Developing Alternative Genetic Systems for

Structural DNA Nanotechnology

and Darwinian Evolution

By

Su Zhang

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved November 2011 by the
Graduate Supervisory Committee:

John C. Chaput, Chair
Giovanna Ghirlanda
Hao Yan

ARIZONA STATE UNIVERSITY

December 2011

ABSTRACT

A major goal of synthetic biology is to recapitulate emergent properties of life. Despite a significant body of work, a longstanding question that remains to be answered is how such a complex system arose? In this dissertation, synthetic nucleic acid molecules with alternative sugar-phosphate backbones were investigated as potential ancestors of DNA and RNA. Threose nucleic acid (TNA) is capable of forming stable helical structures with complementary strands of itself and RNA. This provides a plausible mechanism for genetic information transfer between TNA and RNA. Therefore TNA has been proposed as a potential RNA progenitor. Using molecular evolution, functional sequences were isolated from a pool of random TNA molecules. This implicates a possible chemical framework capable of crosstalk between TNA and RNA. Further, this shows that heredity and evolution are not limited to the natural genetic system based on ribofuranosyl nucleic acids. Another alternative genetic system, glycerol nucleic acid (GNA) undergoes intrasystem pairing with superior thermalstability compared to that of DNA. Inspired by this property, I demonstrated a minimal nanostructure composed of both left- and right-handed mirro image GNA. This work suggested that GNA could be useful as promising orthogonal material in structural DNA nanotechnology.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ix

Chapter 1

INTRODUCTION

The idea that DNA was the molecular structure of life began when James

Watson and Francis Crick elucidated the DNA double helix (Figure 1.1)

(Watson and Crick 1953). Six decades later, we have learned that DNA, in

addition to serving as the "book of life", has many other functions. Nucleic

acids have provided us with useful tools for exploring and manipulating the

activities of complex living system as well as creating potential therapeutic

agents that can be used in the treatment of many human diseases. For example,

the selective amplification of genomic fragments by the polymerase chain

reaction (PCR) using designed synthetic DNA primers has become a common

technology in genome research (Bartlett and Stirling 2003). More recently,

inspired by naturally occurring gene regulation approach, antisense strategy

has been developed to control the gene expression in living cells (Eguchi et al.

1991). All these applications rely on efficient sequence-specific chemical

(Khudyakov and Fields 2002; Muller et al. 2004) and enzymatic synthesis

(Milligan and Uhlenbeck 1989) of RNA and DNA as well as the careful design

of the oligonucleotide that binds the target with high affinity and specificity

following the simple rules of Watson-Crick base pairing (Watson and Crick

1953).
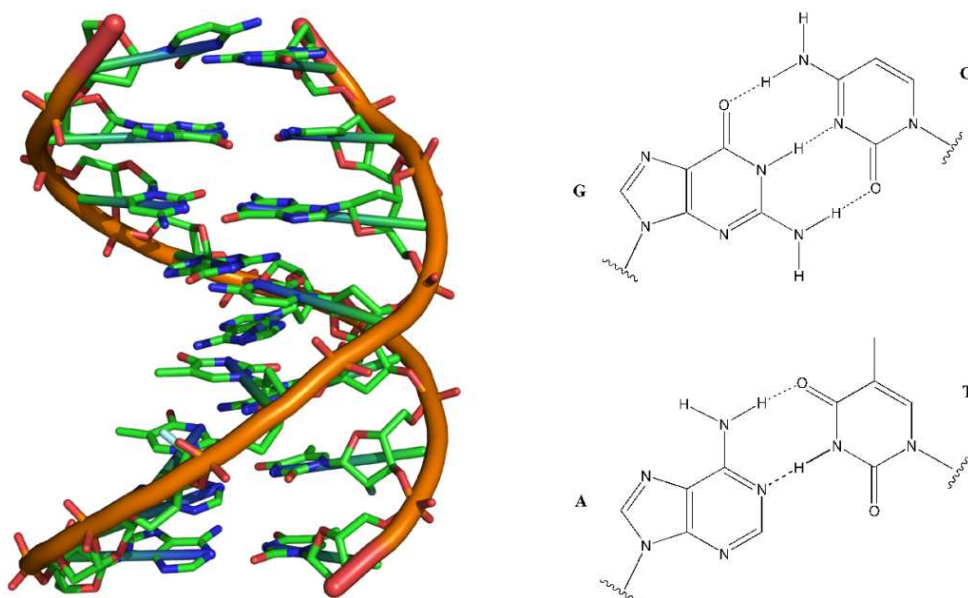
Figure 1.1. Structure of DNA (Watson and Crick model). A) A generic B-form DNA duplex. The crystal structure was obtained from Protein Data Bank (PDB code 436D). The figure was obtained using Pymol. B) The canonical Watson-Crick base pairs.

However, it is becoming increasingly clear that as molecular recognition elements, nucleic acids could be far more useful than forming simple Watson-Crick base pairing structures. A type of special mRNA molecules called riboswitches effects gene expression by recognizing small molecule targets (Nahvi et al. 2002). Moreover, the natural components of nucleic acids have been altered or integrated with other functional and structural elements to expand the utility of nucleic acids into versatile fields such as building probes (Wagenknecht 2008) and material with novel functionalities (Yan 2004) and creating nucleic acids with enhanced or entirely new functions (Famulok 2009; Wilson and Szostak 1999).

This Ph.D. thesis examines the use of synthetic nucleic acid analogs as building blocks for DNA nanotechnology and Darwinian evolution. The

2

introductory chapter focuses on the progress of non-natural nucleic acids in

structural DNA nanotechnology. Templated synthesis of alternative nucleic

acid analogs is then reviewed. Last, the topic of in vitro directed evolution

using synthetic information-carrying polymers is disscussed. The discussion

here represents a preference and interest of the author and does not intend to

be comprehensive.

1.1 Unnatural Nucleic Acids in Structural DNA nanotechnology

In a bottom-up approach, DNA nanotechnology utilizes the specificity of

Watson-Crick base pairing to construct well-defined structures out of DNA

and other nucleic acids and organize material with nanometer precision (Lin et

al 2009). Moreover, DNA nanoarchitectures have been used as probes in

biophysical study (Sacca et al. 2008; Nangreave et al. 2011). While a

significant effort has been made to expand the programmability of natural

B-form DNA (Figure 1.2), very little attention has been given to the use of

unnatural nucleic acids as building blocks for nanotechnology. By expanding

DNA nanotechnology to include alternative polymers, it should be possible to

create nanostructures with chemical and physical properties not found in

natural DNA.

The early work involving unnatural nucleic acids in DNA structural motif

was conducted by Seeman (2004) and Yan (2006), who cleverly and carefully

designed DNA double crossover (DX) molecules and incorporated short
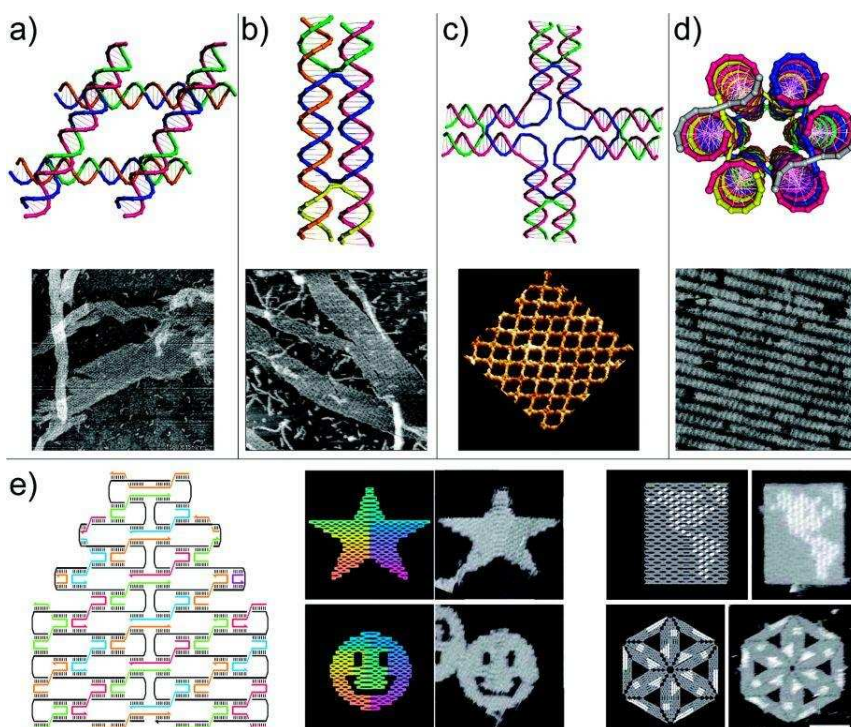
Figure 1.2 . Models of some representative DNA tiles and their assemblies into periodic 2D arrays. (a) Parallelogram DNA tile formed by joining four Holliday junctions in parallel. (b) Double-helix (DX) tile formed through strand exchange between two DNA duplexes. (c) A cross-shaped tile with four arms (4×4 tile). Each arm represents a four-arm junction. (d) Six-helix bundle tube tile viewed from the end of the tube. For panel a–d, representative AFM images of the 2D arrays are shown below the corresponding cartoon models. (e) DNA origami: (left) principle of DNA origami, folding long ssDNA into shapes by multiple helper strands; (middle) star and smiley face DNA origami tiles self-assembled by folding a 7 kb ssDNA with more than 200 helper strands; and (right) hairpin loops (white dots) that can be introduced into certain helper strands to accurately display designated geometries on the fully addressable origami tiles. Reprinted (adapted in part or whole) with permission from Biochemistry (Lin et al. 2009). Copyright 2009 American Chemical Society.
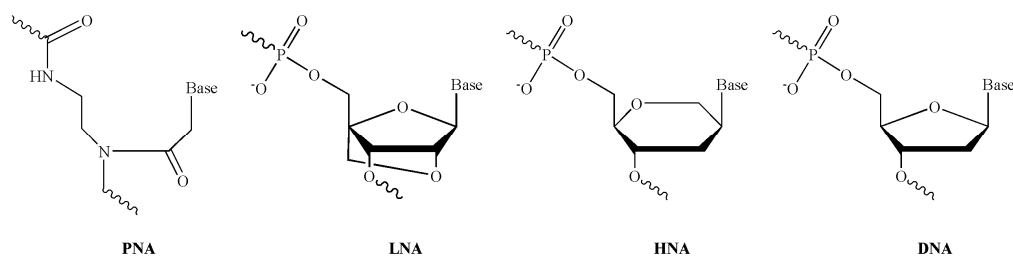
Figure 1.3. The chemical structures of peptide nucleic acid (PNA), locked nucleic acid (LNA), anhydrohexitol nucleic acid (HNA), and deoxyribonucleic acid (DNA).

unnatural nucleic acid strands (peptide nucleic acid (PNA) (Egholm et al. 1993)

and locked nucleic acid (LNA) (Obika et al. 1997; Singh et al. 1998),

respectively) (Figure 1.3). They demonstrated that this system could work as a

general analytical tool to estimate the helicity of unnatural nucleic acids.

