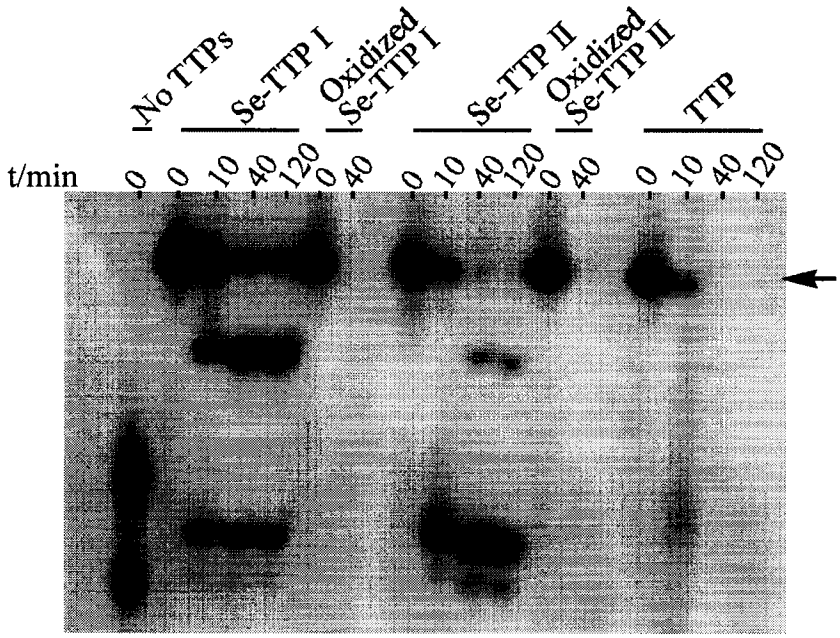


Figure 44. Time-course exonuclease digestion of the extended DNAs. Digestion was done with *E. coli* Exonuclease III containing 0.6 units/μL.

To further confirm that the resistance to exonuclease digestion was only due to the presence of the selenium functionality, the primer extension and exonuclease digestion were conducted after oxidizing both TTPαSe diastereomers to ordinary TTP by bubbling air overnight. The results demonstrated that the DNAs made from oxidized TTPαSe I and II were completely digested. This result was in complete agreement with the previous data.

6.4 Conclusions

A novel approach was developed for the synthesis of phosphoroselenoate DNA. This approach consisted of the enzymatic incorporation of both diastereomerically pure TTPαSe using primer extension and the Klenow fragment of DNA polymerase I under conventional conditions. The design of the synthesis included a 55mer DNA template that coded for the incorporation of three TTPαSe molecules.

The results showed that the Klenow fragment recognized and incorporated both TTPαSe diastereomers to similar levels as natural TTP. Full-length of the extended primers were observed on denaturing gel electrophoresis. These results were confirmed by

exonuclease digestion of the phosphoroselenoate diastereomers. Both phosphoroselenoate diastereomers displayed distinctive pattern of resistance in agreement with the selenium position, relative to the corresponding full-length DNA. In accordance with these results, when both TTPαSe diastereomers were oxidized, the resultant DNA product was completely digested by the exonuclease enzyme.

The fact that both TTPαSe diastereomers are incorporated to the same level by the Klenow fragment opens an interesting question. It has been believed for years that DNA and RNA polymerases are stereospecific enzymes, and that this property has been conserved through evolution (*141*). This specificity has been demonstrated in the case of NTPαS and NTPαB (*123*). The results presented here suggest, however, that the Klenow fragment, at least, has a broader range of substrate specificity. More careful studies comparing the substrate specificity of different nucleic acid polymerases in relation to NTPαSe, NTPαS, and NTPαB under the same experimental conditions should provide further understanding about the biochemistry of these enzymes.

As a laboratory tool, the Klenow fragment facilitates the synthesis of diastereomerically pure, long phosphoroselenoate DNA for

structural studies. The absolute stereochemistry (Rp or Sp) of these phosphoroselenoate DNAs awaits further studies. These results have been published in reference (*142*).

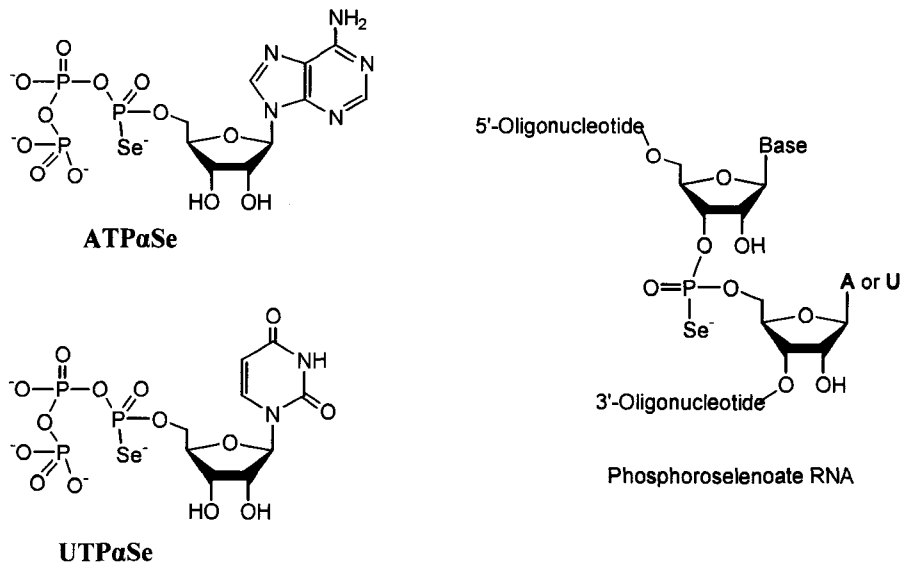
Chapter 7

ENZYMATIC SYNTHESIS OF PHOSPHOROSELENOATE RNA

7.1 Introduction

The enzymatic synthesis of phosphoroselenoate RNA was performed using the conventional method of *in vitro* transcription. The incorporation of ATPαSe and UTPαSe was investigated in connection to the catalytic properties of T7 RNA polymerase. The molar concentrations of diastereomerically pure ATPαSe and UTPαSe were kept the same as that of the ordinary NTPs, 5 mM. The template and the primer used had the same nucleotide sequence as the one used for the synthesis of phosphoroselenoate DNA, with the exception that the T7 RNA promoter is part of this sequence. This promoter region was required in order for T7 RNA polymerase to initiate transcription. The coding sequence of the DNA template consisted of 35 nucleotides and was designed to facilitate the incorporation of twelve "As" (see underlined "Ts") or three "Us" (see bolded "As") **Figure 45.** It is important to note that the primer is not covalently linked to the growing RNA oligonucleotide during polymerization. As a result, the full-length RNA product would consist of 35 nucleotides. The RNA

A)



B)

T7 RNA promoter

DNA Primer:5' GCG TAA TAC GAC TCA CTA TA-3'

DNA Template:3' CGC ATT ATG CTG AGT GAT AT CCG TTG GAC TAC TCC GGC TTT CCG GCT TTG CAT GT 5'

C)

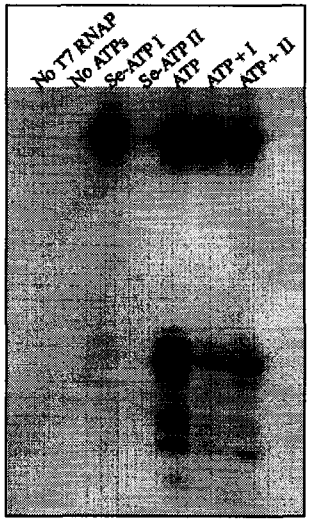


Figure 45. Synthesis of phosphoroselenoate RNA. A) Chemical structures of ATPαSe, UTPαSe, and phosphoroselenoate RNA. B) DNA template containing the promoter region of T7 RNA polymerase. Underlined "Ts" code for ATPαSe while the "As" with the asterisk code for UTPαSe. C) Polyacrylamide gel showing the in vitro transcription of RNA (12.5 % acrylamide).

transcript was body labeled by including radioactive ^{32}P-CTP in the reaction mixture.

7.2 Incorporation of ATPαSe into RNA

The results of a transcription experiment are shown in **Figure 45**. These data clearly suggested that only ATPαSe I was incorporated into the full-length RNA transcript by T7 RNA polymerase. The predominant incorporation of only one diastereomers was expected due to the specific stereochemistry of T7 RNA polymerase. This result is in complete agreement with some reports on the incorporation of thiotriphosphates and boranotriphosphate, in which this enzyme also incorporated only one diastereomer. Additionally, the enzyme was able to recognize ATPαSe I to the same frequency (or better) as the ordinary ATP; ATPαSe I produced much less short abortive fragments as compared to ATP. The data also suggested that ATPαSe II was not an inhibitor of RNA transcription. This can be explained by the fact that full-length RNA transcript was obtained when ATPαSe II was mixed in a 1:1 ratio with ordinary ATP. From this data it can be suggested that in a mixture of both ATPαSe I and II,

only ATPαSe I will get incorporated, and as a result, RP-HPLC separation of two diastereomers may not be necessary.

7.3 Resistance to Enzymatic Digestion of Phosphoroselenoate RNA

In order to confirm the enzymatic synthesis of phosphoroselenoate RNA, enzymatic digestion was conducted using the full-length RNA transcript. Data were collected from two sets of experiment done simultaneously side by side, in which two different DNA templates (55 and 87mers) were used. The RNA transcripts (35 and 67mers) were body-labeled during transcription and then digested. The enzyme used to carry out the digestion experiment was snake venom phosphodiesterase I. Phosphodiesterase I is an exonuclease that degrades DNA and RNA from the 3'-to-5' direction.

The results of time-course experiments are shown in **Figure 46 and Figure 47**. The data clearly suggested that phosphoroselenoate RNA is more resistant toward phosphodiesterase I digestion relative to ordinary DNA. In fact, the results indicated that phosphoroselenoate RNA is about 7 fold more resistant than

unmodified RNA. The results from body labeling using the 55mer and 87mer DNA templates support each other.

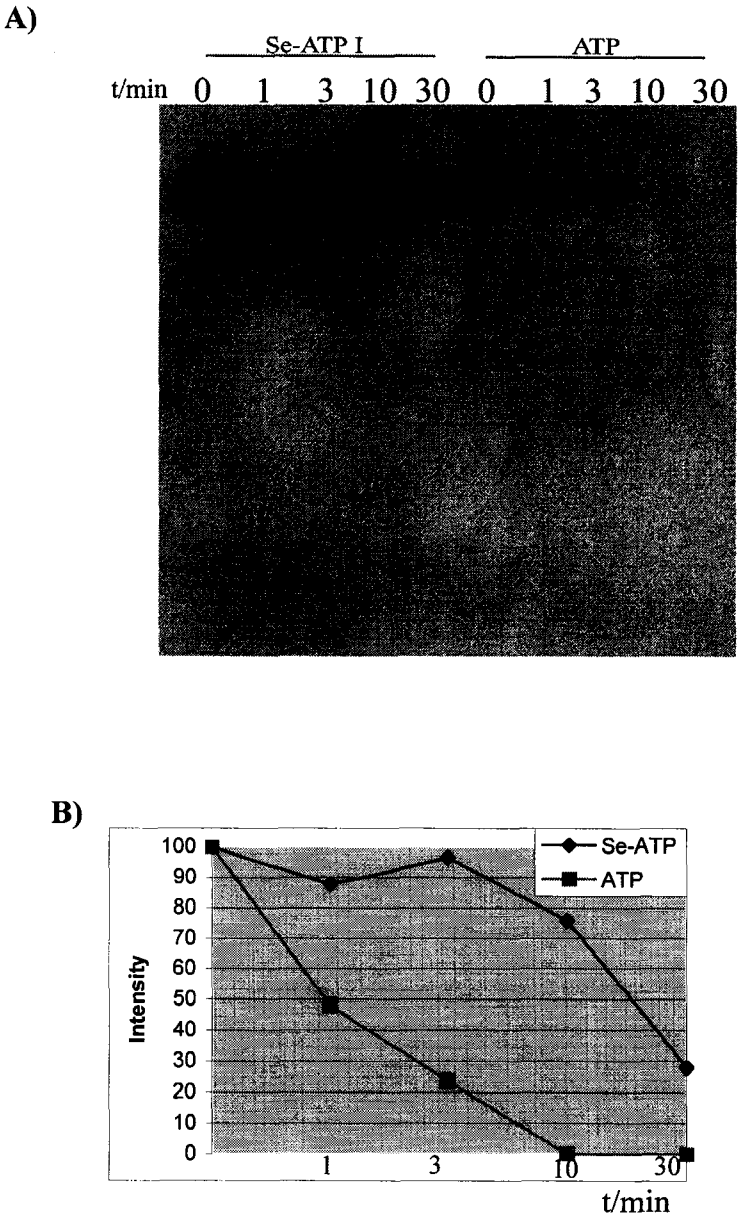


Figure 46. Time-course digestion of 35mer RNA transcripts with snake venom phosphodiesterase I. Transcripts were body-labeled during transcription with ^{32}P-CTP, and then digested. A) Acrylamide electrophoresis (12.5%) was used to analyze the transcripts. B) Quantitation and normalization of the intensities on the gel by densitometry.

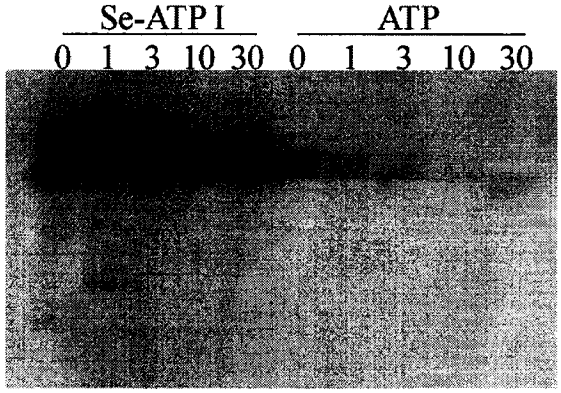


Figure 47. Time-course digestion of the 67mer RNA transcript using phosphodiesterase I. The transcript was body-labeled during transcription and then digested.

7.4 UTPαSe is an Inhibitor of T7 RNA Polymerase

The 55mer DNA template was also used to study the enzymatic compatibility of the UTPαSe diastereomers for the synthesis of phosphoroselenoate RNA. As indicated in **Figure 45b**, this template coded for the incorporation of three "Us" (see "A" with asterisk). The results showed that the UTPαSe diastereomers were not incorporated into RNA by T7 RNA polymerase (**Figure 48**). In order to investigate the effect of Mg^{2+} in the recognition of the diastereomers, its concentration was doubled. The two lanes of UTP were included to act as positive controls. As shown, although full-length RNA

products appeared in the two positive controls, no full-length transcript was synthesized from the selenium diastereomers. This result was unexpected.

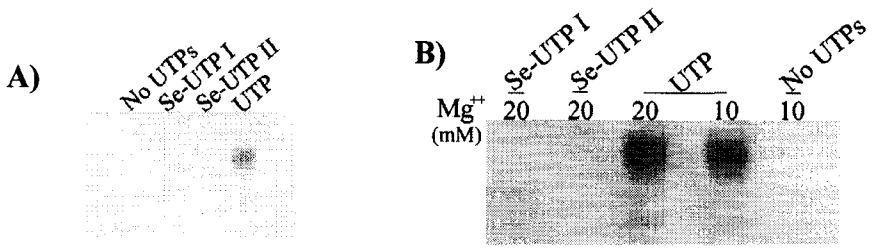


Figure 48. In vitro transcription using UTPαSe and T7 RNA polymerase. The transcript were body-labeled during transcription with ^{32}P-ATP. A) Reactions were done under 10 mM of $Mg.^{2+}$ B) Two different concentrations of Mg^{2+} were studied 10 mM and 20 mM.

The fact that neither UTPαSe diastereomer was incorporated into RNA made us hypothesize that UTPαSe inhibited T7 RNA polymerase. To test this hypothesis, both diastereomers were re-purified separately by RP-HPLC to remove any possible impurities. As another possible control experiment, both UTPαSe diastereomers were oxidized to ordinary UTP by bubbling air to an aqueous solution overnight. Since both diastereomer form the natural product upon air oxidation, they should produce the expected unmodified full-length product during in vitro transcription.

$$\text{UTP}\alpha\text{Se} \xrightarrow{O_2} \text{UTP}$$

Additionally, different ratios of UTPαSe to UTP mixed together in the reaction were included during the transcription experiments. The result of the experiment was conclusive: UTPαSe is an inhibitor of T7 RNA polymerase (**Figure 49**). Both oxidized UTPαSe I and UTPαSe II produced full-length transcripts, which match the corresponding band of ordinary UTP. It is also observed that only when UTPαSe I is mixed in a 1:9 ratio with UTP, full-length product is formed. However, no transcript was observed in the case of UTPαSe II. These results suggested that UTPαSes are an inhibitors of T7 RNA polymerase, and UTPαSe II is a stronger inhibitor than UTPαSe I.

7.5 Conclusions

The enzymatic synthesis of phosphoroselenoate RNA was explored using two diastereomerically pure ATPαSe and UTPαSe. The incorporation efficiency of these nucleoside analogs was studied by *in vitro* transcription using T7 RNA polymerase.

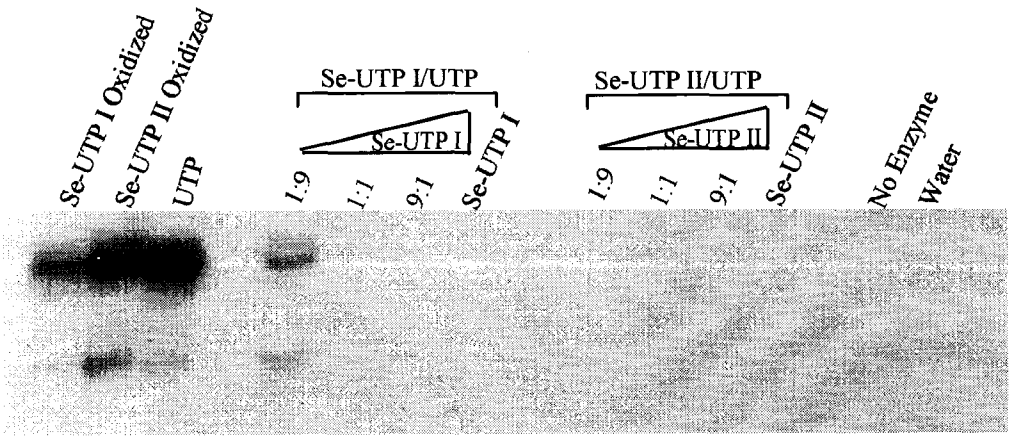


Figure 49. UTPαSe is an inhibitor of T7 RNA polymerase. The transcripts were body-labeled during transcription with ^{32}P-ATP.

In the case of ATPαSe, the results revealed that only the ATPαSe I isomer was recognized and incorporated by the enzyme. The incorporation of ATPαSe I isomer was confirmed by exonuclease digestion with snake venom phosphodiesterase I. The full-length phosphoroselenoate RNA was about 7 times more resistant toward digestion relative to the corresponding native RNA. In the case of UTPαSe, however, it was determined that T7 RNA polymerase did not incorporate any of the diastereomers into RNA. It was shown that UTPαSe is an inhibitor of T7 RNA polymerase.

Chapter 8

CONCLUDING REMARKS

A novel strategy to introduce the selenium functionality into DNA and RNA covalently has been investigated using chemical and enzymatic approaches. This strategy represents a new way to prepare heavy-atom derivatives of DNA and RNA to facilitate structural determination by X-ray crystallography *via* MAD phasing, and involves the replacement of selected oxygen atoms with selenium to minimize perturbations.

The chemical synthesis of selenium-containing DNA and RNA was investigated in connection with the phosphoramidite chemistry on solid support. Three different types of nucleoside phosphoramidite monomers harboring the selenium modification were synthesized and studied. They included 5'-methylseleno thymidine phosphoramidite, 2'-methylseleno uridine phosphoramidite, and 2'-methylseleno cytidine phosphoramidite. 5'-methylseleno thymidine phosphoramidite was only applicable for selenium labeling of DNA or RNA oligonucleotides at the 5'-termini because the methylseleno functionality prevented the continuation of further polymerization.

On the other hand, 2'-methylseleno uridine phosphoramidite, and 2'-methylseleno cytidine phosphoramidite were quite successful for the incorporation of selenium at internal positions. It was determined that these nucleoside analogs had similar coupling efficiencies to the natural nucleoside phosphoramidites under the experimental conditions. In addition, it was found that selenium was relatively resistant toward iodine oxidation during automated synthesis. Other studies aimed to analyze the effect of the methylseleno functionality on the conformation of the sugar concluded that this type of modification does not cause any major perturbation to the conformation of A-DNA or RNA duplexes. These results were further supported by thermodenaturization studies of DNA and RNA duplexes, which indicated that the selenium modification only caused minor changes in the stability of A-DNA and RNA oligonucleotides. Finally, these results set the basis for the general applicability of this chemical approach for solving crystal structures of DNA, RNA, and protein-nucleic acid complexes.

The selenium derivatization strategy was also extended to longer DNAs and RNAs that are not accessible *via* the solid phase synthesis. Although long DNAs or RNAs containing multiple selenium labels

can be prepared by ligation of a synthetic fragment harboring selenium with a long fragment prepared by DNA primer extension or RNA transcription, the efficiency of the ligation step might compromise the applicability of this method.

A more practical approach that was investigated was the enzymatic synthesis of DNAs and RNAs containing the selenium functionality at the phosphate backbone, in which selenium has replaced an oxygen atom at a non-bridging position. This strategy involved the synthesis of nucleoside triphosphates harboring selenium at the α-phosphate and their subsequent incorporation into DNAs or RNAs using primer extension or *in vitro* transcription, respectively. Three different types of selenium nucleoside triphosphate analogs were synthesized and utilized in the enzymatic approach. They included TTPαSe, ATPαSe, and UTPαSe. Their synthesis was accomplished *via* the oxidation with selenium of a nucleoside phosphite intermediate in solution.

The results involving the incorporation of TTPαSe into DNA using primer extension and the Klenow fragment of DNA polymerase I revealed, for the first time ever reported, the flexibility of this polymerase to incorporate both TTPαSe diastereomers to the same levels relative to each other and to unmodified TTP. These results

were supported by exonuclease digestion experiments, in which the selenium-containing DNA duplexes were more resistant toward digestion relative to the corresponding native DNA duplexes. The interesting question to ask is how TTPαSe affects the biochemistry of the Klenow fragment in comparison to TTPαS or TTPαB. X-ray crystallographic studies of the active Klenow fragment complex bound to the DNA template and TTPαSe might provide interesting insights into how this chemistry is accomplished.

The enzymatic incorporations of ATPαSe and UTPαSe into RNAs were also investigated in connection to the catalytic properties of T7 RNA polymerase using *in vitro* transcription. The results using ATPαSe demonstrated that only the ATPαSe I diastereomer was incorporated into RNA. This result was very consistent with the catalytic properties of T7 RNA polymerase previously reported when NTPαS and NTPαB were the substrates. The finding that UTPαSe was an inhibitor of T7 RNA polymerase was unexpected, however.

The results presented here open up new questions about the enzymology of T7 RNA polymerase, and possibly, of other nucleic acid polymerases. The fact that this enzyme is able to discriminate against UTPαSe and not ATPαSe, suggests that the base domain in

nucleoside triphosphates might play a more important role in nucleoside recognition. It would be interesting to find out if this type of base discrimination makes a clear distinction between purine and pyrimidine ribose NTPαSe. Since this type of specificity of T7 RNA polymerase has not been reported with either NTPαS or NTPαB, it is interesting to know how the selenium functionality affects the chemistry of UTPαSe.

Chapter 9

EXPERIMENTAL SECTION

9.1 Materials

Most of the solvents and reagents were purchased from Sigma, Fluka, or Aldrich and used without purification unless mentioned otherwise. Pyridine and triethylamine were refluxed with KOH for 1 hour and then distilled and stored in a desiccator. When necessary, solid reagents were dried prior to use on high vacuum and kept under dry argon. Thin layer chromatography (TLC) was run on Merck 60 F_{254} plates (0.25 mm of thickness). The TLC plates were visualized under UV light or by a Ce-Mo staining solution (phosphomolybdate, 25 g; $Ce(SO_4)_2 \cdot H_2O$, 10 g; H_2SO_4, 60 mL, conc.; H_2O, 940 mL) with heating. Flash column chromatography was performed using the Fluka silica gel 60 (mesh size 0.040-0.063 mm). ^{1}H-NMR and ^{13}C-NMR spectra were recorded on a Bruker 250 MHz or 400 MHz machine using tetramethylsilane as the internal standard in $CDCl_3$. D_2O was also used as the solvent to record NMR spectra of water-soluble compounds. Electrospray mass spectroscopy was recorded on a Hitachi-Perkin-Elmer machine at Hunter College, CUNY; while

high resolution mass spectrum analysis was performed at the Scripps Center for Mass Spectrometry in California. ^{31}P-NMR was carried out in a Bruker 400 MHz machine at Queens College, CUNY. The external reference was a solution of 85% phosphoric acid. The SR value of phosphoric acid when normalized to zero was -343.71 Hz. For the enzymatic synthesis of DNA and RNA, all the enzymes and nucleosides triphosphates used were from New England Biolabs, Inc. ^{32}P radioactive, nucleoside isotopes were purchased from Perkin Elmer.

9.2 Synthesis of 5'-Methylseleno Nucleoside Analogs

1-[(2R, 4S, 5R)-4-*tert*-butyldimethylsilyloxy-5-hydroxymethyl-tetrahydro-furan-2-yl]-thymidine (3)

The starting material **1** (211.2 mg, 0.388 mmol) was placed in a 10-mL round flask and dried on high vacuum for 1 hr. DMF (3.88 mL, dry, final conc. 0.1 M), TEA (540.0 μL, 3.88 mmol, 10 eq.), and TBDMS-Cl (247.2 mg, 1.64 mmol, 4 eq.) were then added sequentially. The reaction mixture was stirred for 4 hr under argon at room temperature and monitored periodically by silica gel TLC (5% MeOH/CH_2Cl_2, R_f = 0.46). After completion of the reaction, MeOH

(0.3 mL) was added to consume any excess reagent, and the mixture was stirred for another 15 min. MeOH was removed under reduced pressure by rotary evaporation at 30°C followed by the evaporation of DMF, which was performed with the high vacuum pump adapted to the rotavapor machine. The residue was filtrated with EtOAc and the solution was evaporated. The resultant compound **2** was then used directly for the synthesis of compound **3** without further purification.

The synthesis of compound **3** was accomplished by treatment of **2** with HOAc (3.9 mL, 80%, approximate conc. 0.1 M, 52 mmol). The reaction mixture was stirred at room temperature for 4 hr while being monitored by TLC periodically (5% MeOH/CH_2Cl_2, R_f = 0.35). Na_2CO_3 (2.74 g, 26 mmol) was then added in small portions to quench the reaction followed by addition of water (3.9 mL). The mixture was stirred until the salt dissolved and the pH was adjusted to 7-8. The resultant mixture was extracted twice with EtOAc (15 mL). The organic phases were combined and washed with $NaHCO_3$ (sat. 15 mL) and NaCl (sat. 15 mL). After drying the combined organic layer over $MgSO_4$ for 1 hr, it was filtrated and evaporated. The crude product was purified by flash column chromatography on silica gel column (20 g of silica gel). The column was washed with CH_2Cl_2 (150 mL)

followed by an increasing step-wise gradient of MeOH/CH_2Cl_2 (1% MeOH, 100 mL; 2%, 50 mL; 3%, 150 mL). The fractions containing the pure compound were combined and the solvent was removed by rotary evaporation at 30°C. The evaporated compound was then dried on high vacuum overnight to offer **3** in 95% yield.