    As part of my Ph.D. thesis, I recently described the synthesis of a

nanostructure composed entirely of glycerol nucleic acid (GNA, Figure 1.4)
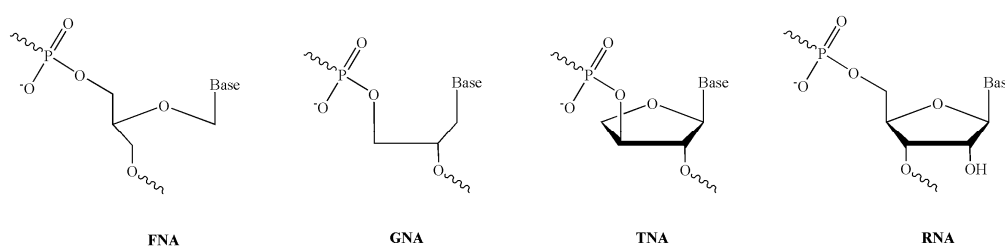
Figure 1.4. The chemical structures of flexible nucleic acid (FNA), glycerol
nucleic acid (GNA), threose nucleic acid (TNA), and ribonucleic acid (RNA).

(Zhang and Meggers 2005). The GNA nanostructure, a 4-helix junction (4HJ)

motif (Zhang et al. 2008) mimics an earlier structure composed of DNA

(Kallenbach et al. 1983). Because the GNA backbone contains only one

stereocenter per repeating unit, it was possible to synthesize two mirror-image

nanostructures using (*S*)- and (*R*)-GNA. A major finding was that the GNA

4HJ was significantly more stable than the earlier 4HJ composed entirely of

DNA (T_m 73°C for GNA versus 37°C for DNA). This feature coupled with the

ability to construct left- and right-handed nanostructures provides new

opportunities for building highly stable nanostructures with unique topologies

that are not readily available to DNA.

Unnatural nucleic acids have expanded the repertoire of building blocks for DNA nanotechnology and may help introduce novel structural, chemical and physical properties into DNA nanostructures assembled by hierarchical methods. However, these potential advantages are offset somewhat by uncertainties in the construction of more complicated nanostructures (Lin et al. 2009). L-DNA (Urata et al. 1991). For example, the enantiomer of the natural D-DNA, could have provided a possible solution to these drawbacks as L-DNA and D-DNA are perfect mirror-images to each other and possess the same duplex conformation but opposite chirality (Urata et al. 1992). Lin et al. in their pioneer work (Lin et al. 2009) described one-strand nanotubes and 2-D nanoarrays self-assembled from L-DNA by the same self-assembly designs and protocol as previously applied to natural D-DNA. Particularly, atomic force microscope (AFM) imaging confirmed that L- and D-DNA nanotubes exhibit opposite chirality. L-DNA nanotubes remained intact after DNA exonuclease digestion while D-DNA nanostructures were degraded. Taking this further, Seeman and co-workers (Ciengshih et al. 2011) demonstrated the formation of a prototype molecular weave by using a combination of D- and L-DNA in the same strand. In this braided nanostructure with alternating positive and negative nodes, L-nucleotides were used to build positive (left-handed) nodes instead of natural Z-DNA (Rich et al. 1984), which was used in the past and is not exactly a mirror image of right-handed B-DNA (Jovin and Soumpasis 1983), and D-nucleotides were used to build negative

(right-handed) nodes. This work provided a more precise way to control the

topology of DNA nanostructures. Based on the results it is possible to build

macroscopically woven materials on the molecular scale.

1.2 Templated Synthesis of alternative nucleic acid analogs

Templated synthesis of non-natural nucleic acids has been intensively

studied, with and without the mediation of enzymes. The purposes of the study

could be categorized into several groups. One of such studies has involved

alternative nucleic acids of evolutionary importance, which aims to understand

chemical routes by which information could be copied and transferred under

primordial world. These alternative nucleic acids include flexible nucleic acid

(FNA) (Joyce et al. 1987; Schneider and Benner 1990), glycerol nucleic acid

(GNA) (Zhang et al. 2005), peptide nucleic acid (PNA) (Egholm et al. 1993),

and threose nucleic acid (TNA) (Schöning et al. 2000) (Figure 1.3; Figure

1.4).

1.2.1 Flexible Nucleic Acid (FNA)

One remarkable challenge facing the RNA world hypothesis (Woese 1967;

Crick 1968; Orgel 1968; Gilbert 1986; Joyce 2002) has been the prebiotic

ribonucleotide synthesis and its assembly into short RNA oligonucleotides

thereafter. Considering the difficulty, several alternative information-carrying

molecules that are more accessible under primordial conditions have been

proposed to precede RNA and play a transitional role eventually leading to the

7

emergence of the RNA world. One of such "RNA progenitors" is FNA (Figure 1.4). FNA monomer with a prochiral carbon becomes chiral once it is incorporated into oligonucleotides and adopts either a D-like or L-like orientation in the chain. However, the sugar backbone might be flexible enough to render the templated synthesis taking place in a manner less discriminated by the absolute configuration of the nucleotides.

The earliest examples of template copying reactions involved the oligomerization of activated FNA *diphosphate* analogs on conventional polynucleotide templates (Schwartz and Orgel 1985). In these cases, successful oligomerization resulted in the formation of a product strand composed of pyrophosphate linkages rather than the more common phosphodiester linkages found in natural DNA and RNA. Activated FNA-G monomers were shown to rapidly oligomerize on poly-C templates. Control reactions performed in the absence of the template yielded only small amounts of the elongated product. Template-directed oligomerization efficiency of the FNA-G diphosphate analog was found to be lower than that of the dG diphosphate analog. Nevertheless, the observed results suggest that atactic products derived from prochiral FNA-G diphosphate monomers are successfully extended during template-directed synthesis. The latter finding is important because it demonstrates that this system avoids the problem of enantiomeric cross-inhibition seen with RNA monomers (Joyce et al. 1984).

8

In contrast to the above studies, Chaput and Switzer examined a series of atactic templates with FNA-C bearing a natural phosphodiester backbone to determine if these molecules might direct the synthesis of activated GMP monomers (Chaput and Switzer 2000). Analysis of the template-copying reactions indicated that atactic FNA-C oligonucleotides support template-directed synthesis, albeit with reduced efficiency when compared to dC templates. A key finding was that both single and multiple FNA-C residues in a template are capable of directing the synthesis of the complementary strand. RNase digestion showed that regioselective control of 3',5'-phosphodiester bond formation in the product strand was gradually lost as more FNA-C residues were introduced into the template. The main conclusion was that atactic FNA templates with natural phosphodiester linkages support template-directed synthesis, and although the template efficiency was less than with natural DNA, it may be presumed to be much higher than what would be obtained from a template bearing a mixture of all possible ribonucleoside isomers.

Despite the inability of FNA to form stable duplex structures with DNA (Schneider and Benner 1990), Heuberger and Switzer (2008) recently demonstrated that both (*S*)- and (*R*)-fNTP antipodes are substrates for several DNA polymerases. Several polymerases accepted racemic, *(R)*- and *(S)*-fNTPs as substrates in primer extension reactions, leading to multiple incorporations of flexible nucleotides. The overall results demonstrated that, on average,

antipodal fNTPs are not distinguished. Mismatched fNTP-template

combinations led to only very limited primer extension. These results show

that functional molecular recognition properties, including Watson-Crick base

pairing, are retained by the structurally simple FNA system even though it has

additional degrees of freedom in its backbone relative to DNA. The above

findings with FNA have implications for the pre-RNA world and molecular

evolution. These polymerase studies highlight the observation that acyclic

structures, in general, and FNA, in particular, might be a means to preserving

the informational properties of a polymer, while avoiding enantiomeric

cross-inhibition (Joyce et al. 1984) during replication.

1.2.2 Threose Nucleic Acid (TNA)

    To understand the chemical etiology of nucleic acid structure,

Eschenmoser (1999) and co-workers (Schöning et al. 2000) systematically

investigated the structural alternatives of RNA and discovered a set of

unnatural nucleic acids that are capable of Watson-Crick base-pairing. One of

the intriguing analogs, (3'→2')-α-L-threofuranosyl nucleic acid (TNA, Figure

1.4) shows similar thermal stability to RNA and DNA in the context of duplex

forming. TNA also forms cross-pairs with either RNA or DNA (Schöning et al.

2000). Structural studies have shown that TNA is a better mimic of A-form

DNA/RNA (Pallan et al. 2003; Ebert et al. 2008), although it could be readily

accommodated into B-form DNA (Wilds et al. 2002). Non-enzymatic

polymerization of activated rGMP against a cytosine TNA template has been

conducted to give short RNA products (Heuberger and Switzer 2006). All the

evidence implicated that TNA might have served a role in the prebiotic world

(Orgel 2000). To understand the functional potential of the TNA sequences,

Szostak and co-workers investigated the templating properties of TNA

(Chaput, Ichida, and Szostak 2003) and the kinetics of incorporating TNA

triphosphates by a variety of DNA polymerases and reverse transcriptases

(Chaput and Szostak 2003; Horhota et al. 2005; Zou et al. 2005). Preliminary

data of a TNA-directed DNA primer-extension assay showed that most

enzymes are capable of copying 1-3 DNA nucleotides on a TNA template. The

primer extension efficiency was dramatically increased in the presence of

Mn^{2+}, which is known to relax the specificity of many DNA polymerases. In

this particular case, Mn^{2+} is believed to partially compromise the progressive

loss of contact between TNA template and the polymerase as the primer is

extended. The error-rate of DNA synthesis was compared with TNA template

and DNA template in order to assess the Mn^{2+} effect on fidelity. Sequencing

result of the full-length DNA products from the primer-extension assay

showed that the error-rate for TNA was 1.5-fold and 3-fold higher than for

DNA, using Sequenase and MMLV-RT, respectively. The authors suggested

that evolution of an improved TNA-directed DNA polymerase would enable

the in vitro selection of TNA sequences (Chaput, Ichida, and Szostak 2003).

In order to perform in vitro TNA selections, it is necessary to synthesize

TNA molecules using DNA templates. Szostak and co-workers (Horhota et al.

11

2005) demonstrated that Deep Vent (exo-) DNA polymerase is capable of

incorporating three TNA nucleotides over a DNA template. Further screening

a number of other thermophilic DNA polymerases revealed Therminator DNA

polymerase (Gardner and Jack 2002), an engineered exonuclease-deficient

form of 9°N DNA polymerase containing A485L mutation as a particularly

efficient DNA-dependent TNA polymerase. Under optimal conditions, the

primer extension using Therminator led to sequences containing 80

consecutive TNA residues as the transcription product (Ichida, Zou, et al.

2005). Error rates were raised compared to the use of natural dNTPs, but low

enough for use in in vitro selection. Based on these results, the authors

demonstrated a conceptual system for in vitro evolution of TNA aptamers and

enzymes (Ichida, Zou, et al. 2005; Ichida, Horhota, et al. 2005).

1.2.3 Glycerol Nucleic Acid (GNA)

Inspired by the discovery of TNA, Meggers and co-workers (Zhang et al.

2005) synthesized a simple nucleic acid composed of a glycerol sugar

connected by phosphodiester linkages (glycerol nucleic acid, GNA, Figure

1.4). The repeating backbone unit contains one stereogenic center, which

results in two enantiomers, namely (*S*)- and (*R*)-GNA. Both (*S*)- and (*R*)-GNA

form antiparallel duplex structures with complementary GNA strands of the

same stereoconfiguration. (*S*)-, but not (*R*)-GNA also forms duplex with RNA

in an antiparallel orientation (Meggers and Zhang 2010). Szostak and

co-workers recently examined the ability of GNA triphosphates (gNTPs) to

function as substrates for natural DNA polymerases (Hornota et al. 2006).

Using a standard primer-template assay, many DNA polymerases were

screened for the ability to extend a DNA primer with GNA. While most DNA

polymerases readily incorporated one GNA residue, further primer extension

beyond a single nucleotide was not observed. The difficulty in synthesizing

longer GNA strands may be due to an incompatibility between gNTPs and

natural polymerases. The situation was very different when the enzymatic

synthesis of DNA was examined on GNA templates (Tsai et al. 2007). In this

case, a variety of DNA polymerases showed GNA-dependent DNA

polymerase activity, with Bst DNA polymerase capable of synthesizing the

full-length DNA product on a GNA template. The authors noted that in

addition to the presence of manganese ions in the reaction buffer, the

substitution of adenosine for diaminopurine in the template and DNA

monomers resulted in higher reaction efficiency. This example demonstrates

that a stable duplex structure between the product and template strands is not

required for enzyme-mediated polymerization, consistent with the FNA results

noted above.

1.2.4 Peptide nucleic acid (PNA)

PNA constitutes a remarkably new class of synthetic analogs of nucleic

acid. The purine and pyrimidine bases are connected by repetitive units of

N-(2-aminoethyl) glycine (Figure 1.3) instead of the ordinary phosphodiester

linkages. PNA backbone is not charged; therefore it undergoes hybridization

13

with RNA, DNA or itself with a high stability (Nielsen 1995). Besides, the constituent components of PNA have been isolated under putative prebiotic conditions, which suggested that PNA polymerizations might have relevance to primordial information storage and transfer.

Orgel and coworkers demonstrated that activated-PNA dimer G_2 could be condensed over a complementary DNA dC_{10} template (Böhler et al. 1995) while the same coupling on a PNA C_{10} template was inhibited. The sequence specificity of polymerization for PNA dimers (A_2, C_2, G_2, and T_2) was also examined on a decamer DNA template sequence 5'-CCCCXXCCCC-3', where X is any of the four nucleotides. The results showed that the building blocks complementary to X are preferentially incorporated with considerable misincorporations and truncations (Schmidt et al. 1997b).

A poly-cytosine PNA template was found to facilitate the polymerization of activated RNA guanosine nucleotides. The reaction is less regioselective for the native 3'-5' phosphodiester bond formation than the polymerization using RNA templates (Böhler et al. 1995). Further studies showed that PNA templates could direct the polymerization of different nucleotides in a sequence-specific manner (Schmidt et al. 1997a).

Inspired by the reductive amination of DNA and amido DNA (Goodwin and Lynn 1992), Liu and coworkers creatively developed the sequence-specific DNA-templated polymerization of tetrameric PNA aldehydes (Rosenbaum and Liu 2003). The polymerization exhibited high

efficiency on DNA templates containing as many as 40 nucleotides and high

fidelity even in the presence of a mixture of different PNA building blocks

closely related in sequence. The authors suggested that these findings are a

key step toward the evolution of sequence-defined synthetic polymers other

than natural biopolymers, using iterative in vitro selections.

Another category of templated-synthesis of modified nucleic acids aims at

identifying DNA/RNA sequences with novel functions beyond that of native

DNA and RNA using in vitro selection (Keefe and Cload 2008; Knudsen et al.

2002). The majority of data reported has focused on the selection using

ribonucleotides modified at 2'-position, which increases the stability of DNA

and RNA to chemical and nuclease-mediated degradation. To be useful in the

selection, modified nucleotides must serve as substrates for polymerases in the

enzymatic steps (Loakes and Holliger 2009), which include direct PCR

amplification using a DNA polymerase and reverse transcription of the

modified oligonucleotide library into DNA followed by transcription into the

appropriately modified oligonucleotide library (Figure 1.5). Remarkable

efforts have been made to discover polymerases or their variants that are

amenable to these processes. This work led to the identification of several

mutants of T7 RNA polymerases able to accept 2'-modifications reasonably

well (Keefe and Cload 2008). Modifications at 5-position of T/U base are also

tolerated by a variety of T7 RNA polymerases.
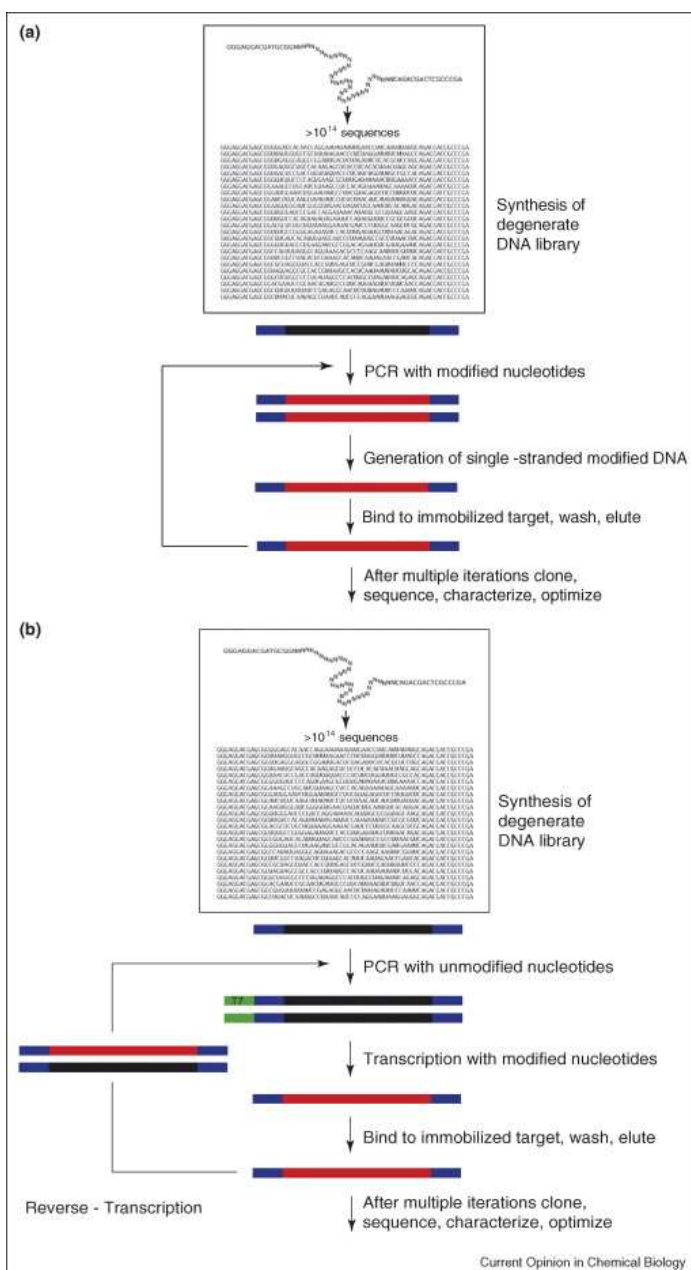
Figure 1.5. (a) Schematic representation of the Modified SELEX process in which a DNA polymerase is used to incorporate modified nucleotides into the oligonucleotide library. Black indicates the degenerate region of the unmodified oligonucleotide library, red indicates the degenerate region of the modified oligonucleotide library, blue indicates the fixed primer-binding region. (b) Schematic representation of the modified SELEX process in which a DNA polymerase is used to incorporate modified nucleotides into the oligonucleotide library. Black indicates the degenerate region of the unmodified oligonucleotide library, red indicates the degenerate region of the modified oligonucleotide library, blue indicates the fixed primer-binding region, and green indicates the transcription promoter sequence. Reprinted (adapted in part or whole) with permission from Curr. Opin. Chem. Biol. (Keefe and Cload 2008). Copyright 2008 Elsevier.

16

Besides increasing the stability of aptamers, modified nucleotides has been proved to have more advantages in the selection against challenging targets than native DNA/RNA in some cases. Using dUTP derivatives, Eaton and co-workers (Vaught et al. 2010) successfully identified aptamers to a protein target TNFRSF9 (Croft 2009) that had previously been found refractory to in vitro selection using DNA. Other chemical modifications of the nucleotides employed in in vitro selection include boranophosphate and phosphorothioate internucleotide linkages, and 4'-thio pyrimidines, all of which have shown enhanced nuclease-resistance (Keefe and Cload 2008). Chemical structures of the modified nucleotides used in in vitro selection are summarized in Figure 1.6.
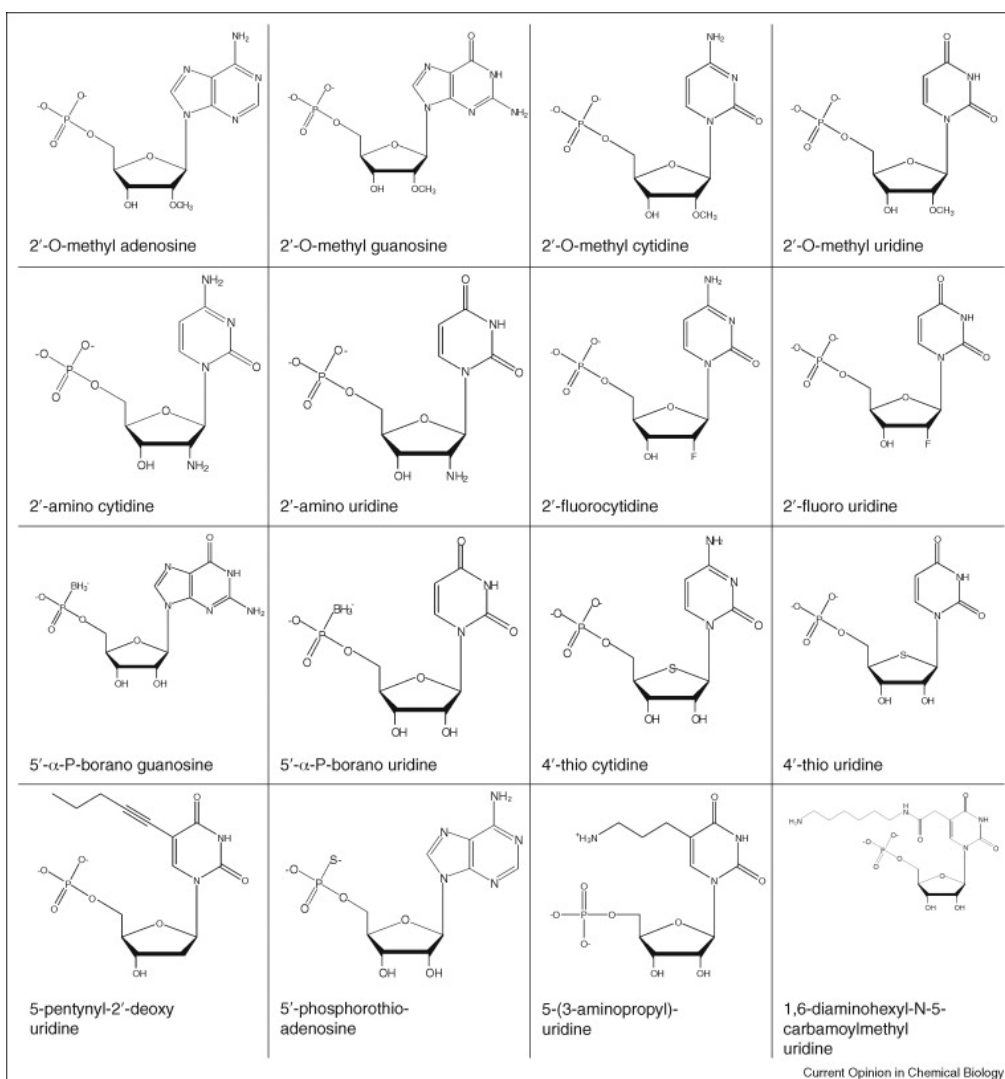
Figure 1.6. Chemical structures of the modified nucleotides which have been reported in the process of SELEX. Reprinted (adapted in part or whole) with permission from Curr. Opin. Chem. Biol. (Keefe and Cload 2008). Copyright 2008 Elsevier.