^{1}H-NMR ($CDCl_3$) δ: 0.11 [s, 6H, $(CH_3)_2Si$], 0.81 [s, 9H, $(CH_3)_3C$], 1.62 (s, 3H, 5-CH_3), 2.11-2.42 (m, 2H, 2'-H), 2.73-2.85 (m, 1H, HO, exchangeable by D_2O), 3.66-3.78 (m, 1H, 4'-H), 3.82-3.89 (m, 2H, 5'-H), 4.38-4.49 (m, 1H, 3'-H), 6.06 (t, J = 6.7 Hz, 1H, 1'-H), 7.24-7.38 (s, 1H, 6-H), 8.85-8.98 (b, 1H, NH, exchangeable by D_2O);

^{13}C-NMR ($CDCl_3$) δ: 12.40 (5- CH_3), 17.86 (5- CH_3), 25.64 [$(\underline{CH_3})_3C$], 40.51 ($C_{2'}$), 61.76 ($C_{5'}$), 71.52 ($C_{4'}$), 86.43 ($C_{3'}$), 87.57 ($C_{1'}$), 110.83 (C_5), 136.99 (C_6), 150.47 (C_2), 164.15 (C_4).

1-[(2R, 4S, 5R)-4-*tert*-butyldimethylsilyloxy-5-bromomethyl-tetrahydro-furan-2-yl]-thymidine (4).

Compound **3** (261.1 mg, 0.733 mmol) and Ph_3P (577.5 mg, 2.2 mmol, 3 eq.) were placed in a 25-mL round-bottom flask and dried on high vacuum for 1 hr. THF (7.33 mL, final conc. 0.1 M), TEA (614 μL, 6

eq.), and CBr_4 (729.72 mg, 2.2 mmol, 3 eq.) were then added sequentially. The reaction mixture was stirred at room temperature under dry argon. The reaction was completed after 15 min as indicated by silica gel TLC (5% MeOH/CH_2Cl_2, R_f = 0.46). MeOH (0.5 mL) was then added to consume any excess reagent, and the reaction mixture was stirred for another 15 min. All the solvents were removed by rotary evaporation under reduced pressure at 30°C. The crude product was then dissolved in EtOAc, the salt was removed by filtration, and the solvent was evaporated. The residue was directly applied to a silica gel column (25 g of silica gel), and the column was eluded with EtOAc/Hexane (3:7). This solution precipitated the majority of triphenylphosphine oxide, which facilitated this purification. The fractions containing the pure product were combined and evaporated under reduced pressure, and the resultant product was dried on high vacuum overnight to give a brownish foamy product (298 mg, 97% yield).

^{1}H-NMR ($CDCl_3$) δ: 0.11 [(s, 6H, $(CH_3)_2Si$], 0.90 (s, 9H, t-Bu), 1.94 (s, 3H, 5-CH_3), 2.11-2.39 (m, 2H, 2'-H), 3.58-3.72 (m, 2H, 5'-H), 4.0-4.13 (m, 1H, 4'-H), 4.35-4.48 (m, 1H, 3'-H), 6.29 (t, J = 6.75

Hz, 1H, 1'-H), 7.45 (s, 1H, 6-H), 9.65-9.78 (b, 1H, NH, exchangeable by D_2O).

^{13}C-NMR ($CDCl_3$) δ: 12.60 (5-CH_3), 17.81 (CH_3-Si), 25.81 [($\underline{CH_3}$)$_3$C], 33.03 ($C_{5'}$), 40.52 ($C_{2'}$), 73.13 ($C_{4'}$), 84.60 ($C_{3'}$), 84.76 ($C_{1'}$), 111.20 (C_5), 135.58 (C_6), 150.35 (C_2), 163.95 (C_4).

IR (KBr): 3162, 3040, 2951, 2929, 2857, 1692, 1470, 1426, 1276, 1198, 1054, 993, 904, 838, 782, 671, 561 cm^{-1}.

UV (in acetonitrile): 263.8 nm.

1-[(2R, 4S, 5R)-4-*tert*-butyldimethylsilyloxy-5- methanesulfonyl-methyl-tetrahydro-furan-2-yl]-thymidine (5).

The starting material **3** (302 mg, 0.843 mmol) was placed in a 25-mL round flask and dried on high vacuum for 2 hr. THF (8.5 mL, final conc. 0.1M), TEA (473 μL, 3.4 mmol, 4 eq.), and methanesulfonyl chloride (131 μL, 1.31 mmol, 2 eq.) were then added sequentially. The reaction mixture was stirred at 0°C in an ice-water bath under a dry argon atmosphere while being monitored by silica gel TLC (5% MeOH/CH_2Cl_2. The reaction was completed after in 1.5 hours. MeOH (0.5 mL) was then added to consume any excess reagent, and the mixture was stirred for another 15 min. All the solvents were

removed under reduced pressure by rotary evaporation at 30° C. The residue was filtrated with EtOAc and the filtrate was evaporated. The crude product was purified by flash column chromatography on a silica gel column (18g of silica gel). The column was washed with CH_2Cl_2 (100 mL) followed by an increasing step-wise gradient of MeOH/CH_2Cl_2 (0.5%, 1%, 2% and 3%). The fractions containing the pure compound were combined and the solvent was evaporated. The resultant residue was dried overnight on high vacuum to offer **5** in 95 % yield.

^{1}H-NMR ($CDCl_3$) δ: 0.12 [(s, 6H, $(CH_3)_2Si$], 0.79 (s, 9H, t-Bu), 1.63 (s, 3H, 5-CH_3), 2.05-2.19 (m, 2H, 2'-H), 2.98 (s, 3H, CH_3SO_3), 3.94-3.96 (b, 1H, 4'-H), 4.24-4.29 (m, 2H, 5'-H), 4.31-4.39 (m, 1H, 3'-H), 6.19 (t, J = 6.6 Hz, 1H, 1'-H), 7.22 (s, 1H, 6-H), 9.36-9.468 (b, 1H, NH, exchangeable by D_2O).

^{13}C-NMR ($CDCl_3$) δ: 12.36 (5-CH_3), 17.83 (CH_3-Si), 25.60 [($\underline{CH_3}$)$_3$C], 37.59 (CH_3SO_3), 40.37 ($C_{2'}$), 68.33 ($C_{5'}$), 71.35 ($C_{4'}$), 84.17 ($C_{3'}$), 85.27 ($C_{1'}$), 111.47 (C_5), 135.47 (C_6), 150.36 (C_2), 163.86 (C_4).

N^2-*iso*Butyryl-1-[(2R, 4S, 5R)-4-*tert*-butyldimethylsilyloxy-5-methane-sulfonyl-methyl-tetrahydrofuran-2-yl]-guanine (5-Guanosine).

The same procedure for **5** was used to prepare 5-Guanosine. The mesylation reaction was complete in 1.5 hours. After the column chromatography over silica gel, it was obtained in 94% yield.

^{1}H-NMR ($CDCl_3$) δ: 0.15 [s, 6H, $(CH_3)_2Si$], 0.93 (s, 9H, *t*-Bu), 1.25 [s, 6H, $(CH_3)_2C$], 2.35-2.64 (m, 2H, 2'-H), 2.76 [sept, J = 6.9 Hz, 1H, $C\underline{H}(CH_3)_2$], 3.07 (s, 3H, CH_3-SO_3), 4.12-4.25 (m, 1H, 4'-H), 4.31-4.52 (m, 1H, 3'-H), 4.58-4.72 (m, 2H, 5'-H), 6.21 (dd, J = 6.0, 6.8Hz, 1H, 1'-H), 7.73 (s, 1H, 8-H), 8.87-8.95 (br, 1H, NH, exchangeable by D_2O).

^{13}C-NMR ($CDCl_3$) δ: 17.84 (CH_3-Si), 18.82 [$(CH(\underline{C}H_3)_2)$], 25.60 [$(\underline{C}H_3)_3C$], 27.59 (CH_3SO_3), 36.49 [$(\underline{C}H(CH_3)_2)$], 41.46 ($C_{2'}$), 66.35 ($C_{5'}$), 71.36 ($C_{4'}$), 86.81 ($C_{3'}$), 89.27 ($C_{1'}$), 122.48 (C_5), 138.47 (C_8), 146.89 (C_4) 147.36 (C_2), 155.86 (C_6).

UV (in acetonitrile): 266.8 nm.

N^6-Benzoyl-1-[(2R, 4S, 5R)-4-*tert*-butyldimethylsilyloxy-5-methanesulfonyl-methyl-tetrahydrofuran-2-yl]-adenine (5-deoxy-Adenosine).

5-deoxy-Adenosine was prepared in similar way as describe above. After column chromatography, the pure product was obtained as a colorless foamy material 98% yield.

^{1}H-NMR ($CDCl_3$). δ: 0.13 [s, 6H, $(CH_3)_2Si$], 0.92 (s, 9H, t-Bu), 2.45-2.55 (m, 2H, 3'-H), 2.98 (s, 3H, CH_3SO_3), 4.20-4.26 (m, 1H, 5'-H), 4.40-4.52 (m, 2H, 5"-H), 4.70-4.78 (m, 1H, 4'-H), 6.45-6.53 (t, J = 6.6Hz, 1H, 2'-H), 7.49-7.65 (m, 3H, Ar), 8.01-8.08 (m, 2H, Ar), 8.20 (s, 1H, 8-H), 8.80 (s, 1H, 2-H), 8.95-9.20 (b, 1H, NH, exchangeable with D_2O).

The molecular weight of **5**-deoxy-Adenosine is 547. Electrospray experiment showed molecular peaks at 548 $[M + H]^+$, 570 $[M + Na]^+$.

1-[(2R,4S,5R)-4-*tert*-butyldimethylsilyloxy-5-methselenomethyl-tetrahydro-furan-2-yl]-thymidine (6).

Method: Two-Phase System: $NaBH_4$ (63 mg, 1.65 mmol) was dissolved in 1.5 ml water and sealed in a 10-mL flask under nitrogen. Dimethyl diselenide (54 μL, 0.55 mmol) was slowly injected into the flask. Addition of 0.2 mL of ethanol with vigorous stirring helped to dissolve dimethyl diselenide completely; a colorless homogeneous solution was formed in 5 minutes. A solution of precursor **4** or **5**

(47.7mg, 0.110mmol) in toluene (1.1 mL) containing tetrahexylammonium hydrogen sulfate (1 mg) was then added to the sodium methyl selenide solution described above. The two-phase mixture (toluene and water) was stirred under nitrogen. The reaction was complete in 3 hour, as indicated by silica gel TLC (5% $MeOH/CH_2Cl_2$, $R_f = 0.40$). The organic phase was then removed, and the aqueous phase was extracted twice with EtOAc (10 mL each time). The combined organic phase was washed twice with saturated NaCl solution (10 mL each time) and dried over anhydrous $MgSO_4$. The solvents were removed by rotary evaporation under reduced pressure at 40°C. The crude product was purified on TLC (5% $MeOH/CH_2Cl_2$). The pure product (**6**) was dried on high vacuum overnight to afford 44.2 mg (93% yield).

Method 2: One-Phase System: Compound **6** was also made successfully in a one-phase system reaction of THF and ethanol. The general procedure consisted of the reduction of dimethyl diselenide with 3 mol equivalents of $NaBH_4$ in THF containing a minimum amount of ethanol (less than 25 % v/v) to enhance the solubility of $NaBH_4$. After the solution turned clear, it was added to a THF

solution of **4** or **6**. The reaction was usually completed in 3 hours as indicated by TLC analyses. All the solvents were evaporated under reduced pressure, and the compound was re-dissolved in 15 mL of EtOAc and washed three times with a saturated solution of NaCl. The aqueous phase was extracted with EtOAc (3X) and the combined organic layer was dried over anhydrous $MgSO_4$. Evaporation of the filtrate, followed by column chromatography afforded the desired produce in more 94 % yield.

Method 3: One-Phase System: Another method to synthesize compound **6** was also explored. This method involved reduction of the dialkylated diselenide nucleoside **9** with $NaBH_4$ in ethanol. After the yellow solution of the diselenide turned colorless, indicating production of the selenol **10** (approximately 5 min), 3 molar equivalents of CH_3I was added to the reaction mixture, followed by the same work-up procedures as indicated above. This approach gave **6** in quantitative yield over 95 %.

^{1}H-NMR ($CDCl_3$) δ: 0.05 [(s, 6H, $(CH_3)_2Si$], 0.88 (s, 9H, t-Bu), 1.94 (s, 3H, 5-CH_3), 2.08 (s, 3H, CH_3-Se), 2.05-2.18 and 2.28-2.40 (2m, 2H, 2'-H), 2.78-2.93 (m, 2H, 5'-H), 4.02-4.09 (m, 1H, 4'-H),

4.28-4.38 (m, 1H, 3'-H), 6.26 (t, J = 6.5 Hz, 1H, 1'-H), 7.42 (s, 1H, 6-H), 8.95-9.06 (b, 1H, NH, exchangeable in D_2O).

^{13}C-NMR ($CDCl_3$) δ: 5.66 (CH_3-Se), 12.55 (5-CH_3), 17.68 (CH_3-Si), 25.68 [($\underline{CH_3}$)$_3$C], 27.63 ($C_{5'}$), 40.63 ($C_{2'}$), 73.94 ($C_{4'}$), 84.57 ($C_{3'}$), 85.60 ($C_{1'}$), 111.11 (C_5), 135.59 (C_6), 150.18 (C_2), 163.69 (C_4).