Other alternative nucleic acids that have been used in Aptamer sequences include Locked nucleic acid (LNA, Figure 1.3) (Obika et al. 1997; Singh et al. 1998) and anhydrohexitol nucleic acid (HNA, Figure 1.3) (Hendrix et al. 1997). Short pieces of LNA/HNA strands have been incorporated into aptamers and/or ribozymes to enhance nuclease resistance and improve

activity (Kang et al. 2004; Lebars et al 2007; Kolb et al. 2005), though they are not directly involved in the selection process.

1.2.5 Locked Nucleic Acid (LNA)

LNA was first described by Imanishi (1997) and Wengel (1998). The methylene bridge between 2'-O and C4' of the ribose moiety locks the sugar pucker into a 3'-endo conformation. LNA binds tightly to complementary DNA and RNA and induces A-form like duplexes (Nielsen et al. 2004; Vester and Wengel 2004). LNA residues could be bypassed within a DNA template by DNA polymerases if the LNA nucleotides are distributed along the strand. LNA triphosphates could be recognized by several DNA polymerases as substrates (Kuwahara et al. 2008). The most efficient KOD DNA polymerase has incorporated up to 8 consecutive LNA residues.

1.2.6 Hexitol Nucleic Acid (HNA)

Among a set of six-membered carbohydrate mimics of RNA developed by Herdewijn and co-workers (Kerremans et al. 2001), anhydrohexitol nucleic acid (HNA) forms the most stable cross-pair with RNA (even more stable than RNA self-pairing). Structural studies showed that both HNA-HNA and HNA-RNA duplexes adopt A-form helices (Declercq et al. 2002; Maier et al. 2005). While several polymerases were able to incorporate a single HNA triphosphate, only Deep Vent (exo-) DNA polymerase could incorporate up to 6 consecutive HNA residues at high enzyme concentration (Vastmans et al. 2000). Under such conditions the fidelity decreased after the first two HNA

triphosphates were inserted. HNA oligonucleotides have been used in

aptamers (Kolb et al. 2005) and antisense (Kang et al. 2004) with nuclease

resistance. However, pure HNA antisense was shown ineffective possibly due

to the lack of RNase H activity (Kang et al. 2004).

1.3 Artificial Information-Carrying Polymers Capable of Darwinian Evolution

Historically in vitro evolution has been limited to two types of biological

polymers, namely nucleic acids and proteins, due to its heavy dependency on

the use of biological machinery. Protein engineering dedicated to creating

ribosome and polymerases tolerant to unnatural susbstrates has been proved

useful (Wang et al. 2006). Some successful examples of modified aptamer

selections were discussed in previous section. However, in most of the cases,

only modest structural modifications have been proved suitable. It is hard to

develop a general strategy for the directed evolution of a completely synthetic

polymer system (Appella 2009; Brudno and Liu 2009).

One approach inspired by biological evolution (Figure 1.7A) generates the

synthetic polymers covalently linked to the encoding nucleic acid templates.

The resulting libraries of the synthetic polymers are subjected to selection and

the encoding template nucleic acids of the surviving molecules are amplified

for a subsequent round of selection (Figure 1.7B). The templated synthesis of

synthetic polymers could be either nonenzymatic or enzymatic.
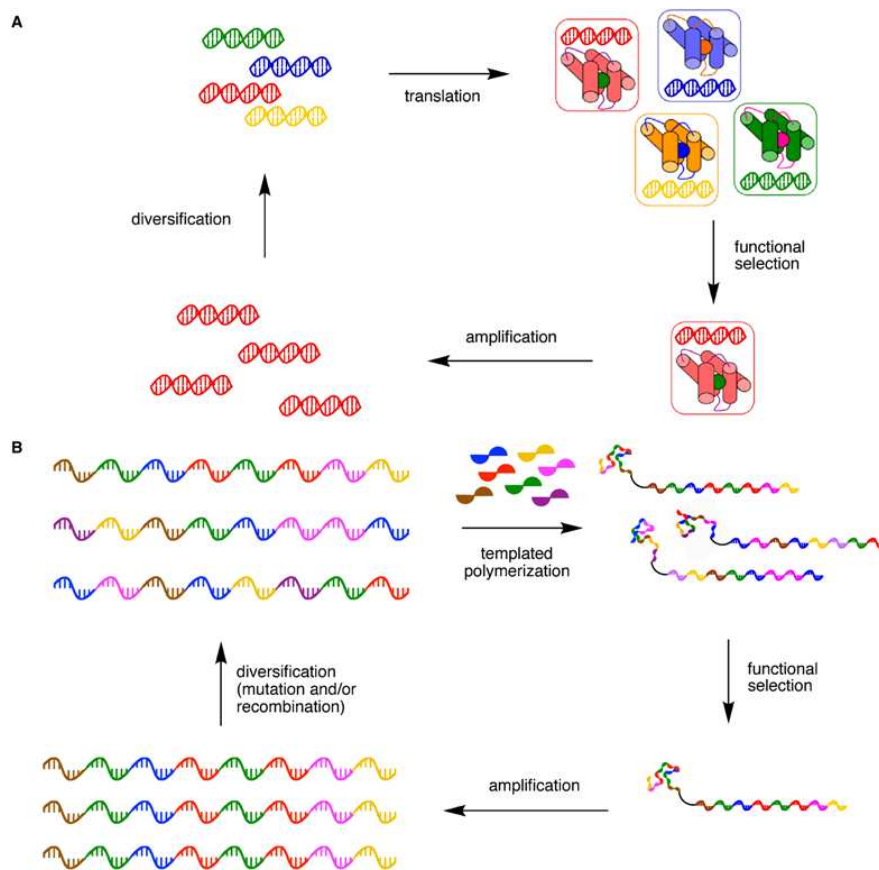
Figure 1.7. Key Components of the Evolution of Biological or Synthetic Polymers. (A) Biopolymer evolution as it occurs in nature or in the laboratory. A polymer (such as a protein) is translated from and spatially associated with an information carrier (such as DNA). The resulting biopolymers undergo selection based on their functional properties. The information encoding surviving biopolymers replicates and mutates, resulting in a second generation of biopolymer variants related to those that survived selection. (B) Synthetic polymers can in principle undergo a similar evolutionary process. Translation could be effected by nonenzymatic or enzymatic templated synthesis in a manner that associates each synthetic polymer with its information carrier. Following selection, the information carriers encoding surviving synthetic polymers are amplified and mutated to generate templates for a subsequent round of translation. Reprinted (adapted in part or whole) with permission from Chem. Biol. (Brudno and Liu 2009). Copyright 2009 Elsevier.

Liu and co-workers (Brudno et al. 2010) demonstrated a peptide nucleic acid (PNA) in vitro selection system, in which PNA library was generated by reductive amination-based polymerization. Next, the PNA sequences need to be liberated from the DNA templates to which they are bound so that the PNA

strands could fold and be subjected to selection. This is a challenging problem as PNA binds tightly to complementary DNA. After extensive screening, the Herculase II DNA polymerase was found to efficiently displace a 40-nucleotide PNA strand from the DNA template. The resulting encoding template DNA could be successfully isolated and amplified using PCR even after a 15000-fold dilution. This strategy was employed in a mock selection to enrich a unique DNA sequence encoding biotinylated PNA pentamer, which was highly underrepresented in the starting library ($1:10^6$). After six rounds of selection the results showed that primary sequences identified were the unique encoding sequence.

As part of my PhD. dissertation, I developed the methodology necessary to evolve functional TNA molecules by in vitro selection. Working in collaboration with another student in the lab, I identified a TNA aptamer binding human α-thrombin with a Kd around 200 nM after three iterative rounds of selection using this library, which will be discussed in Chapter 6 in this dissertation. Other examples of evolution-based technologies for synthetic polymer selection, including DNA encoded library of macrocycles (Gartner et al. 2004) and YoctoReactor technology (Hansen et al. 2009), are beyond the scope of this review and have been discussed elsewhere (Li and Liu 2004).

1.4 References

Appella, D.H. 2009. Non-natural nucleic acids for synthetic biology. *Curr. Opin. Chem. Biol.* 13: 687-696.

Bartlett, J.M.S., and D. Stirling. 2003. A short history of the polymerase chain reaction. *Methods Mol. Biol.* 226: 3-6.

Böhler, C., P.E. Nielsen, and L.E. Orgel. 1995. Templated switching between PNA and RNA oligonucleotides. *Nature* 376: 578-581.

Brudno, Y., M. Birnbaum, R.E. Kleiner, and D.R. Liu. 2010. An in vitro translation, selection and amplification system for peptide nucleic acids. *Nat. Chem. Biol.* 6: 148-155.

Brudno, Y., and D.R. Liu. 2009. Recent progress towards the templated synthesis and directed evolution of sequence-defined synthetic polymers. *Chem. Biol.* 16: 265-276.

Chaput, J.C., J.K. Ichida, and J.W. Szostak. 2003a. DNA polymerase-mediated DNA synthesis on a TNA template. *J. Am. Chem. Soc.* 125: 856-857.

Chaput, J.C., and C. Switzer. 2000. Nonenzymatic oligomerization on templates containing phosphodiester-linked acyclic glycerol nucleic acid analogues. *J. Mol. Evol.* 51: 464-470.