The MS spectrum of **6** was determined by electrospray, positive ion experiment). The molecular weight ($C_{17}H_{30}N_2O_4SiSe$) is 434 with adjustment for ^{80}Se isotope. Because of binding of H or Na ions, two sets of molecular peaks are observed, in which all selenium isotopic peaks [Se 76 (9%), 77 (7%), 78 (23%), 80 (49%), 82 (9.2%)] are also observed: 431 [M (^{76}Se) + H]$^+$, 432 [M (^{77}Se) + H]$^+$, 433 [M (^{78}Se) + H]$^+$, 435 [M (^{80}Se) + H]$^+$, 437 [M (^{82}Se) + H]$^+$, 453 [M (^{76}Se) + Na]$^+$, 454 [M (^{77}Se) + Na]$^+$, 455 [M (^{78}Se) + Na]$^+$, 457 [M (^{80}Se) + Na]$^+$, and 459 [M (^{82}Se) + Na]$^+$.

IR (KBr): 3157, 3034, 2960, 2940, 2862, 2363, 1698, 1470, 1426, 1370, 1276, 1198, 1120, 1089, 1049, 827, 777, 682, 621 cm^{-1}.

UV (in acetonitrile): 265.2 nm.

1-[(2R, 4S, 5R)-4-hydroxy-5-methselenomethyl-tetrahydrofuran-2-yl]-thymidine (7)

Following standard procedures, compound **6** (50 mg, 0.115 mmol) was dissolved in THF (345 μL), and *tert*-butyl ammonium fluoride (1 M in THF 230 μL, 2 eq.) was added. The deprotection reaction was complete in 2 hours (monitored by TLC, 7.5% $MeOH/CH_2Cl_2$). The product was purified on TLC to give quantitative yield, and the structure of this product was confirmed by spectroscopic analysis, including ^{77}Se-NMR.

^{1}H-NMR ($CD_3OD/CDCl_3$=1:1) δ: 1.93 (s, 3H, 5-CH_3), 2.06 (s, 3H, CH_3-Se), 2.15-2.28 and 2.32-2.46 (2m, 2H, 2'-H), 2.82-2.96 (m, 2H, 5'-H), 4.05-4.15 (m, 1H, 4'-H), 4.28-4.38 (m, 1H, 3'-H), 6.26 (t, J = 6.7 Hz, 1H, 1'-H), 7.49 (s, 1H, 6-H), 8.95-9.06 (b, 1H, NH, exchangeable in D_2O).

^{13}C-NMR ($CDCl_3$) δ: 5.31 (CH_3-Se), 12.11 (5-CH_3), 27.49 ($C_{5'}$), 39.74 ($C_{2'}$), 72.88 ($C_{4'}$), 84.21 ($C_{3'}$), 84.14 ($C_{1'}$), 110.89 (C_5), 135.70 (C_6), 150.44 (C_2), 163.69 (C_4).

As ^{77}Se NMR active (M_I=½). ^{77}Se-NMR ($CDCl_3$) δ: 362.55 ppm (reference: dibenzyl diselenide, 133.25 ppm)

IR (KBr): 3473, 3167, 3095, 2960, 2923, 2812, 2679, 1703, 1476, 1410, 1259, 1071, 1015, 950, 888, 816, 632, 570 cm^{-1}.

UV (in acetonitrile): 265.0 nm.

N^6-Benzoyl-1-[(2R,4S,5R)-4-*tert*-butyldimethylsilyloxy-5-methselenomethyl-tetrahydrofuran-2-yl]-adenine (6-deoxy-Adenosine)

$NaBH_4$ (20.0 mg, 0.525 mmol) was placed in 10mL-round flask under nitrogen. Water (1.5 mL) and dimethyl diselenide (17.0 μL, 0.175 mmol) were sequentially and slowly injected into the flask. Vigorous stirring helped to dissolve dimethyl diselenide completely, forming a colorless homogeneous solution after 5 –10 minutes; the pH of the solution was higher than 11. Since high pH caused the hydrolysis of the protecting benzoyl group on the adenine base, and pH 7.0 or lower made the following selenide substitution very slow, the pH was adjusted to 8.0 by adding dilute HCl dropwise. **4** or 5-deoxy-Adenosine (19.2mg, 0.0350mmol) and tetrahexylammonium hydrogen sulfate (0.5 mg) dissolved in toluene (0.7 mL) were then added to the sodium methylselenide solution described above, and the two-phase mixture (toluene and water) was stirred under nitrogen. The reaction was complete after 5 hours, forming 6-deoxy-Adenosine (silica gel TLC, 5% $MeOH/CH_2Cl_2$, $R_f = 0.52$). Longer reaction time caused slow hydrolysis of the benzoyl group (the hydrolyzed product $R_f = 0.33$ on TLC, 5% $MeOH/CH_2Cl_2$). The organic phase was

removed and the aqueous phase was extracted twice with EtOAc; the combined organic phase was washed twice with saturated NaCl solution, and dried over $MgSO_4$. The solvents were removed by rotary evaporation under reduced pressure at 40°C. The crude product was dissolved in CH_2Cl_2 and loaded on a silica gel TLC plate (5% $MeOH/CH_2Cl_2$). Colorless product was recovered from this purification (18.2 mg, 95% yield).

^{1}H-NMR ($CDCl_3$). δ: 0.12 [s, 6H, $(CH_3)_2Si$], 0.92 (s, 9H, t-Bu), 2.02 (s, 3H, CH_3Se), 2.43-2.53 (m, 1H, 2'-H), 2.78-3.00 (m, 3H, 2'-H and 5'-H), 4.15-4.23 (m, 1H, 4'-H), 4.55-4.63 (m, 1H, 3'-H), 6.42-6.50 (t, J = 6.6Hz, 1H, 1'-H), 7.48-7.57 (m, 2H, Ar), 7.57-7.65 (m, 1H, Ar), 8.00-8.07 (m, 2H, Ar), 8.27 (s, 1H, 8-H), 8.80 (s, 1H, 2-H), 8.94-9.00 (b, 1H, NH, exchangeable by D_2O).

The molecular weight of 5-deoxy-Adenosine ($C_{24}H_{33}N_5O_2SiSe$) is 547 with adjustment for ^{80}Se isotope [average atomic weight of Se is 79, including 76 (9%), 77 (7%), 78 (23%), 80 (49%), 82 (9.2%)]. The molecular peaks are: 546 $[M(^{78}Se) + H]^+$, 548 $[M(^{80}Se) + H]^+$, 550 $[M(^{82}Se) + H]^+$, 568 $[M(^{78}Se) + Na]^+$, 570 $[M(^{80}Se) + Na]^+$, 572 $[M(^{82}Se) + Na]^+$.

1-[(2R, 4S, 5S)-4-*tert*-butyldimethylsilyloxy-5-diselenomethyl-tetrahydro-furan-2-yl]-thymidine (9).

Sodium borohydride (45.1 mg, 1.2 mmol) dissolved in water (1.2 mL) was added to a 25-mL flask containing a suspension of selenium metal powder (94.3 mg, 1.19 mmol) in water (1.2 mL). The reaction was placed in an ice-bath for the first few minutes to slow down the reaction; the reaction mixture was later stirred at room temperature under argon. After the vigorous reaction had subsided (approximately 10 min), additional selenium metal powder (94.3 mg, 1.19 mmol) was added. The mixture was stirred for another 10 min and then warmed on a steam bath for 5 min to completely dissolve all the selenium and to decompose the excess $NaBH_4$. The color of the solution was brownish red and its pH was about 10-11. This aqueous solution was injected into a 25-mL flask containing **4** or **5** (100 mg, 0.239 mmol), tetrahexylammonium hydrogen sulfate (10.7 mg, 0.0239 mmol, 0.1 eq.), and toluene (4.8 mL). The reaction was closely monitored by silica gel TLC (5% MeOH/CH_2Cl_2, product $R_f = 0.33$). After 45 min the reaction was complete. A current of air was then passed through the reaction mixture to oxidize any excess of sodium diselenide to selenium metal, which precipitated.

The crude product mixture was centrifuged to remove the precipitated selenium metal, followed by extraction twice with toluene (15 mL), and the respective organic layers were combined. This organic phase was then washed with $NaHCO_3$ (15 mL, sat.) and NaCl (15 mL, sat.). The resultant yellowish organic phase was dried over anhydrous $MgSO_4$ for 30 min, the solution was filtered, and the solvents were removed by rotary evaporation under reduced pressure at 40 °C. The crude residue was then dissolved in CH_2Cl_2, and the solution was loaded onto a silica gel column (6.0 g silica gel). The column was eluted with CH_2Cl_2 (50 mL) followed by a stepwise gradient of MeOH (0-3%). The fractions containing the product were combined and the solvents were removed by rotary evaporation at 30°C. After drying under high vacuum overnight, 96 mg of the yellowish diselenide product (**9**) was obtained (96% yield).

^{1}H-NMR ($CDCl_3$) δ: 0.11 [s, 6H, $(CH_3)_2Si$], 0.93 (s, 9H, *t*-Bu), 1.94 (s, 3H, 5-CH_3), 2.25-2.4 (m, 2H, 2'-H), 3.33-3.36 (m, 2H, 5'-H), 4.03-4.12 (m, 1H, 4'-H), 4.28-4.35 (m, 1H, 3'-H) 6.19 (t, J = 6.6 Hz, 1H, 1'-H) 7.25 (s, 1H, 6-H). 8.80-8.92 (b, 1H, NH, exchangeable by D_2O).

^{13}C-NMR ($CDCl_3$) δ: 12.66 (CH_3-Si), 17.92 [($\underline{CH_3}$)$_3$C], 25.71 (5-CH_3), 40.41 ($C_{2'}$), 52.79 ($C_{5'}$), 74.14 ($C_{4'}$), 85.38 ($C_{3'}$), 86.04 ($C_{1'}$), 111.18 (C_5), 135.72 (C_6), 150.16 (C_2), 164.10 (C_4).

IR (KBr): 3428, 3179, 3057, 2962, 2934, 2862, 2363, 1703, 1476, 1370, 1281, 1204, 1104, 838, 782, 666 cm^{-1}.

UV (in acetonitrile): 264.4 nm.

Electrospray mass analysis (positive mode): Calculated mass ($C_{32}H_{54}N_4O_8Se_2Si_2$) 839.2 [M(^{80}Se) + H]$^+$; measured: 837.1 [M(^{78}Se) + H]$^+$; 839.1 [M(^{80}Se) + H]$^+$; 841.1 [M(^{82}Se) + H]$^+$.

9.3 Synthesis of the 2'-Selenium Derivatized Uridine and Cytidine Phosphoramidites

A) Using Partially Protected Nucleoside Precursors:

2'-mesyl-3'-*tert*-butyldimethylsilyloxy-5'-dimethoxytrityl-uridine (13)

Partially protected uridine 2 (0.429 g, 0.65 mmol) was placed in a 25-mL round flask and dissolved in dry THF (6.5 mL, 0.1 M). Under an ice-water bath, triethylamine (269 mL, 1.95 mmol) and methanesulfonyl chloride (76 mL, 0.975 mmol) were then added. The reaction mixture was stirred under argon at 0 °C for 20 min [monitored on TLC, CH_3OH/CH_2Cl_2 (1:19), starting material R_f =

0.37, product R_f = 0.39]. After the reaction was complete, MeOH (0.5 mL) was added to quench the reaction; it was stirred for another 15 min. The solvents were removed under reduced pressure. The crude product was purified by flash chromatography on a silica gel column (CH_3OH/CH_2Cl_2; the gradients, 0.5% to 2%) to give 3 (0.455 g, 95% yield) as a white foam.

^{1}H-NMR (CD_3OD) δ (ppm): 0.08 and 0.21 [s, s, 2x3H, $(CH_3)_2Si$], 0.92 [m, 9H, $(CH_3)_3CSi$], 3.37 (s, 3H, CH_3SO_2), 3.50-3.54 (m, 2H, H-5'), 3.88 (s, 6H, 2xCH_3O), 4.20-4.25 (m, 1H, H-4'), 4.63-4.67 (m, 1H, H-3'), 5.35-5.42 (m, 1H, H-2'), 5.41 (d, J=8.1 Hz, 1H, H-5), 6.19 (d, J=2.7 Hz, 1H, H-1'), 6.98-6.92 (m, 4H, Ar-H), 7.37-7.56 (m, 9H, Ar-H), 8.21 (d, J=8.2 Hz, 1H, H-6).

^{13}C-NMR (CD_3OD) δ (ppm): 19.23 (Si-CH_3), 26.72 [$SiC(CH_3)_3$], 39.57 (-SO_3CH_3), 56.28 (OCH_3), 62.01 (C-5'), 71.14 (C-3'), 82.52 (C-4'), 85.11 (C-2'), 89.06 (C-1'), 103.53 (C-5), 114.80, 132.04, 136.70, 145.95, 160.94 (Ar-C), 142.40 (C-6), 152.52 (C-2), 166.31 (C-4).

IR (KBr): 3450 (br.), 3068, 3030, 2950, 2839, 1702, 1610, 1519, 1460, 1380, 1256, 1188, 1110, 1054, 1010, 933, 904, 858, 782, 763, 715, 561 cm^{-1}.

UV (in acetonitrile), λ_{max}: 236.2, 268.6 nm.

FAB-HRMS: $C_{37}H_{47}N_2O_{10}SiS$ ($M + H^+$), 739.2719 (calc. 739.2721).

2,2'-anhydro-1-(2'-deoxy-3'-*tert*-butyldimethylsilyloxy-5'-O-dimethoxytrityl-β-Dribofuranosyl)-uracil (15)

13 (407 mg, 0.551 mmol) and tetrahexylammonium hydrogen sulfate (25 mg, 0.1 eq.) were dissolved in toluene (11.0 mL, 0.05M). A saturated Na_2CO_3 solution (11 mL, pH~11) was added to the previous solution. The suspension was vigorously stirred at room temperature for 3 hrs while being periodically monitored by TLC [CH_3OH/CH_2Cl_2 (1:19), product R_f = 0.32]. After the reaction was complete, the suspension was extracted three times with ethyl acetate (3x15 mL). The resultant organic phases were combined, washed with NaCl (15 mL, sat.), and dried over anhydrous $MgSO_4$. After filtration, the solvent were evaporated under reduced pressure, and the crude product was purified by flash chromatography on a silica gel column (CH_3OH/CH_2Cl_2; the gradients, 0.5% to 3%) to afford 4 (339 mg, 96% yield) as a white foam.

^{1}H-NMR ($CDCl_3$) δ (ppm): 0.08 and 0.14 [s, s, 2x3H, $(CH_3)_2Si$], 0.88 [m, 9H, $(CH_3)_3CSi$], 2.98-3.14 (m, 2H, H-5'), 3.82 (s, 6H, 2xCH_3O), 4.22-4.29 (m, 1H, H-4'), 4.48-4.52 (m, 1H, H-3'), 5.03-

5.09 (m, 1H, H-2'), 5.98 (d, J=7.6 Hz, 1H, H-5), 6.10 (d, J=5.7 Hz, 1H, H-1'), 6.72-6.84 (m, 4H, Ar-H), 7.18-7.36 (m, 10H, H-6, 9 Ar-H).

^{13}C-NMR ($CDCl_3$) δ (ppm): 18.22 (Si-CH_3), 25.94 [$SiC(CH_3)_3$], 55.61 (OCH_3), 62.51 (C-5'), 77.17 (C-3'), 86.95 (Ar.-C), 87.90 (C-4'), 89.22 (C-2'), 90.04 (C-1'), 110.87 (C-5), 113.62, 127.38, 128.22, 130.13, 134.54, 135.56, 144.50, 158.98 (Ar-C), 137.40 (C-6), 159.51 (C-2), 171.68 (C-4).

IR (KBr): 3450, 3035, 2930, 2860, 1670, 1530, 1505, 1460, 1250, 1190, 1085, 1060, 820, 790, 770, 710, 610 cm^{-1}.

UV (in acetonitrile), λ_{max}: 233.2, 281.0 nm.

FAB-HRMS: $C_{36}H_{43}N_2O_7Si$ ($M + H^+$), 643.2838 (calc. 643.2839).

2,2'-anhydro-1-(2'-deoxy-5'-O-dimethoxytrityl-β-D-ribofuranosyl)-uracil (16)

15 (140.1 mg, 0.218 mmol) was placed in a 10-mL round flask and dissolved in THF (1.6 mL). A 1M solution of *tert*-butylammonium fluoride in THF (0.43 mL, 0.43 mmol) was then injected. The reaction mixture was stirred at room temperature for 1 hr, monitored by TLC [CH_3OH/CH_2Cl_2 (1.5:18.5), product R_f = 0.15)]. The solvent was evaporated under reduced pressure and the crude product was purified by flash chromatography on a silica gel column (CH_3OH/CH_2Cl_2; the

gradients, 0.5% to 3%) to give 5 (109 mg, 95% yield) as a white, foamy product.

^{1}H-NMR ($CDCl_3$) δ (ppm): 2.99-3.18 (m, 2H, H-5'), 3.69 (s, 6H, 2xCH_3O), 4.29-4.36 (m, 1H, H-4'), 4.42-4.46 (m, 1H, H-3'), 5.20-5.25 (m, 1H, H-2'), 5.92 (d, J=7.5 Hz,1H, H-5), 6.07 (d, J=5.7 Hz, 1H, H-1'), 6.98-6.79 (m, 4H, Ar-H), 7.11-7.30 (m, 10H, H-6, 9 Ar-H).

^{13}C-NMR ($CDCl_3$) δ (ppm): 55.20 (OCH_3), 62.96 (C-5'), 75.60 (C-3'), 86.16(Ar.-C), 87.64 (C-4'), 89.42 (C-2'), 90.31 (C-1'), 109.64 (C-5), 113.18, 126.91, 127.91,129.83, 135.40, 144.39, 158.45 (Ar-C), 135.93 (C-6), 159.75 (C-2), 172.69 (C-4).

IR (KBr): 3400 (br.), 3030, 2920, 2850, 1670, 1520, 1510, 1490, 1460, 1260, 1190, 1095, 1055, 820, 770, 705, 580 cm^{-1}.

UV (in acetonitrile), λ_{max}: 233.8, 281.2 nm.

FAB-HRMS: $C_{30}H_{29}N_2O_7$ ($M + H^+$), 529.1976 (calc. 529.1974).

B) Using Unprotected Nucleoside Precursors:

2,2'-O-anhydro-1-(β-D-arabinofuranosyl)-uracil (18)

Uridine (50 g, 201 mmol) and diphenyl carbonate (48 g, 224 mmol) were placed in a round bottle flask, and N,N-dimethylformamide (50 mL) was added. The slurry was heated in an oil bath at 100°C. Dry

sodium bicarbonate (400 mg) was then added, and a watch glass was used to cover the flask. The reaction mixture was heated at 130-140°C for one hour while being monitored on TLC (methanol/methylene chloride, 2:8). After completion, the reaction was cooled to room temperature, filtered, and washed by methanol three times (each time 30 mL). The white powdered product was dried on high vacuum overnight. (37.9 g, 82%; wt. 226). ^{1}H-NMR (d_6-DMSO) δ (ppm): 3.09-3.21 (m, 2H, H-5'), 3.95-4.11 (m, 1H, H-4'), 4.25-4.41 (m, 1H, H-3'), 4.91 (m, HO), 5`. 10 (d, J=6.5 Hz, 1H, H-2'), 5.81 (m, HO), 5.96 (d, J=7.5 Hz, 1H, H-5), 6.22 (d, J=5.7 Hz, 1H, H-1'), 7.78 (d, J=7.5 Hz, 1H, H-6).

2,2'-anhydro-1-[2'-deoxy-5'-O-(4,4-dimethoxytrity1)-β-D-ribofuranosyl]-uracil ***(16)***

Dimethoxytrityl chloride (18.6 g, 55.0 mmol) and 2,2'-O-anhydrouridine (12.4 g, 54.9 mmol) were placed in a 250-mL round flask and dried on high vacuum for half an hour before addition of dry pyridine (110 mL, 0.5 M). The reaction was stirred for 3 hr at room temperature and monitored by TLC (methanol/methylene chloride = 7.5%). After the reaction was complete, MeOH (10 mL) was added to quench the reaction, and the mixture was stirred for 10 minutes. NaCl

(50 mL, sat.) and ethyl acetate (EtOAc, 100 mL) were then added to the flask. The aqueous phase was extracted with EtOAc (3x 100 mL). The combined organic phase was dried over anhydrous $MgSO_4$ for 30 min, the salt was filtrated, and the organic solvents were evaporated under reduced pressure. The residue was purified on a silica gel column (equilibrated with EtOAc) and eluted with a methanol/EtOAc gradient [EtOAc, and CH_3OH in EtOAc (0-10%)]. The fractions containing the product were combined, evaporated, and dried on high vacuum overnight to yield a white foamy product. (78% yield; 22.7 g, wt. 528). UV (in acetonitrile), λ_{max}: 233.8, 281.2 nm. IR (KBr): 3400 (br.), 3030, 2920, 2850, 1670, 1520, 1510, 1490, 1460, 1260, 1190, 1095, 1055, 820, 770, 705, 580 cm^{-1}. ^{1}H-NMR ($CDCl_3$) δ (ppm): 2.99-3.18 (m, 2H, H-5'), 3.69 (s, 6H, 2xCH_3O), 4.29-4.36 (m, 1H, H-4'), 4.42-4.46 (m, 1H, H-3'), 5.20-5.25 (m, 1H, H-2'), 5.92 (d, J=7.5 Hz, 1H, H-5), 6.07 (d, J=5.7 Hz, 1H, H-1'), 6.98-6.79 (m, 4H, Ar-H), 7.11-7.30 (m, 10H, H-6, 9 Ar-H). ^{13}C-NMR ($CDCl_3$) δ (ppm): 55.20 (OCH_3), 62.96 (C-5'), 75.60 (C-3'), 86.16 (Ar.-C), 87.64 (C-4'), 89.42 (C-2'), 90.31 (C-1'), 109.64 (C-5), 113.18, 126.91, 127.91, 129.83, 135.40, 144.39, 158.45 (Ar-C), 135.93 (C-6), 159.75 (C-2),

172.69 (C-4). FAB-HRMS: $C_{30}H_{29}N_2O_7$ ($M + H^+$), 529.1976 (calc. 529.1974).

5'-O-(4,4-dimethoxytrity1)-2'-methylseleno-2'-deoxy-uridine (19)

$NaBH_4$ (1.1 g, 6 mmol) was placed in a 250-mL round flask and suspended in dry THF (45 mL) under vigorous stirring. Dimethyl diselenide ($CH_3SeSeCH_3$, 1.99 mL, 19.9 mmol) was slowly injected, and the suspension was placed in an ice-water bath under dry argon. Anhydrous ethanol (5 mL) was added dropwise. Gas bubbles started to occur in the yellow mixture. The reaction mixture turned colorless after 60 min, and a solution of **16** (10.5 g, 19.9 mmol) in THF (20 mL) was injected to the flask. The flask for **16** was rinsed twice (2x5mL) and the washings were injected to the reaction flask. The ice-water bath was removed, and the reaction was stirred under argon and monitored by TLC (5% CH_3OH/CH_2Cl_2, product R_f = 0.35). The reaction was complete in 1 hr. NaCl (sat, 50 mL) was added to the reaction, followed by dropwise addition of 20% HOAc (approximate 12 mL) until pH reached 7. The crude product was extracted by EtOAc (3x 100 mL), and the combined organic layer was dried over anhydrous $MgSO_4$, followed by filtration and solvent evaporation.

The residue was purified on a silica gel column (equilibrated with hexane/CH_2Cl_2, 1:1) and eluted with a methanol/methylene chloride gradient [(hexane/CH_2Cl_2, 1:1, hexane/CH_2Cl_2, 1:3, CH_2Cl_2, and CH_3OH in CH_2Cl_2 (0.5, and 1%)] to afford the desired product **19** as a light yellow foam (92% yield; 11.4 g, wt. 622).

^{1}H-NMR ($CDCl_3$) δ: 2.08 (s, 3H, CH_3Se), 3.42-3.47 (m, 2H, H-5'), 3.49-3.54 (m, 1H, H-2'), 3.78 (s, 6H, CH_3O), 4.14-4.18 (m, 1H, H-4'), 4.34-4.39 (m, 1H, H-3'), 5.36 (d, J=8.0 Hz, 1H, H-5), 6.19 (d, 1H, J=3.3 Hz, H-1'), 6.86-6.92 (m, 4H, aromatic), 7.19-7.38 (m, 9H, aromatic), 7.77 (d, J=7.8 Hz, 1H, H-6), 8.48 (br, 1H, NH).

^{13}C-NMR ($CDCl_3$) δ: 5.18 ($SeCH_3$), 51.12 (C-2'), 55.63 (OCH_3), 62.8 (C-5'), 71.90 (C-3'), 84.84 (C-4'), 87.66 (Ar.-C), 87.98 (C-1'), 103.05 (C-5), 113.69, 128.43, 130.42, 135.50 (Ar-C), 139.45 (C-6), 144.36 (Ar-C), 150.55 (C-2), 158.55 (Ar-C), 163.03 (C-4).

UV (in acetonitrile), λ_{max}: 236.2, 273.4 nm. IR (KBr): 3450 (br.), 3080, 3030, 2940, 1705, 1610, 1520, 1460, 1390, 1245, 1190, 1090, 1045, 850, 782, 710, 590 cm^{-1}.

ESI-MS (positive mode): $[M(^{80}Se)+NH_3]^+$ calculated 641, observed 641.2; $[M(^{78}Se)+NH_3]^+$ calculated 639, observed 639.1 FAB-HRMS: $C_{31}H_{32}N_2O_7Se$ (M^+), 624.1376 (calc. 624.1374).

3'-O-(2-cyanoethyl-N,N-diisopropylphosphoramidite)-5'-O-(4,4-dimethoxytrityl)-2'-methylseleno-2'-deoxy-uridine (20).
The starting material **19** (10 g, 16.1 mmol) was placed in a 250-mL round flask and dried on high vacuum. Dry CH_2Cl_2 (80 mL, final conc. 0.2 M), N,N-diisopropylethylamine (8.4 mL, 48.3 mmol), and 2-cyanoethyl N,N-diisopropylchloro-phosphoramidite (80 mL, 32.2 mmol, Aldrich) were then added sequentially. The reaction mixture was stirred at 0°C in an ice-water bath under dry nitrogen for 30 min and the ice bath was then removed. The mixture was further stirred for 2 hr at RT. Reaction completion was indicated by TLC [5% CH_3OH/CH_2Cl_2, product $R_f = 0.37$]. The reaction mixture was then quenched with $NaHCO_3$ (20 mL, sat.), stirred for 15 min, and extracted with CH_2Cl_2 (3x100 mL). The combined organic layer was washed with NaCl (100 mL, sat.) and dried over anhydrous $MgSO_4$ for 15 min, followed by filtration and solvent evaporation. The crude product was re-dissolved in CH_2Cl_2 (20 mL), and this solution was added dropwise to petroleum ether (1000 mL) under vigorous stirring; a white precipitate was formed. The petroleum ether solution was descanted carefully (sometimes filtration was necessary). The crude product was then loaded into a silica gel column that was equilibrated with 20% EtOAc/Hexane containing 0.5% triethylamine. The column

was eluded with an increasing step-wise gradient of EtOAc/Hexane in the presence of 0.5% triethylamine (20%, 30%, 40%, 50%, and 60%, 400-mL each). The pooled fractions containing the pure compound were combined and evaporated under reduced pressure, and re-dissolved in 20 mL of CH_2Cl_2. This solution was precipitated again in petroleum ether as indicated above. The precipitate was re-dissolved in CH_2Cl_2, transferred into a small round flask, evaporated, and dried on high vacuum overnight to yield a white foamy product (12 g, 91%).