Chaput, J.C., and J.W. Szostak. 2003b. TNA synthesis by DNA polymerases. *J. Am. Chem. Soc.* 125: 9274-9275.

Ciengshih, T., R. Sha, and N.C. Seeman. 2011. Automatic molecular weaving prototyped by using single-stranded DNA. *Angew. Chem. Int. Ed.* 50: 4419-4422.

Crick, F.H.C. 1968. The origin of the genetic code. *J. Mol. Biol.* 38: 367-369.

Croft, M. 2009. The role of TNF superfamily members in T-cell function and diseases. *Nat. Rev. Immunol.* 9: 271-285.

Declercq, R., A. Van Aerschot, R.J. Read, P. Herdewijn, and L. Van Meervelt. 2002. Crystal structure of double helical hexitol nucleic acids. *J. Am. Chem. Soc.* 124: 928-933.

Ebert, M.-O., C. Mang, R. Krishnamurthy, A. Eschenmoser, and B. Jaun. 2008. The structure of a TNA-TNA complex in solution: NMR study of the octamer

duplex derived from α-(L)-threofuranosyl-(3'-2')-CGAATTCG. *J. Am. Chem. Soc.* 130: 15105-15115.

Egholm, M., O. Buchardt, L. Christensen, C. Behrens, S.M. Freier, D.A. Driver, R.H. Berg, S.K. Kim, P.E. Norden, and P.E. Nielsen. 1993. PNA hybridizes to complementary oligonucleotides obeying the Watson-Crick hydrogen-bonding rules. *Nature* 365: 566-568.

Eguchi, Y., T. Itoh, and J. Tomizawa. 1991. Antisense RNA. *Annu. Rev. Biochem.* 60: 631-652.

Eschenmose, A. 1999. Chemical etiology of nucleic acid structure. *Science* 284: 2118-2124.

Famulok, M. 2009. Exploring chemical space with aptamers. *J. Med. Chem.* 52: 6951-6957.

Gardner, A.F., and W.E. Jack. 2002. Acyclic and dideoxy terminator preferences denote divergent sugar recognition by archaeon and *Taq* DNA polymerases. *Nucleic Acid Res.* 30: 605-613.

Gartner, Z.J., B.N. Tse, R. Grubina, J.B. Doyon, T.M. Snyder, and D.R. Liu. 2004. DNA-templated organic synthesis and selection of a library of macrocycles. *Science* 305: 1601-1605.

Gilbert, W. 1986. The RNA world. *Nature* 319: 618.

Goodwin, J.T., and D.G. Lynn. 1992. Template-directed synthesis: use of a reversible reaction. *J. Am. Chem. Soc.* 114: 9197-9198.

Hansen, M.H., P. Blakskjær, L.K. Peterson, T.H. Hansen, J.W. Højfeldt, K.V. Gothelf, and N.J.V. Hansen. 2009. A yoctoliter-scale DNA reactor for small-molecule evolution. *J. Am. Chem. Soc.* 131: 1322-1327.

Hendrix, C., H. Rosemeyer, I. Verheggen, F. Seela, A. Van Aerschot, and P. Herdewijn. 1997. 1',5'-Anhydrohexitol oligonucleotides: synthesis, base pairing and recognition by regular oligodeoxyribonucleotides and oligoribonucleotides. Chem. Eur. J. 3: 110-120.

Heuberger, B.D., and C. Switzer. 2006. Nonenzymatic oligomerization of RNA by TNA templates. *Org. Lett.* 8: 5809-5811.

Heuberger, B.D., and C. Switzer. 2008. A pre-RNA candidate revisited: both enantiomers of flexible nucleoside triphosphates are DNA polymerase substrates. *J. Am. Chem. Soc.* 130: 412-413.

Hornota, A.T., J.W. Szostak, and L.W. McLaughlin. 2006. Glycerol nucleoside triphosphates: synthesis and polymerase substrate activities. *Org. Lett.* 8: 5345-5347.

Horhota, A., K. Zou, J.K. Ichida, B. Yu, L.W. McLaughlin, J.W. Szostak, and J.C. Chaput. 2005. Kinetics analysis of an efficient DNA-dependent TNA polymerase. *J. Am. Chem. Soc.* 127: 7427-7434.

Ichida, J.K., A. Horhota, K. Zou, L.W. McLaughlin, and J.W. Szostak. 2005a. High fidelity TNA synthesis by Therminator polymerase. *Nucleic Acids Res.* 33: 5219-5225.

Ichida, J.K., K. Zou, A. Horhota, B. Yu, L.W. McLaughlin, and J.W. Szostak. 2005b. An in vitro selection system for TNA. *J. Am. Chem. Soc.* 127: 2802-2803.

Johnson, A.T., M.K. Schlegel, E. Meggers, L.-O. Essen, and O. Wiest. 2011. On the structure and dynamics of the duplex GNA. *J. Org. Chem.* 76: 7964-7974.

Jovin, T.M., and D.M. Soumpasis. 1983. The transition between B-DNA and Z-DNA. *Annu. Rev. Phys. Chem.* 38: 521-560.

Joyce, G.F. 2002. The antiquity of RNA-based evolution. *Nature* 418: 214-221.

Joyce, G.F., A.W. Schwartz, S.L. Miller, and L.E. Orgel. 1987. The case for an ancestral genetic system involving simple analogues of the nucleotides. *Proc. Natl. Acad. Sci. USA* 84: 4398-4402.

Joyce, G.F., G.M. Visser, C.A.A. van Boeckel, J.H. van Boom, L.E. Orgel, and J. van Westrenen. 1984. Chiral selection in poly(C)-directed synthesis of oligo(G). *Nature* 310: 602-604.

Kallenbach, N.R., R.-I. Ma, and N.C. Seeman. 1983. An immobile nucleic acid junction constructed from oligonucleotides. *Nature* 305: 829-831.

Kang, H., M.H. Fisher, D. Xu, Y.J. Miyamoto, A. Marchand, A. Van Aerschot, P. Herdewijn, and R.L. Juliano. 2004. Inhibition of MDR1 gene expression by chimeric HNA antisense oligonucleotides. *Nucleic Acid Res.* 32: 4411-4419.

Keefe, A.D., and S.T. Cload. 2008. SELEX with modified nucleotides. *Curr. Opin. Chem. Biol.* 12: 448-456.

Kerremans, L., G. Schepers, J. Rozenski, R. Busson, A. Van Aerschot, and P. Herdewijn. 2001. Hybridization between "six-membered" nucleic acids: RNA as a universal information system. *Org. Lett.* 3: 4129-4132.

Khudyakov, Y.E., and H.A. Fields. eds. 2002. *Artificial DNA: methods and applications*. Boca Raton, Florida: CRC Press.

Knudsen, S.M., M.P. Robertson, and A.D. Ellington. 2002. In vitro selection using modified or unnatural nucleotides. *Curr. Protoc. Nucleic. Acid Chem.* Chapter 9: unit 9.6.

Kolb, G., S. Reigadas, C. Boiziau, A. Van Aerschot, A. Arzumanov, M.J. Gait, P. Herdewijn, and J.-J. Toulme. 2005. Hexitol nucleic acid-containing aptamers are efficient ligands of HIV-1 TAR RNA. *Biochemistry* 44: 2926-2933.

Kuwahara, M., S. Obika, J.-I. Nagashima, Y. Ohta, Y. Suto, H. Ozaki, H. Sawai, and T. Imanishi. 2008. Systematic analysis of enzymatic DNA polymerization using oligo-DNA templates and triphosphate analogs involving 2',4'-bridged nucleosides. *Nucleic Acid Res.* 36: 4257-4265.

Lebars, I., T. Richard, C. Di Primo, and J.J. Toulme. 2007. LNA derivatives of a kissing aptamer targeted to the *trans*-activating responsive RNA element of HIV-1. *Blood Cells Mol. Dis.* 38: 204-209.

Li, X., and D.R. Liu. 2004. DNA-templated organic synthesis: nature's strategy for controlling chemical reactivity applied to synthetic molecules. *Angew. Chem. Int. Ed.* 43: 4848-4870.

Lin, C., Y. Ke, Z. Li, J.H. Wang, Y. Liu, and H. Yan. 2009 Mirror image DNA nanostructures for chiral supramolecular assemblies. *Nano Lett.* 9: 433-436.

Lin, C., Y. Liu, and H. Yan. 2009. Designer DNA nanoarchitectures. *Biochemistry* 48: 1663-1674.

Loakes, D., and P. Holliger. 2009. Polymerase engineer: towards the encoded synthesis of unnatural biopolymers. *Chem. Commun.* 4619-4631.

Lukeman, P.S., A.C. Mittal, and N.C. Seeman. 2004. Two dimensional PNA/DNA arrays: estimating the helicity of unusual nucleic acid polymers. *Chem. Commun.* 1694-1695.

Maier, T., I. Przylas, N. Strater, P. Herdewijn, and W. Saenger. 2005. Reinforced HNA backbone hydration in the crystal structure of a decameric HNA/RNA hybrid. *J. Am. Chem. Soc.* 127: 2937-2934.

Megger, E., and L. Zhang. 2010. Synthesis and properties of the simplified nucleic acid GNA. *Acc. Chem. Res.* 43: 1092-1102.

Milligan, J.F., and O.C. Uhlenbeck. 1989. Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol.* 180: 51-62.

Muller, S., J. Wolf, and S.A. Ivanov. 2004. Current strategies for the synthesis of RNA. *Curr. Org. Syn.* 1: 293-307.

Nahvi, A., N. Sudarsan, M.S. Ebert, X. Zou, K.L. Brown, and R.R. Breaker. 2002. Genetic control by a metabolite binding mRNA. *Chem. Biol.* 9: 1043-1049.

Nangreave, J., Y. Liu, and H. Yan. 2011. DNA nanostructures as models for evaluating the role of enthalpy and entropy in polyvalent binding. *J. Am. Chem. Soc.* 133: 4490-4497.

Nielsen, P.E. 1995. DNA analogues with nonphosphodiester backbones. *Annu. Rev. Biophys. Biomol. Struct.* 24: 167-183.

Nielsen, K.E., J. Rasmussen, R. Kumar, J. Wengel, J.P. Jacobsen, and M. Peterson. 2004. NMR studies of fully modified locked nucleic acid (LNA) hybrids: solution structure of an LNA: RNA hybrid and characterization of an LNA: DNA hybrid. *Bioconjug. Chem.* 15: 449-457.

Obika, S., M. Dorio, D. Naubu, and T. Imanishi. 1997. Synthesis and conformation of 3'-O,4'-C-methyleneribonucleotides, novelbicyclic nucleoside analogues for 2',5'-linked oligonucleotide modification. *Chem. Commun.* 1643-1644.

Orgel, L.E. 1968. Evolution of the genetic apparatus. *J. Mol. Biol.* 38: 381-393.

Orgel, L. 2000. A simpler nucleic acid. *Science* 17: 1306-1307.