^{1}H-NMR ($CDCl_3$) δ: 1.05-1.37 (m, 24H, 8xCH_3-ipr), 2.04 (s, 6H, 2xCH_3CO), 2.09 (2x s, 6H, 2xCH_3Se), 2.42 and 2.68 (2x t, J=7.5 Hz, 4H, 2xCH_2-CN), 3.42-3.75 (m, 12H, 4x CH-ipr, $C\underline{H_2}$-CH_2-CN, 2x H-2', 2x 2H-5'), 3.84 (s, 12H, 4x CH_3O), 3.90-4.05 (m, $C\underline{H_2}$-CH_2-CN), 4.22 and 4.29 (2x m, 2H, H-4'), 4.63-4.75 (m, 2H, H-3'), 5.27 and 5.35 (2x d, J=9.6 Hz, 2H, 2x H-5), 6.38 (d, 2H, J=8.4 Hz, 2x H-1'), 6.84-6.95 (d, J=8.7 Hz, 8H, aromatic), 7.26-7.48 (m, 18H, aromatic), 8.34 (d, J=9.8 Hz, 2H, H-6), 9.23 (br, 2H, NH).

^{13}C-NMR ($CDCl_3$) δ: 4.78 ($SeCH_3$), 19.14, 19.67, 20.87, 24.83, 24.46, 43.32, 43.48, 46.32, 47.56, 51.47 and 51.84 (C-2'), 55.61 (OCH_3), 57.05, 58.38, 62.51 (C-5'), 73.35 and 73.66 (C-3'), 84.56 (C-

4'), 87.35 (Ar.-C), 88.12 (C-1'), 103.35 (C-5), 113.42, 117.28, 127.21, 127.94, 128.36, 130.19, 130.25, 135.17, 135.38 (Ar-C), 139.46 (C-6), 144.18 (Ar-C), 150.39 (C-2), 158.81 (Ar-C), 163.19 (C-4).

^{31}P-NMR ($CDCl_3$) δ: 148.68, 149.12.

UV (in acetonitrile), λ_{max}: 236.6, 267.8 nm. IR (KBr): 3450 (br.), 3070, 3030, 2930, 2850, 1715, 1610, 1505, 1480, 1390, 1250, 1180, 1090, 1045, 785, 715, 580 cm^{-1}.

HRMS (MALDI-FTMS): $C_{40}H_{50}N_4O_8PSe$ ($M+H^+$), 825.2538 (calc. 825.2532).

5'-O-(4,4-dimethoxytrity1)-2'-methylseleno-2'-deoxy-cytidine ***(24)***.

Phosphorus oxychloride (1.1 mL, 12 mmol) was added to a solution of 1,2,4-triazole (3.32 g, 48 mmol) in dry acetonitrile (40 mL) under argon, and the reaction was stirred for 1 hr at RT. Dry triethylamine (13.4 mL, 96 mmol) was added and the reaction was stirred for another hr. The suspension was then filtered directly into a flask containing both compound **19** (2.49 g, 4 mmol) and compound 1-(trimethylsilyl)imidazole (1.17 mL, 8 mmol,) dissolved in dry acetonitrile (20 mL) under argon. The reaction mixture was allowed to

run for 2 hr. The formation of N^4-triazolide **19a** was revealed by a fluorescent spot on TLC (7.5% MeOH in CH_2Cl_2) that moved slightly slower than the silylated intermediate. After the reaction was complete, concentrated NH_3-water (20 mL) was then injected to the solution and stirred. After 15 min, the mixture was concentrated to a solution of about 10 mL. Concentrated NH_3-water (30 mL) and dioxane (30 mL) were then added to the flask, and the solution was stirred overnight (monitored on TLC, 5% MeOH in CH_2Cl_2). The reaction mixture was then evaporated to approximate 20 mL before extraction with ethyl acetate (3x100 mL). The combined organic layer was washed with saturated NaCl (100 mL) and dried over $MgSO_4$ (s) before evaporation. The crude product was purified on silica gel column equilibrated with hexane/ CH_2Cl_2, 1:1. The column was eluted with hexane/CH_2Cl_2, 1:1, 1:3, and pure CH_2Cl_2), and then with methanol/CH_2Cl_2 gradient (0.5, 1, 2, 3 and 4% methanol in CH_2Cl_2) to afford product **24** (2.13 g, Fw. 622.6) as a white foam (86% yield).

^{1}H-NMR ($CDCl_3$) δ: 1.2 (br., 1H, OH), 1.99 (s, 3H, CH_3Se), 3.38 and 3.46 (2xdd, J=1.8, 9.6 Hz, 2H, H-5'), 3.51 (dd, J=4.5, 4.8 Hz, 1H, H-2'), 3.68 (s, 6H, CH_3O), 4.13-4.17 (m, 1H, H-4'), 4.44-4.47 (m, 1H, H-3'), 5.40 (d, J=7.5 Hz, 1H, H-5), 6.32 (d, 1H, J=5.9 Hz, H-1'), 6.76-

6.82 (d, J=5.9 Hz, 4H, aromatic), 7.09-7.31 (m, 9H, aromatic), 7.87 (d, J=7.5 Hz, 1H, H-6), 12.1 (br, 2H, NH_2).

^{13}C-NMR ($CDCl_3$) δ: 5.25 ($SeCH_3$), 51.56 (C-2'), 56.23 (OCH_3), 64.50 (C-5'), 72.04 (C-3'), 85.52 (C-4'), 86.97 (Ar.-C), 88.76 (C-1'), 96.10 (C-5), 114.23, 128.84, 129.08, 129.60, 130.46, 135.73, 136.95, 142.23 (C-6), 144.05, 158.40, (Ar-C), 160.40 (C-2), 167.26 (C-4).

HRMS (MALDI-FTMS): $C_{31}H_{33}N_3O_6Se$ $[M+Na]^+$: 646.1429 (calc. 646.1427).

N^4-acetyl-5'-O-(4,4-dimethoxytrity1)-2'-methylseleno-2'-deoxycytosine (25i)

THF (10 mL), dry triethylamine (1.11 mL, 8 mmol) and 1-(trimethylsilyl)imidazole (0.44 mL, 3 mmol) were added to a 25-mL flask containing compound **24** (dry 1.25 g, 2 mmol) under argon at RT. After 15 min of stirring, a catalytic amount of N,N-dimethyl-aminopyridine (DMAP, 20 mg) and acetic anhydride (371 μL, 4 mmol) were added and the reaction was left for 2 hr while being monitored by silica gel TLC (5% MeOH/CH_2Cl_2). After the reaction was complete, MeOH (2 mL) was added, and the mixture was stirred for another 20 min. The solvents were evaporated under reduced pressure and the resultant residue was dissolved in EtOAc (50 mL).

The precipitated salt was removed by filtration and the filtrate was evaporated again under reduced pressure. This residue was dissolved in THF (10 mL), and tetrabutylammonium fluoride (4 mL, 1 M, 4 mmol) was added. The mixture was stirred for one hour to remove the 3'-TMS group (monitored on silica gel TLC in 5% MeOH/CH_2Cl_2). After evaporation of THF, the crude product was dissolved in CH_2Cl_2 and purified on silica gel column, equilibrated with hexane/CH_2Cl_2, 1:1. The column was first eluted with hexane/CH_2Cl_2, 1:1, 1:3, and pure CH_2Cl_2), and then with a methanol/CH_2Cl_2 gradient (0.5, 1, 2 and 3% methanol in CH_2Cl_2) to afford the desired product **25i** (1.26 g, Fw. 664.6) as a white foam (95% yield).

^{1}H-NMR ($CDCl_3$) δ: 2.15 (s, 3H, CH_3Se), 2.23 (s, 3H, CH_3CO), 3.29 (br., 1H, OH), 3.49 and 3.57 (2x dd, J=1.9, 9.8 Hz, 2H, H-5'), 3.68 (dd, J=4.4, 4.7 Hz, 1H, H-2'), 3.79 (s, 6H, CH_3O), 4.16-4.27 (m, 1H, H-4'), 4.42-4.53 (m, 1H, H-3'), 6.34 (d, 1H, J=6.1 Hz, H-1'), 6.81-6.95 (d, J=6.1 Hz, 4H, aromatic), 7.19 (d, J=7.5 Hz, 1H, H-5), 7.26-7.52 (m, 9H, aromatic), 8.32 (d, J=7.5 Hz, 1H, H-6), 9.89 (br, 1H, NH).

^{13}C-NMR ($CDCl_3$) δ: 4.58 ($SeCH_3$), 24.79 ($\underline{C}H_3CO$), 46.23 (C-2'), 55.23 (OCH_3), 62.23 (C-5'), 69.92 (C-3'), 84.47 (C-4'), 87.21 (Ar.-

C), 90.26 (C-1'), 96.98 (C-5), 113.35, 127.17, 128.02, 128.15, 130.08, 135.20, 135.40, 144.24 (Ar-C), 144.45 (C-6), 155.37 (C-2), 158.76 (Ar-C), 162.85 (C-4), 170.75 ($\underline{C}OCH_3$).

HRMS (MALDI-FTMS): $C_{33}H_{35}N_3O_7Se$ $[M+Na]^+$: 688.1553 (calc. 688.1532).

N^4-acetyl-3'-O-(2-cyanoethyl-N,N-diisopropylphosphor-amidite)-5'-O-dimethoxytrityl-2'-methylseleno-2'-deoxycytosine (25d)
See the synthesis of compound **20**. A white foamy product resulted (Fw.: 864.8; 1.28 g, 92%).

^{1}H-NMR ($CDCl_3$) δ: 1.04-1.35 (m, 24H, 8xCH_3-ipr), 2.13 (s, 6H, 2xCH_3CO), 2.16 and 2.19 (2x s, 6H, 2xCH_3Se), 2.41 and 2.67 (2x t, J=7.5 Hz, 4H, 2xCH_2-CN), 3.45-3.78 (m, 12H, 4x CH-ipr, $C\underline{H_2}$-CH_2-CN, 2x H-2', 2x 2H-5'), 3.83 (s, 12H, 4x CH_3O), 3.92-4.03 (m, $C\underline{H_2}$-CH_2-CN), 4.36-4.41 (m, 2H, H-4'), 4.65-4.71 (m, 2H, H-3'), 6.44 (d, 2H, J=5.1 Hz, H-1'), 6.82-6.94 (d, J=8.7 Hz, 8H, aromatic), 7.09 and 7.14 (2x d, J=7.5 Hz, 2H, H-5), 7.27-7.46 (m, 18H, aromatic), 8.31 and 8.35 (2x d, J=7.5 Hz, 2H, H-6), 10.56 (br, 2H, NH).

^{13}C-NMR ($CDCl_3$) δ: 4.65 ($SeCH_3$), 19.12, 19.89, 20.56, 22.94 and 22.96 (CH_3CO), 24.77, 24.83, 43.34, 43.47, 45.32 and 45.38 (C-2'), 46.77, 47.64, 55.25 (OCH_3), 57.95, 58.34, 61.94 (C-5'), 73.32 and

73.46 (C-3'), 84.36 (C-4'), 87.15 (Ar.-C), 90.92 (C-1'), 97.22 (C-5), 113.32, 117.25, 127.23, 127.99, 128.30, 130.09, 130.23, 135.14, 13 5.32, 144.16 (Ar-C), 144.46 (C-6), 155.39 (C-2), 158.81 (Ar-C), 163.09 (C-4), 171.20 ($\underline{C}OCH_3$).

^{31}P-NMR ($CDCl_3$) δ: 148.67, 149.05.

HRMS (MALDI-FTMS): $C_{42}H_{52}N_5O_8PSe$ $[M+H]^+$: 866.2795 (calc. 866.2791).

9.4 Synthesis of 2'-Se-functionalized RNA and DNA Oligonucleotides

All the DNA and RNA oligonucleotides were synthesized chemically on a 1.0 μmol scale using an Applied Biosystems 392 DNA/RNA Synthesizer as reported previously (24,25) with the following modifications: the concentration of the nucleoside selenium phosphoramidites was identical to that of the conventional phosphoramidites, 0.1 M in acetonitrile. Coupling was carried out using a 0.3 M solution of 5-(benzylmercapto)-1H-tetrazole (5-BMT) in acetonitrile (26). The coupling time for DNA was 20 sec, and that for RNA was 15 min; 5'-detritylation was done in 0.3 M trichloroacetic acid in methylene chloride. Syntheses were performed on polystyrene supports (CPG-500) connected with the appropriate

nucleoside through a succinate linker (Glen Research). All the oligonucleotides were prepared DMTr-on. In the case of RNA, the syntheses were done with nucleoside phosphoramidites containing the 2-*O*-Triisopropylsilyloxymehtyl (TOM) functionalities. After synthesis, the DNA oligonucleotides were released from the solid support and fully deprotected by aqueous ammonia treatment for 14 hr at 55°C. Similarly, the 2'-*O*-TOM protected RNA oligonucleotides were deprotected following recommendations from Glen Research. Briefly, 1.5 mL of a methylamine solution (prepared by mixing 40% aqueous methylamine with 33% ethanolic methylamine (Fluka) in a 1:1 ration) was added to an eppendorf tube containing the appropriate RNA sample, followed by incubation for 6 hr at 35°C. The solution was chilled in a -20°C freezer for 10 min and then transferred into two sterile eppendorf tubes. After evaporating the solvents in speed vacuum, each portion was treated with 25 equivalents of tetrabutylammonium fluoride (1.0 M in THF, Fluka) and incubated for 6 hr at 35°C. Following evaporation of THF, the portions were combined in 1.0 mL of 1.0 M Tris-HCl buffer (RNase-free, pH 7.5), incubated with shaking overnight at 25°C, and filtrated through a 0.20 μm pore-sized filter. The 5'-DMTr deprotection step of both DNA and

RNA oligonucleotides was performed in a 2% trichloroacetic acid solution (from 10% w/w, 0.9 M in water) for 1.5 min, followed by neutralization to pH 7.0 with a freshly made 1.1 M aqueous solution of triethylamine.