Pallan, P.S., C.J. Wilds, Z. Wawrzak, R. Krishnamurthy, A. Eschenmoser, and M. Egli. 2003. Why does TNA cross-pair more strongly with RNA than with DNA? An answer from X-ray analysis. *Angew. Chem. Int. Ed.* 42: 5893-5895.

Rich, A., A. Nordheim, and A.H. Wang. 1984. The chemistry and biology of left-handed Z-DNA. *Annu. Rev. Biochem.* 53: 791-846.

Rinker, S., Y. Liu, and H. Yan. 2006. Two dimensional LNA/DNA arrays: estimating the helicity of LNA/DNA hybrid duplex. *Chem. Commun.* 2675-2677.

Rosenbaum, D.M., and D.R. Liu. 2003. Efficient amd sequence-specific DNA-templated polymerization of peptide nucleic acid aldehydes. *J. Am. Chem. Soc.* 125: 13924-13925.

Sacca, B., R. Meyer, U. Feldkamp, H. Schroeder, and C.M. Niemeyer. 2008. High-throughput, real-time monitoring of the self-assembly of DNA nanostructures by FRET spectroscopy. *Angew. Chem. Int. Ed.* 47: 2135-2137.

Schmidt, J.G., P.E. Nielsen, and L.E. Orgel. 1997a. Information transfer from peptide nucleic acids to RNA by template-directed syntheses. *Nucleic Acid Res.* 25: 4797-4802.

Schmidt, J.G., L. Christensen, P.E. Nielsen, and L.E. Orgel. 1997b. Information transfer from DNA to peptide nucleic acids by template-directed syntheses. *Nucleic Acid Res.* 25: 4792-4796.

Schneider, K.C., and S.A. Benner. 1990. Oligonucleotides containing flexible nucleoside analogs. *J. Am. Chem. Soc.* 112: 453-455.

Schöning, K.-U., P. Scholz, S. Guntha, X. Wu, R. Krishnamurthy, and A. Eschenmoser. 2000. Chemical etiology of nucleic acid structure: the α-threofuranosyl-(3'→2') oligonucleotide system. *Science* 290: 1347-1351.

Schwartz, A.W., and L.E. Orgel. 1985. Template-directed synthesis of novel, nucleic acid like structures. *Science*. 228: 585-587

Singh, S.K., P. Nielson, A.A. Koshkin, and J. Wengel. 1998. LNA (locked nucleic acid): synthesis and high-affinity nucleic acid recognition *Chem. Commun.* 455-456.

Tsai, C.H., J. Chen, and J.W. Szostak. 2007. Enzymatic synthesis of DNA on glycerol nucleic acid templates without stable duplex formation between product and template. *Proc. Natl. Acad. Sci. USA* 104: 14598-14603.

Urata, H., E. Ogura, K. Shinohara, Y. Ueda, and M. Akagi. 1992. Synthesis and properties of mirror-image DNA. *Nucleic Acid Res.* 20: 3325-3332.

Urata, H., K. Shinohara, E. Ogura, Y. Ueda, and M. Akagi. 1991. Mirror-image DNA. *J. Am. Chem. Soc.* 113: 8174-8175.

Vastmans, K., S. Pochet, A. Peys, L. Kerremans, A. Van Aerschot, C. Hendrix, P. Marliere, and P. Herdewijn. 2000. Enzymatic incorporation in DNA of 1,5-anhydrohexitol nucleotides. *Biochemistry* 39: 12757-12765.

Vaught, J.D., C. Bock, J. Carter, T. Fitzwater, M. Otis, D. Schneider, J. Rolando, S. Waugh, S.K. Wilcox, and B.E. Eaton. 2010. Expanding the chemistry of DNA for in vitro selection. *J. Am. Chem. Soc.* 132: 4141-4151.

Vester, B., and J. Wengel. 2004. LNA (locked nucleic acid): high-affinity targeting of complementary RNA and DNA. *Biochemistry* 43: 13233-13241.

Wagenknecht, H.-A. 2008. Fluorescent DNA base modifications and substitutes: multiple fluorophore labeling and the DETEQ concept. *Ann. N.Y. Acad. Sci.* 1130: 122-130.

Wang, L., J. Xie, and P.G. Schultz. 2006. Expanding the genetic code. *Annu. Rev. Biophys. Biomol. Struct.* 35: 225-249.

Watson, J.D., and F.H.C. Crick. 1953. Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature* 171: 737-738.

Wilds, C.J., Z. Warwrzak, R. Krishnamurthy, A. Eschenmoser, and M. Egli. 2002. Crystal structure of a B-form DNA duplex containing (L)-α-threofuranosyl (3' →2') nucleosides: a four-carbon sugar is easily accommodated into the backbone of DNA. *J. Am. Chem. Soc.* 124: 13716-13721.

Wilson, D.S., and J.W. Szostak. 1999. In vitro selection of functional nucleic acids. *Annu. Rev. Biochem.* 68: 611-647.

Woese, C.R. 1967. *The genetic code: The molecular basis for genetic expression.* p.179. New York: Harper & Row.

Yan, H. 2004. Nucleic acid nanotechnology. *Science* 17: 2048-2049.

Zhang, L., A.E. Peritz, P.J. Carroll, and E. Meggers. 2006. Synthesis of glycol nucleic acids. *Synthesis* 4: 645-653.

Zhang, L., A. Peritz, and E. Meggers. 2005. A simple glycol nucleic acid. *J. Am. Chem. Soc.* 127: 4174-4175.

Zhang, R.S., E.O. McCullum, and J.C. Chaput. 2008. Synthesis of two mirror image 4-helix junctions derived from glycerol nucleic acid. *J. Am. Chem. Soc.* 130: 5846-5847.

Zou, K., A. Horhota, B. Yu, J.W. Szostak, and L.W. McLaughlin. 2005. Synthesis of α–L-threofuranosyl nucleoside triphosphates (tNTPs). *Org. Lett.* 7: 1485-1487.

Chapter 2

SYNTHESIS OF GLYCEROL NUCLEIC ACID (GNA)

PHOSPHORAMIDITE MONOMERS

AND OLIGONUCLEOTIDE POLYMERS

2.1 Contributions

The following chapter describes a published protocol for the synthesis of GNA phosphoramidite monomers and oligonucleotide polymers. All reported experiments were performed by Su Zhang under the guidance of Professor John Chaput. Su Zhang and John Chaput wrote the manuscript. The results from this work were published in *Current Protocols in Nucleic Acid Chemistry.*

2.2 Abstract

This unit describes a straightforward method for preparing glycerol nucleic acid (GNA) phosphoramidite monomers and oligonucleotide polymers using standard cyanoethyl phosphoramidite chemistry. GNA is an unnatural nucleic acid analogue composed of an acyclic three-carbon sugar-phosphate backbone that contains one stereogenic center per repeating unit. GNA has attracted significant attention as a nucleic acid derivative due its unique ability to form stable Watson-Crick antiparallel duplex structures with thermal and thermodynamic stabilities rivaling those of natural DNA and RNA. The chemical simplicity of this nucleic acid structure provides access to

enantiomerically pure forms of right- and left-handed helical structures that

can be used as unnatural building blocks in DNA nanotechnology.

2.3 Introduction

Acyclic oliogonucleotides are experiencing a tremendous resurgence in

basic and applied research due to their unique structural and biophysical

properties (for review see Zhang et al., 2010). This unit contains procedures

that describe the chemical synthesis of one type of acyclic nucleic acid

polymer commonly referred as glycerol nucleic acid or GNA. The chemical

synthesis and purification of glycerol nucleoside analogues bearing adenine

(A), cytosine (C), guanine (G), and thymine (T) as the bases and of

oligonucleotides thereof (Figure 2.1) are described in detail. Unless otherwise
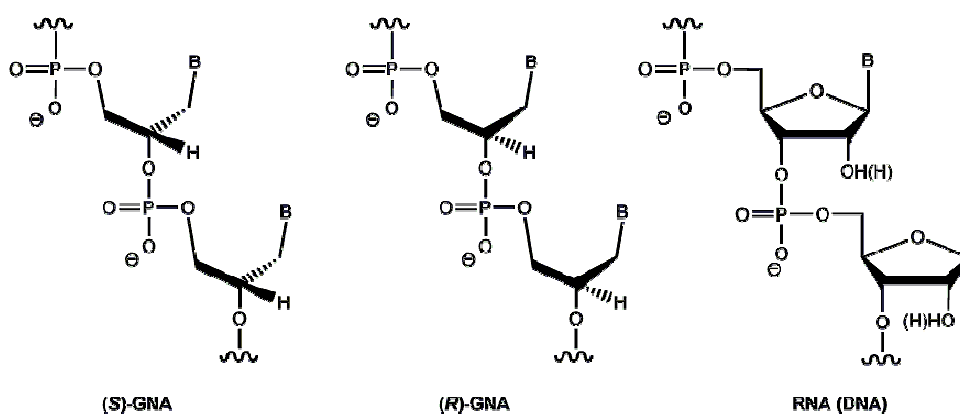
Figure 2.1. Chemical Structures of GNA, DNA and RNA. Reprinted with
permission from Current Protocols in Nucleic Acid Chemistry. Copyright 2010
John Wiley & Sons, Inc.

stated, all of the procedures start from (*R*)-(+)-glycidol and yield (*S*)-GNA.

The same chemistry can also be applied to (*S*)-(-)-glycidol to produce

(*R*)-GNA. Commercially available glycidol or tritylated glycidol is used to

31

obtain the 2,3-dihydroxypropyl derivatives of each nucleobase via an epoxide

ring opening reaction (see Basic Protocol 1). The glycerol nucleosides are then

converted to their corresponding nucleotide phosphoramidites (see Basic

Protocol 2 to 5), which can then be used as building blocks to synthesize GNA

oligonucleotides (see Basic Protocol 6).

2.4 Basic Protocol 1

Synthesis of Enantiomerically Pure dimethoxytrityl-*O*-(S)-glycidol

2,3-dihydroxypropyl derivatives of nucleobases have been synthesized by

several different methods. The current approach is based on a modified

version of Acevedo's procedure (Acevedo et al., 1996), and involves a direct

ring opening of the stable glycidol intermediate by nucleophilic attack of one

of the four natural nucleobases. In this protocol, pure (R)-(+)-glycidol is

tritylated using 4,4'-dimethoxytritylchloride (DMT-Cl) in dichloromethane as

illustrated in Figure 2.2. The tritylated glycidol is used to make the A, C, and T
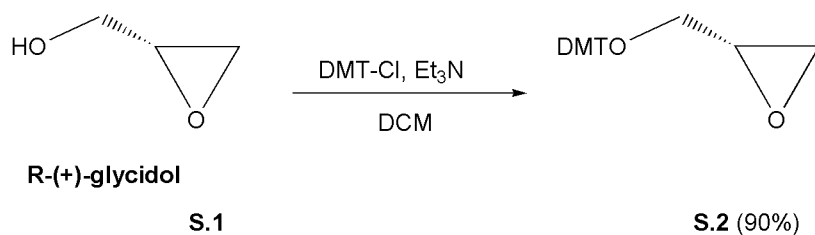
Figure 2.2. Preparation of (S)-DMT-O-glycidol (S.2). Abbreviations: DMT-Cl, 4,4'-dimethoxytritylchloride; Et$_3$N, triethylamine; DCM, dichloromethane. Reprinted with permission from Current Protocols in Nucleic Acid Chemistry. Copyright 2010 John Wiley & Sons, Inc.

glycerol nucleoside phosphoramidites (see Basic Protocols 2, 3, and 5, respectively). Synthesis of the G glycerol nucleoside phosphoramidite requires an alternative strategy (see Basic Protocol 4). Note that (R)-(+)-glycidol substrates provide (S)-GNA phosphoramidite monomers that can be used to make GNA oligonucleotides with the natural right-handed stereoconfiguration.