9.5 HPLC Analysis and Purification

The DNA and RNA oligonucleotides were analyzed and purified by reverse phase high performance liquid chromatography (RP-HPLC) both DMTr-on and DMTr-off. Purification was carried out in a 21.2 mm X 250 mm Zorbax, RX-C8 column at a flow rate of 10 mL/min. Buffer A consisted of 50 mM triethylammonium acetate (TEAAc, pH 7.1, RNase-free water), while buffer B contained 50% aqueous acetonitrile and 50 mM TEAAc, pH 7.1. Similarly, analysis was performed on a Zorbax SB-C18 column (4.6 mm X 250 mm) at a flow of 1.0 mL/min using the same buffer system. DMTr-on oligonucleotides were eluded with a linear gradient of up to 90% buffer B in 25 min, while DMTr-off oligonucleotides were eluted using a gradient of up to 40% of buffer B in the same period of time. The appropriate solutions of pure products were lyophilized, the compounds re-dissolved in RNase-free water, and pH adjusted to 7.0.

9.6 Electrospray Mass Spectrometry Analysis

Crude and purified oligonucleotides containing the selenium derivatization were analyzed by LC-MS using electrospray negative ion mode. The general analytical procedures for the liquid chromatography were elution (1 mL/min, 4.5 X 150 mm 300SB-C8 column) with buffer A (5 mM ammonium acetate, pH 6.5) for two minutes, and then elution with a linear gradient from buffer A to 100% buffer B (60% acetonitrile and 40% of buffer A) in 13 minutes.

9.7 Thermodenaturization of Duplex DNAs

Solutions of the duplex DNAs (2 μM) were prepared by dissolving the DNAs in a buffer containing NaCl (90 mM), sodium phosphate (10 mM, pH 7.2), and EDTA (1mM). The solutions were then heated to 95°C for 2 min, cooled slowly to room temperature, and stored at 5°C overnight before measurement. Prior to thermal denaturation, helium was bubbled through the samples. Denaturation curves were acquired at 254 nm at a heating rate of 0.5°C/min using 8453 UV-Visible Spectrometer from Agilent Technologies. This system is equipped with a Peltier Temperature Controller. The data were analyzed in accordance with the convention of Puglisi and Tinoco (*143*).

9.8 Crystallization of Se-derivatized Oligonucleotides

The purified oligonucleotides (2 mM) containing selenium labels were first heated to 90°C for 1 minute, and the samples were then allowed to slowly cool to 25°C. Crystallization conditions were first screened using the nucleic acid screening kits from Hampton Research. To minimize the amount of material used, 1 or 2 μL of the appropriate oligonucleotide solutions were typically used in each of these screens. Crystallization was carried out using the hanging drop method by vapor diffusion at 25°C and at 4°C.

9.9 Synthesis of α-Se-Nucleoside Triphosphates (58)

Both diastereomers of TTPαSe were synthesized by modifying the method reported by Ludwig and Eckstein for the synthesis of nucleoside 5'-thiotriphoshates (*133*). Compound **54** (28.46 mg, 0.10 mmol) was placed in a 5-mL round flask, dried on high vacuum overnight and dissolved in a mixture of freshly distilled pyridine (0.10 mL) and dioxane (0.30 mL). The resultant solution was then injected drop-wise in a course of about 5 minutes into another 5-mL round flask containing a solution of 2-chloro-4H-1,3,2-benzodioxaphosphorin-4-one (21.0 mg, 0.10 mmol, 1 eq.) in dioxane

(0.10 mL). The reaction was stirred at room temperature under a constant flow of dry argon for 10 min.

A solution of pyrophosphate (64.1 mg, 0.14 mmol, 1.5 eq.) in dry DMF (0.28 mL) containing tributylamine (0.10 mL) was then injected and the reaction was stirred for another 10 min. 3H-1,2-benzothiaselenol-3-one (43.0 mg, 0.20 mmol, 1.5 eq.) was dissolved in dioxane (0.22 mL), triethylamine (42 μL, 0.3 mmol, 3 eq.) and TMS-Cl (1.0 eq.) were then added, and this solution was injected into the reaction flask. After selenization was complete (30 min, monitored by ^{31}P-NMR), the reaction was quenched with water (1.0 mL) for 2 hr.

The acetyl protecting group was then removed by aqueous ammonia (conc. 3.0 mL) at 60°C for 1.5 hr. Most of the solvents were removed by rotary evaporation, and the pH was adjusted to 7.0. A solution of DTT was then added to a final concentration of 20 mM, and the crude product was transferred into a 15-mL centrifuge tube and spun for 3 min to remove the selenium metal. The supernatant was transferred into another centrifuge tube, NaCl (3.0 M, 1.3 mL) was added, and the content was divided into two equal portions. Three volumes of absolute ethanol (thoroughly purged with argon) were added to each portion and the samples were placed in a -20°C freezer

for 10 min. before centrifugation (10 min. at 6000 rpm). After removing the supernatant, the precipitate was re-dissolved in water and purified by RP-HPLC using Zorbox C18 column (9.4 x 250 mm). Samples were eluded (5 mL/min) with a linear gradient from buffer A [10 mM triethylammonium acetate (TEAAc), pH 6.5] to 20% buffer B (30% acetonitrile in water, 10 mM TEAAc, pH 6.5) in 25 min. The purified TTPαSe diastereomers were then analyzed by RP-HPLC and by HRMS and stored at -80°C in a solution of 10 mM Tris-Cl (pH 7.5) containing 20 mM of DTT. This general procedure was used to synthesize and purify the diastereomers of ATPαSe and UTPαSe.

^{1}H-NMR (D_2O) δ: 1.86 (s, 3H, 5-CH_3), 2.38-2.30 (br, 2H, H-2'), 4.12-4.16 (m, 2H, H-5'), 4.19-4.21(m, 1H, H-3'), 4.61-4.62 (m, 1H, H-4'), 6.23-6.29 (m, 1H, H-1'), 7.65-7.68 (m, 1H, H-6).

^{31}P-NMR (85% H_3PO_4 in D_2O): see Table 6, pg. 101.

UV (H_2O, λ_{max}) TTPαSe I, 267.6; TTPαSe II, 267.6 nm.

HRMS (MALDI-FTMS): $C_{10}H_{17}N_2O_{13}P_3Se$ ($M - H^+$), 544.9037 (calc. 544.9036).

REFERENCES:

1. Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E., and Cech, T. R. (1982) *Cell 31*, 147-157.

2. Hegg, L. A., and Fedor, M. J. (1995) *Biochemistry 34*, 15813-15828.

3. Herschlag, D., and Cech, T. R. (1990) *Biochemistry 29*, 10159-10171.

4. Hertel, K. J., Herschlag, D., and Uhlenbeck, O. C. (1994) *Biochemistry 33*, 3374-3385.

5. Cech, T. R. (1985) *International Review of Cytology 93*, 3-22.

6. Joyce, G. F. (1989) *Nature 338*, 217-224.

7. Johnston, W. K., Unrau, P. J., Lawrence, M. S., Glasner, M. E., and Bartel, D. P. (2001) *Science 292*, 1319-1325.

8. Zhang, B., and Cech, T. R. (1997) *Nature 390*, 96-100.

9. Wilson, C., and Szostak, J. W. (1995) *Nature 374*, 777-782.

10. Lohse, P. A., and Szostak, J. W. (1996) *Nature 381*, 442-444.

11. Lee, N., Bessho, Y., Wei, K., Szostak, J. W., and Suga, H. (2000) *Nature Structural Biology 7*, 28-33.

12. Tarasow, T. M., Tarasow, S. L., and Eaton, B. E. (1997) *Nature 389*, 54-57.

13. Baskerville, S., and Bartel, D. P. (2002) *PNAS 99*, 9154-9159.

14. Lee, R. C., and Ambros, V. (2001) *Science 294*, 862-864.

15. Lagos-Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. (2001) *Science 294*, 853-858.

16. Lee, R. C., Feinbaum, R. L., and Ambros, V. (1993) *Cell 75*, 843-854.

17. Pasquinelli, A. E., Reinhart, B. J., Slack, F., Martindale, M. Q., Kuroda, M. I., Maller, B., Hayward, D. C., Ball, E. E., Degnan, B., and Muller et, a. (2000) *Nature 408*, 86-89.

18. Storz, G. (2002) *Science 296*, 1260-1263.

19. Platt, T. (1998) *RNA Structure and Function*, Cold Spring Harbor Laboratory Press, New York.

20. Watson, J., and Crick, F. (1953) *Nature 171*, 737-738.

21. Stryer, L. (1995) *Biochemistry*, 4th ed., W.H. Freeman and Company, New York.

22. Kim, S., Sussman, J. L., Suddath, F., Quigley, G., McPherson, A., Wang, A., Seeman, N., and Rich, A. (1974) *Proc. Nat. Acad. Sci. USA 71*, 4970-4974.

23. Woese, C. R., Magrum, L. J., Gupta, R., Siegel, R. B., Stahl, D. A., Kop, J., Crawford, N., Brosius, J., Gutell, R., and Hogan et, a. (1980) *Nucleic Acids Research 8*, 2275-2293.

24. Noller, H. F., Kop, J., Wheaton, V., Brosius, J., Gutell, R. R., Kopylov, A. M., Dohme, F., Herr, W., Stahl, D. A., and Gupta et, a. (1981) *Nucleic Acids Research 9*, 6167-6189.

25. Noller, H. F., and Woese, C. R. (1981) *Science 212*, 403-411.

26. Michel, F., and Westhof, E. (1990) *Journal of Molecular Biology 216*, 585-610.

27. Pyle, A. M., Murphy, F. L., and Cech, T. R. (1992) *Nature 358*, 123-128.

28. Michel, F., Hanna, M., Green, R., Bartel, D. P., and Szostak, J. W. (1989) *Nature 342*, 391-395.

29. Wang, J. F., Downs, W. D., and Cech, T. R. (1993) *Science 260*, 504-508.

30. Szewczak, A. A., Ortoleva-Donnelly, L., Ryder, S. P., Moncoeur, E., and Strobel, S. A. (1998) *Nature Structural Biology 5*, 1037-1042.

31. Strobel, S. A., and Shetty, K. (1997) *Proceedings of the National Academy of Sciences, USA 94*, 2903-2908.

32. Capel, M. S., Engelman, D. M., Freeborn, B. R., Kjeldgaard, M., Langer, J. A., Ramakrishnan, V., Schindler, D. G., Schneider, D. K., Schoenborn, B. P., and Sillers et, a. (1987) *Science 238*, 1403-1406.

33. Capel, M. S., Kjeldgaard, M., Engelman, D. M., and Moore, P. B. (1988) *Journal of Molecular Biology 200*, 65-87.

34. Mueller, F., Sommer, I., Baranov, P., Matadeen, R., Stoldt, M., Wohnert, J., Gorlach, M., van Heel, M., and Brimacombe, R. (2000) *Journal of Molecular Biology 298*, 35-59.

35. Hoogstraten, C. G., Legault, P., and Pardi, A. (1998) *Journal of Molecular Biology 284*, 337-350.

36. Cai, Z., and Tinoco, I., Jr. (1996) *Biochemistry 35*, 6026-6036.

37. Mirmira, S. R., and Tinoco, I., Jr. (1996) *Biochemistry 35*, 7664-7674.

38. Butcher, S. E., Allain, F. H., and Feigon, J. (1999) *Nature Structural Biology 6*, 212-216.

39. Kim, S. H., and Rich, A. (1968) *Science 162*, 1381-1384.

40. Kim, S. H., Quigley, G., Suddath, F. L., and Rich, A. (1971) *Proceedings of the National Academy of Sciences, USA 68*, 841-845.

41. Kim, S. H., Suddath, F. L., Quigley, G. J., McPherson, A., Sussman, J. L., Wang, A. H., Seeman, N. C., and Rich, A. (1974) *Science 185*, 435-440.

42. Kim, S. H., Sussman, J. L., Suddath, F. L., Quigley, G. J., McPherson, A., Wang, A. H., Seeman, N. C., and RICH, A. (1974) *Proceedings of the National Academy of Sciences, USA 71*, 4970-4974.

43. Robertus, J. D., Ladner, J. E., Finch, J. T., Rhodes, D., Brown, R. S., Clark, B. F., and Klug, A. (1974) *Nature 250*, 546-551.

44. Stout, C. D., Mizuno, H., Rubin, J., Brennan, T., Rao, S. T., and Sundaralingam, M. (1976) *Nucleic Acids Research 3*, 1111-1123.

45. Sussman, J. L., Holbrook, S. R., Warrant, R. W., Church, G. M., and Kim, S. H. (1978) *Journal of Molecular Biology 123*, 607-630.