Materials

(R)-(+)-Glycidol

4,4'-Dimethoxytritylchloride (DMT-Cl), 95%

Triethylamine (Et_3N), 99.5%

Dichloromethane (DCM)

Hexanes

Ethyl acetate (AcOEt)

Saturated aqueous sodium bicarbonate solution (sat. aq. $NaHCO_3$)

Brine (sat. aq. NaCl)

Sodium sulfate (Na_2SO_4)

100-ml round-bottom flask

Magnetic stir plate and stir bar

Gas balloon

Rotary evaporator equipped with a vacuum pump

Thin layer chromatography (TLC) plate, EMD silica gel 60 F_{254}

2.5 x 25-cm chromatography column

33

Silica gel (60 Å, 230-400 mesh) (Whatman Inc.)

200-ml separatory funnel

Büchner funnel

Filter paper

UV lamp, 254 nm

Additional reagents and equipment for thin layer chromatography (TLC) and

column chromatography

1. To a 100-ml round bottom flask equipped with a magnetic stir bar, add 780

    mg (10.6 mmol) of (R)-(+)-glycidol, 24 ml of DCM (freshly distilled

    over calcium hydride), 3.8 ml of Et_3N and 4.54 g (13.4 mmol) of DMT-Cl.

    Stir the reaction at r. t. overnight under an argon atmosphere.

2. Remove the precipitate by vacuum filtration and wash the residue with

    DCM.

3. Wash the filtrate sequentially with sat. aq. $NaHCO_3$, water, and brine.

4. Dry the organic layer over Na_2SO_4, filter, and evaporate to dryness using a

    rotary evaporator equipped with a vacuum pump and cooling trap.

5. Purify the oily residue by column chromatography on 20 g of silica gel in a

    5 x 25-cm column. Deactivate the column with hexanes/Et_3N 97:3 and

    elute the column using a step-wise gradient of 99:1 (v/v) hexanes/Et_3N to

    hexanes/AcOEt/Et_3N 18:1:1. Combine the product fractions and evaporate

    to dryness.

The resulting product, DMT-*O*-(S)-glycidol (S.2) should be obtained in a

90% yield (3.56 g, 9.47 mmol) as colorless oil. TLC (hexanes/AcOEt 10:1): R_f

= 0.12. ^{1}H NMR (300 MHz, CDCl$_3$): δ=2.61 (dd, J=2.4, 5.1 Hz, 1 H), 2.76 (m,

1 H), 3.12 (m, 2 H), 3.29 (m, 1 H), 3.79 (s, 6 H), 6.83 (m, 4 H), 7.05-7.40 (m,

7 H), 7.45 (m, 2 H).

2.5 Basic Protocol 2

Synthesis of 2'-*O*-(2-cyanoethoxy)(diisopropylamino)phosphino-3'-*O*-

(4,4'-dimethoxytriphenyl)methyl-*N*[6]-benzoyl-(S)-9-(2,3-dihydroxypropyl)-

adenine

This protocol describes the synthesis of N[6]-benzoyl-protected (S)-GNA

adenosine nucleoside phosphoramidite S.5 from S.2 (Figure 2.3).
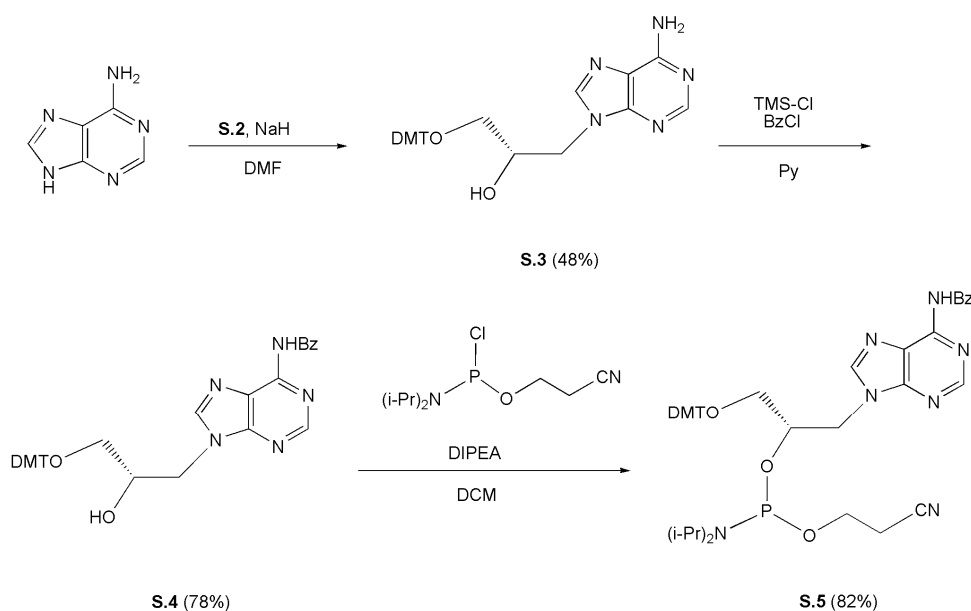
Figure 2.3. Preparation of S.5. Abbreviations: NaH, sodium hydride; DMF,
dimethylformamide; TMSCl, trimethylsilyl chloride; BzCl, benzoyl chloride;
Py, pyridine; DIPEA, diisopropylethylamine; DCM, dichloromethane.
Reprinted with permission from Current Protocols in Nucleic Acid Chemistry.
Copyright 2010 John Wiley & Sons, Inc.

Materials

DMT-*O*-(S)-glycidol (S.2)

Adenine

Sodium hydride (60% in mineral oil) (NaH)

Dimethylformamide (DMF), anhydrous

Pyridine, anhydrous

Trimethylsilyl chloride (TMS-Cl)

Benzoyl chloride (BzCl)

Ammonium hydroxide, concentrated (conc. NH_4OH)

Dichloromethane (DCM)

Diisopropylethylamine, redistilled (DIPEA)

Chloro(2-cyanoethoxy)(diisopropylamino)phosphine

Hexanes

Ethyl acetate (AcOEt)

Triethylamine (Et_3N)

Methanol (MeOH)

50- and 100-ml round bottom flasks

Oil bath

Magnetic stir plate and stir bar

Gas balloon

Graham condenser

Rotary evaporator equipped with a vacuum pump

1.3 x 30-cm chromatography column

6.4 x 45-cm chromatography column

Silica gel (60 Å, 230-400 mesh)

TLC plate, EMD silica gel 60 F$_{254}$

UV lamp, 254 nm

Büchner funnel

Filter paper

Additional reagents and equipment for TLC and column chromatography

Prepare S.3

1. Add 0.71 g (5.25 mmol) of adenine and 43 mg (1.07 mmol) of NaH (60% in mineral oil) in 8 ml of DMF to a 50-ml round bottom flask.

2. Allow the reaction to stir at r. t. for 2 hr under an argon atmosphere.

3. Dissolve 1.86 g (4.95 mmol) of S.2 in 13.5 ml of DMF in a 50-ml round bottom flask.

4. Add the solution from step 3 into the 50-ml round bottom flask from step 1, attach a condenser to the flask, and stir the reaction mixture at 105 °C for 15 hr under an argon atmosphere.

5. Allow the solution to cool to r. t. and remove the precipitate by vacuum filtration. Wash the residue with AcOEt.

6. Evaporate the organic solvent from step 5 to dryness using a rotary evaporator equipped with a vacuum pump and a cooling trap.

7. Purify the residue by column chromatography on 50 g of silica gel in a 6.4 x 45-cm column. Deactivate the column with DCM/Et$_3$N 97:3 and elute the product using a step-wise gradient of DCM/MeOH/Et$_3$N 35:1:0.01 to 25:1:0.01 (v/v). Collect fractions containing the product and evaporate to dryness.

The resulting product, 3'-O-(4,4'-Dimethoxytriphenyl)methyl-(S)-9-(2,3-dihydroxypropyl)adenine (S.3), should be obtained in a 48% yield (1.21 g, 2.38 mmol) as a colorless foam. TLC (AcOEt/MeOH 25:1): R$_f$ = 0.19. ^{1}H NMR (300 MHz, CDCl$_3$): δ=3.02 (dd, J=6.3, 9.6 Hz, 1H), 3.25 (dd, J=5.7, 9.3 Hz, 1H), 3.78 (s, 6H), 4.18 (m, 1H), 4.29 (dd, J=6.6, 14.4 Hz, 1H), 4.40 (dd, J=2.4, 14.2 Hz, 1H), 5.75 (br s, 2H), 6.80 (m, 4H), 7.20-7.30 (m, 7H), 7.37-7.40 (m, 2H), 7.72 (s, 1H), 8.26 (s, 1H).

Prepare S.4

1. Dissolve 1.69 g (3.29 mmol) of S.3 in 25 ml of anhydrous pyridine in a 100-ml round bottom flask and add 1.7 ml (13.55 mmol) of TMS-Cl to the solution.

2. Stir the reaction at r. t. for 2 hr under an argon atmosphere.

3. Cool the mixture in a water-ice bath to 0°C and add in a drop-wise fashion 0.59 ml (5.1 mmol) of BzCl.

4. Continue stirring the reaction at 0°C for 0.5 hr, then at r. t. for another 2 hr.

5. Cool the solution in a water-ice bath to 0°C and add 4 ml of water to quench the reaction.

38

6. Keep the reaction at 0°C and add 8 ml of conc. NH$_4$OH by injection after 15 min.

7. Stir the reaction in the ice bath for an additional 50 min.

8. Evaporate the solvent to dryness using a rotary evaporator equipped with a vacuum pump and a cooling trap.

9. Re-suspend the residue in DCM and adsorb the mixture onto silica gel (sample/silica gel ratio 1:1.2, w/w).

10. Purify the sample from step 9 by column chromatography (Deactivate the column with hexanes/Et$_3$N 97:3 and elute using hexanes/AcOEt/Et$_3$N 1:2:0.01, then AcOEt/Et$_3$N 100:1, and finally AcOEt/MeOH/Et$_3$N 25:1:0.01). Collect fractions containing the product and evaporate to dryness.