46. Holbrook, S. R., Sussman, J. L., Warrant, R. W., and Kim, S. H. (1978) *Journal of Molecular Biology 123*, 631-660.

47. Holbrook, S. R., and Kim, S. H. (1997) *Biopolymers 44*, 3-21.

48. Holbrook, S. R., Holbrook, E. L., and Walukiewicz, H. E. (2001) *Cell Mol Life Sci 58*, 234-43.

49. Pley, H. W., Flaherty, K. M., and McKay, D. B. (1994) *Nature 372*, 68-74.

50. Scott, W. G., Finch, J. T., and Klug, A. (1995) *Cell 81*, 991-1002.

51. Ferre-D'Amare, A. R., Zhou, K., and Doudna, J. A. (1998) *Nature 395*, 567-574.

52. Rupert, P. B., and Ferre-D'Amare, A. R. (2001) *Nature 410*, 780-6.

53. Golden, B. L., Gooding, A. R., Podell, E. R., and Cech, T. R. (1998) *Science 282*, 259-264.

54. Ban, N., Nissen, P., Hansen, J., Moore, P. B., and Steitz, T. A. (2000) *Science 289*, 905-920.

55. Wedekind, J. E., and McKay, D. B. (1999) *Nature Structural Biology 6*, 261-268.

56. Klosterman, P. S., Tamura, M., Holbrook, S. R., and Brenner, S. E. (2002) *Nucleic Acids Res 30*, 392-4.

57. Milligan, J. F., Groebe, D. R., Witherell, G. W., and Uhenbeck, O. C. (1987) *Nucleic Acids Research 15*, 8783-8798.

58. Wyatt, J. R., Chastein, M., and Puglisi, J. D. (1991) *Biotechniques 11*, 764-769.

59. Kim, R., Holbrook, E. L., and Jancarik, S.-H. K. (1995) *BioTechniques 18*, 992-994.

60. Golden, B. L., Podell, E. R., Gooding, A. R., and Cech, T. R. (1997) *Journal of Molecular Biology 270*, 711-23.

61. Ferre-D'Amare, A. R., Zhou, K., and Doudna, J. A. (1998) *J Mol Biol 279*, 621-31.

62. Doudna, J. A., Grosshans, C., Gooding, A., and Kundrot, C. E. (1993) *Proc Natl Acad Sci U S A 90*, 7829-7833.

63. Ferre-D'Amare, A. R., and Doudna, J. A. (1996) *Nucleic Acids Research 24*, 977-8.

64. Schurer, H., Lang, K., Schuster, J., and Morl, M. (2002) *Nucleic Acids Research 30*, e56.

65. Anderson, A. C., S.A., S., Earp, B. E., and Frederick, C. A. (1996) *RNA 2*, 110-117.

66. Wahl, M. C., Ramakrishnan, B., Ban, C., Chen, X., and Sundaralingam, M. (1996) *Acta Crystallographica, Section D 52*, 668-675.

67. Usman, N., Egli, M., and Rich, A. (1992) *Nucleic Acids Research 20*, 6695-6699.

68. Moffat, K., and Ren, Z. (1997) *Current Opinion in Structural Biology 7*, 689-696.

69. Hendrickson, W. A. (2000) *Trends in Biochemical Sciences 25*, 637-643.

70. Guss, J. M., Merritt, E. A., Phizackerley, R. P., Hedman, B., Murata, M., Hodgson, K. O., and Freeman, H. C. (1988) *Science 241*, 806-811.

71. Minor, W., Tomchick, D., and Otwinowski, Z. (2000) *Structure with Folding & Design 8*, R105-R110.

72. Hauptman, H. (1997) *Current Opinion in Structural Biology 7*, 672-680.

73. Boggon, T. J., Helliwell, J. R., Judge, R. A., Olczak, A., Siddons, D. P., Snell, E. H., and Stojanoff, V. (2000) *Acta Crystallographica. Section D, Biological Crystallography 56 (Pt 7)*, 868-880.

74. Hendrickson, W. A. (1991) *Science 254*, 51-8.

75. Hendrickson, W. A., Horton, J. R., and LeMaster, D. M. (1990) *Embo J 9*, 1665-72.

76. Hendrickson, W. A., Pahler, A., Smith, J. L., Satow, Y., Merritt, E. A., and Phizackerley, R. P. (1989) *Proc Natl Acad Sci U S A 86*, 2190-4.

77. Yang, W., Hendrickson, W. A., Crouch, R. J., and Satow, Y. (1990) *Science 249*, 1398-405.

78. Ogata, C. M. (1998) *Nature Structural Biology 5*, 638-640.

79. Hendrickson, W. A. (1999) *Journal of Synchrotron Radiation 6*, 845-851.

80. Ealick, S. E. (2000) *Current Opinion in Chemical Biology 4*, 495-499.

81. Walsh, M. A., Evans, G., Sanishvili, R., Dementieva, I., and Joachimiak, A. (1999) *Acta Crystallographica. Section D, Biological Crystallography 55 (Pt 10)*, 1726-1732.

82. Wery, J.-P., and Schevitz, R. W. U. (1997) *Current Opinion in Chemical Biology 1*, 365-369.

83. Doublie, S. (1997) *Methods in Enzymology 276*, 523-530.

84. Bellizzi, J. J., Widom, J., Kemp, C. W., and Clardy, J. (1999) *Structure with Folding & Design 7*, R263-R267.

85. Bushnell, D. A., Cramer, P., and Kornberg, R. D. (2001) *Structure with Folding & Design 9*, R11-R14.

86. Budisa, N., Karnbrock, W., Steinbacher, S., Humm, A., Prade, L., Neuefeind, T., Moroder, L., and Huber, R. (1997) *Journal of Molecular Biology 270*, 616-623.

87. Bae, J. H., Alefelder, S., Kaiser, J. T., Friedrich, R., Moroder, L., Huber, R., and Budisa, N. (2001) *Journal of Molecular Biology 309*, 925-936.

88. Ferre-D'Amare, A. R., and Doudna, J. A. (2000) *J Mol Biol 295*, 541-56.

89. Batey, R. T., Sagar, M. B., and Doudna, J. A. (2001) *J Mol Biol 307*, 229-46.

90. Nikulin, A., Serganov, A., Ennifar, E., Tishchenko, S., Nevskaya, N., Shepard, W., Portier, C., Garber, M., Ehresmann, B., and Ehresmann et, a. (2000) *Nature Structural Biology 7*, 273-277.

91. Shah, K., Wu, H., and Rana, T. M. (1994) *Bioconjugate Chemistry 5*, 508-512.

92. Hung, L. W., Holbrook, E. L., and Holbrook, S. R. (2000) *Proceedings of the National Academy of Sciences, USA 97*, 5107-5112.

93. Deng, J., Xiong, Y., and Sundaralingam, M. (2001) *Proceedings of the National Academy of Sciences, USA 98*, 13665-13670.

94. Shui, X., Peek, M. E., Lipscomb, L. A., Wilkinson, A. P., Williams, L. D., Gao, M., Ogata, C., Roques, B. P., Garbay-Jaureguiberry, C., and Wilkinson et, a. (2000) *Current Medicinal Chemistry 7*, 59-71.

95. Gott, J. M., Willis, M. C., Koch, T. H., and Uhlenbeck, O. C. (1991) *Biochemistry 30*, 6290-6295.

96. Ennifar, E., Carpentier, P., Ferrer, J. L., Walter, P., and Dumas, P. (2002) *Acta Crystallographica. Section D, Biological Crystallography 58*, 1262-1268.

97. Ching, W.-M., and Stadtman, T. C. (1982) *Proc. Nat. Acad. Sci. USA 79*, 374-377.

98. Wittwer, A. J. (1983) *J. Biol. Chem. 258*, 8637-8641.

99. Veres, Z., Tsai, L., Scholz, T. D., Politino, M., Balaban, R. S., and Stadtman, T. C. (1992) *Proc. Nat. Acad. Sci. USA 89*, 2975-2979.

100. Teplova, M., Wilds, C. J., Wawrzak, Z., Tereshko, V., Du, Q., Carrasco, N., Huang, Z., and Egli, M. (2002) *Biochimie 84*, 849-858.

101. Scaringe, S., Francklyn, C., and Usman, N. (1990) *Nucleic Acids Res 18*, 5433-5441.

102. Mitsunobu, O. (1981) *Synthesis*.

103. Klayman, D., and Griffin, T. (1973) *J. Am. Chem. Soc. 95*, 197-199.

104. Syper, L., and Mlochowski, J. (1984) *Synthesis*, 439-442.

105. Cosford, N., and Schinazi, R. (1991) *J. Org. Chem. 56*, 2161-2165.

106. Carrasco, N., Ginsburg, D., Du, Q., and Huang, Z. (2001) *Nucleosides Nucleotides Nucleic Acids 20*, 1723-34.

107. McGee, D. P. C., Vanghn-Settle, A., Vargeese, C., and Zhai, Y. (1996) *J. Org. Chem. 61*, 781-785.

108. Du, Q., Carrasco, N., Teplova, M., Wilds, C. J., Egli, M., and Huang, Z. (2002) *J Am Chem Soc 124*, 24-5.

109. Hamm, M. L., and Piccirilli, J. A. (1997) *J. Org. Chem. 62*, 3415-3420.

110. Ogawa, A., Tanaka, M., Sasaki, T., and Matsuda, A. (1998) *J. Med. Chem. 41*, 5094-5107.

111. Egli, M., Tereshko, V., Teplova, M., Minasov, G., Joachimiak, A., Sanishvili, R., Weeks, C. M., Miller, R., Maier, M. A., An, H., and Cook, P. D. (2000) *Biopolymers, Nucleic Acid Sciences 48*, 234–252.

112. Welz, R., and Mueller, S. (2002) *Tetrahedron Letters 43*, 793-797.

113. Boggon, T. J., Hancox, E. L., McAuley-Hecht, K. E., Connolly, B. A., Hunter, W. N., Brown, T., Walker, R. T., and Leonard, G. A. (1996) *Nucleic Acids Research 24*, 951-961.

114. De Mesmaeker, A., Häner, R., Martin, P., and Moser, H. E. (1995) *Acc. Chem. Res. 28*, 366–374.

115. Carrasco, N., Buzin, Y., Tyson, E., Halpert, E., and Huang, Z. (2004) *Nucleic Acids Research 32*, 1638-1646.

116. Buzin, Y., Carrasco, N., and Huang, Z. (2004) *Organic Letter 6*, 1099-1102.

117. Jankowska, J., Sobkowska, A., Cieslak, J., Sobkowski, M., Kraszewski, A., Stawinski, J., and Shugar, D. (1998) *J. Org. Chem. 63*, 8150-8156.

118. Jankowska, J., Sobkowski, M., Stawinski, J., and Kraszewski, A. (1994) *Tetrahedron Letters 35*, 3355-3358.

119. He, K., Sergueev, D., Sergueeva, Z., and Shaw, B. (1999) *Tetrahedron Letters 40*, 4601-4604.

120. Yoshikawa, M., Kato, T., and Takenishi, T. (1969) *Bull. Chem. Soc. Jpn 42*, 3505-3508.

121. Ikemoto, T., Haze, A., Hatano, H., Kitamoto, Y., Ishida, M., and Nara, K. (1995) *Chem. Pharm. Bull 43*, 210-215.

122. Murray, A. W., and Atkinson, M. R. (1968) *Biochemistry 11*, 4023-4029.

123. Eckstein, F. (1985) *Annu Rev Biochem 54*, 367-402.

124. Lindh, I., and Stawinski, J. (1989) *J. Org. Chem. 54*, 1338-1342.

125. Mori, K., Boiziau, C., Cazenave, C., Marsukura, M., Subasinghe, C., Cohen, J., Broder, S., Toulme′, J., Stein, C., and . (1989) *Nucleic Acids Research 17*, 8207-8219.

126. Stawinski, J., and Thelin, M. (1992) *Tetrahedron Letters 33*, 7255-7258.

127. Stawinski, J., and Thelin, M. (1994) *J. Org. Chem. 59*, 130-136.

128. Stec, W., Zon, G., Egan, W., and Stec, B. (1992) *Tetrahedron Letters 33*, 7255-7258.

129. Bollmark, M., and Stawinski, J. (2001) *Chem. Commun*, 771-772.

130. Stawinski, J., and Thelin, M. (1991) *J. Org. Chem. 56*, 5169-5175.

131. Wilds, C., Pattanayek, R., Pan, C., Wawrzak, Z., and Egli, M. (2002) *J. Am. Chem. Soc. 124*, 14910-14916.

132. Wolf, W. M., and Bartczak, T. J. (1989) *Acta Crystallographica, Section C 45*, 1767-1770.

133. Ludwig, J., and Eckstein, F. (1989) *Journal of Organic Chemistry 54*, 631-635.

134. He, K., Hasan, A., Krzyzanowska, B., and Shaw, B. R. (1998) *J Org Chem 63*, 5769-5773.

135. Lin, J., and Shaw, B. R. (2001) *Nucleosides Nucleotides Nucleic Acids 20*, 1325-8.

136. He, K., Porter, K. W., Hasan, A., Briley, Jd, and Shaw, B. R. (1999) *Nucleic Acids Res 27*, 1788-94.

137. Eckstein, F., and Gish, G. (1989) *Trends Biochem Sci 14*, 97-100.

138. Li, H., Porter, K., Huang, F., and Shaw, B. R. (1995) *Nucleic Acids Res 23*, 4495-501.

139. Rogers, G. G., and Weiss, B. (1980) *Methods in Enzymology*, Vol. 65, Academic Press: New York.

140. Gupta, A., DeBrosse, C., and Benkovic, S. J. (1982) *J. Biol. Chem. 257*, 7689-7692.

141. Joyce, C. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 1619-1622.

142. Carrasco, N., and Huang, Z. (2004) *J. Am. Chem. Soc 126*, 448-449.

143. Puglisi, J. D., and Tinoco, I., Jr. (1989) *Methods in Enzymology 180*, 304-325.