The resulting product, 3'-*O*-(4,4'-Dimethoxytriphenyl)methyl-*N*^{6}-benzoyl-(S)-9-(2,3-dihydroxypropyl)adenine (S.4), should be obtained in a 78% yield (1.6 g, 2.57 mmol) as a colorless foam. TLC (AcOEt/MeOH 25:1): R$_f$ = 0.41. ^{1}H NMR (500 MHz, CDCl$_3$): δ=3.14 (dd, J=6.0, 9.5 Hz, 1H), 3.21 (dd, J=6.0, 9.5 Hz, 1H), 3.78 (s, 6H), 4.19 (m, 1H), 4.34 (dd, J=7.2, 14.2 Hz, 1H), 4.48 (dd, J=2.8, 14.2 Hz, 1H), 6.81 (d, J=9.0 Hz, 4H), 7.21 (t, J=7.2 Hz, 1H), 7.25-7.27 (m, 6H), 7.39 (d, J=7.0 Hz, 2H), 7.53 (t, J=7.8 Hz, 2H), 7.61 (t, J=7.5 Hz, 1H), 8.00 (s, 1H), 8.04 (d, J=7.5 Hz, 2H), 8.74 (s, 1H), 9.03 (br s, 1H).

Prepare S.5

1. Dry 1.20 g (1.95 mmol) of S.4 overnight under vacuum of 0.05 mmHg in a 100-ml round bottom flask.

2. Dissolve S.4 from step 1 and 2.1 ml (11.2 mmol) of DIPEA in 33 ml of DCM (freshly distilled over CaH_2) under an argon atmosphere.

3. In a drop-wise fashion, add 0.56 ml (2.53 mmol) of chloro(2-cyanoethoxy)(diisopropylamino)phosphine to the reaction mixture.

4. After 2 hr, add another 0.06 ml (0.25 mmol) of chloro(2-cyano-ethoxy)-(diisopropylamino)phosphine (1/10 the volume from step 3) to the reaction mixture.

5. Stir the mixture at r. t. under an argon atmosphere for a total of 4 hr.

6. Evaporate the solvent and purify the residue by column chromatography. Deactivate the column with hexanes/Et_3N 97:3 and elute the product using a step-wise gradient of hexanes/AcOEt/Et_3N 1:2:0.01 to hexanes/AcOEt/MeOH/Et_3N 1:2:0.1:0.01. Combine the product fractions and evaporate to dryness.

The resulting product, 2'-O-(2-cyanoethoxy)(diisopropylamino)-phosphino-3'-O-(4,4'-dimethoxytriphenyl)methyl-N^6-benzoyl-(S)-9-(2,3-dihy droxypropyl)adenine (S.5) should be obtained in a 82% yield (1.30 g, 1.60 mmol) as a colorless foam. TLC (hexanes/AcOEt/MeOH 16:32:1): R_f = 0.18, 0.34. ^{1}H NMR (500 MHz, $CDCl_3$): δ=1.00 (d, J=6.5 Hz, 6H), 1.06 (d, J=7.0

Hz, 6H), 1.10 (d, J=7.0 Hz, 12H), 2.39 (t, J=6.2 Hz, 2H), 2.47 (q, J=6.2 Hz,

2H), 3.15 (dd, J=6.0, 10.0 Hz, 1H), 3.21-3.26 (m. 2H), 3.30 (dd, J=4.0, 12.5

Hz, 1H), 3.45-3.69 (m, 8H), 3.78 (s, 6H), 3.79 (s, 6H), 4.37 (m, 2H), 4.47-4.58

(m, 4H), 6.82 (m, 8H), 7.17-7.33 (m, 14H), 7.45 (m, 4H), 7.53 (m, 4H), 7.60

(m, 2H), 8.04 (m, 4H), 8.05 (s, 1H), 8.06 (s, 1H), 8.78 (s, 1H), 8.79 (s, 1H),

9.03 (br, 2H). ^{31}P NMR (202 MHz, CDCl$_3$): δ=148.98, 149.59.

2.6 Basic Protocol 3

Synthesis of 2'-*O*-(2-cyanoethoxy)(diisopropylamino)phosphino-3'-*O*-

(4,4'-dimethoxytriphenyl)methyl-*N*^{4}-benzoyl-(S)-1-(2,3-dihydroxypropyl)-

cytosine

This protocol describes the synthesis of *N*^{4}-benzoyl-protected (S)-GNA

cytidine nucleoside phosphoramidite S.7 from S.2 (Figure 2.4).
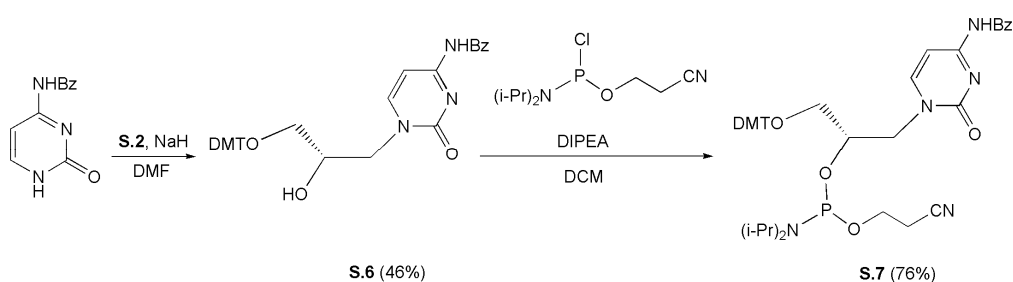
Figure 2.4. Preparation of S.7. Abbreviations: NaH, sodium hydride; DMF, dimethylformamide; DIPEA, diisopropylethylamine; DCM, dichloromethane. Reprinted with permission from Current Protocols in Nucleic Acid Chemistry. Copyright 2010 John Wiley & Sons, Inc.

Materials

DMT-O-(S)-glycidol (S.2)

N^4-benzoylcytosine

41

Sodium hydride (60% in mineral oil) (NaH)

Dimethylformamide (DMF), anhydrous

Dichloromethane (DCM)

Diisopropylethylamine (DIPEA)

Chloro(2-cyanoethoxy)(diisopropylamino)phosphine

Hexanes

Acetone

Ethyl acetate (AcOEt)

Triethylamine (Et$_3$N)

50- and 100-ml round bottom flasks

Oil bath

Magnetic stir plate and stir bar

Gas balloon

Graham condenser

Rotary evaporator equipped with a vacuum pump

1.3 x 30-cm chromatography column

6.4 x 45-cm chromatography column

Silica gel (60 Å, 230-400 mesh)

TLC plate, EMD silica gel 60 F$_{254}$

200-ml separatory funnel

UV lamp, 254 nm

Büchner funnel

Filter paper

Additional reagents and equipment for TLC and column chromatography

Prepare S.6

1. Add 1.94 g (8.99 mmol) of N^4-benzoyl protected cytosine and 73 mg (1.82 mmol) of NaH (60% in mineral oil) in 18 ml of DMF in a 50-ml round bottom flask.

2. Stir the reaction at r. t. for 1 hr under an argon atmosphere using a gas balloon.

3. Dissolve 3.02 g (8.03 mmol) of S.2 in 19 ml of DMF in a 50-ml round bottom flask.

4. Add the solution from step 3 to the 50-ml round bottom flask from step 1, attach a condenser to the flask, and stir the mixture at 110°C for 20 hr.

5. Cool the reaction mixture to r. t. and remove the precipitate by vacuum filtration. Wash the residue with AcOEt.

6. Evaporate the organic solvent from step 5 to dryness using a rotary evaporator equipped with a vacuum pump and a cooling trap.

7. Purify the residue by column chromatography on 80 g of silica gel in a 6.4 x 45-cm column. Deactivate the column with hexanes/Et$_3$N 97:3 and elute using hexanes/acetone/Et$_3$N 3:2:0.01. Collect the product fractions and evaporate to dryness.

The resulting product, 3'-*O*-(4,4'-dimethoxytriphenyl)methyl-*N*[4]-benzoyl- (S)-1-(2,3-dihydroxypropyl)cytosine (S.6), should be obtained in a 46% yield (2.2 g, 3.70 mmol) as a light-yellow foam. TLC (DCM/MeOH 25:1): R_f = 0.21. ^{1}H NMR (300 MHz, CDCl$_3$): δ= 3.12 (dd, J=5.7, 9,6 Hz, 1H), 3.26 (dd, J=5.6, 9.4 Hz, 1H), 3.78 (s, 6H), 3.84 (dd, J=6.8, 13.8 Hz, 1H), 4.22 (m, 1H), 4.36 (dd, J=2.7, 13.8 Hz, 1H), 6.84 (m, 4H), 7.20-7.64 (m, 14H), 7.89 (s, 1H), 7.90 (s, 1H), 8.67 (s br, 1H).

Prepare S.7

1. Dry 1.0 g (1.7 mmol) of S.6 overnight under vacuum of 0.05 mmHg in a 100-ml round bottom flask.

2. Dissolve S.6 from step 1 and 1.9 ml (10.2 mmol) of DIPEA in 29 ml of DCM (freshly distilled over CaH$_2$) under an argon atmosphere.

3. Add 0.5 ml (2.2 mmol) of chloro(2-cyanoethoxy)(diisopropylamino)-phosphine to the reaction mixture in a drop-wise fashion.

4. After 2 hr, add another 0.05 ml (0.22 mmol) of chloro(2-cyanoethoxy)-(diisopropylamino)phosphine (1/10 the volume from step 3) to the reaction mixture.

5. Allow the reaction to stir at r. t. under an argon atmosphere for a total of 3 hr.

6. Evaporate the solvent and purify the residue by column chromatography. Deactivate the column with hexanes/Et$_3$N 97:3 and elute using hexanes/acetone/Et$_3$N 2:1:0.01. Combine the product fractions and evaporate to dryness.

The resulting product, 2'-O-(2-cyanoethoxy)(diisopropylamino)-phosphino-3'-O-(4,4'-dimethoxytriphenyl)methyl-N^4-benzoyl-(S)-1-(2,3-di-hydroxypropyl)cytosine (S.7) should be obtained in a 76% yield (1.02 g, 1.29 mmol) as a light-yellow foam. TLC (hexanes/actone/Et$_3$N 2:1:0.01): R_f = 0.21. ^{1}H NMR (500 MHz, CDCl$_3$): δ=1.12-1.16 (m, 16H), 1.27 (t, J=7.2 Hz, 8H), 2.43 (t, J=6.4 Hz, 2H), 2.55-2.77 (m, 3H), 3.15 (dd, J=3.0, 10.0 Hz, 1H), 3.22 (dd, J=4.5, 10.0 Hz, 1H), 3.28 (dd, J=4.0, 10.0 Hz, 1H), 3.42-3.90 (m, 10H), 3.78 (s, 6H), 3.79 (s, 6H), 4.31-4.45 (m, 4H), 6.82-6.85 (m, 8H), 7.19-7.36 (m, 16 H), 7.45-7.54 (m, 8H), 7.58 (m, 4H), 7.90 (d, J=7.5 Hz, 4H), 8.60 (br, 2H). ^{31}P NMR (202 MHz, CDCl$_3$): δ=149.12, 149.18.

2.7 Basic Protocol 4

Synthesis of 2'-O-(2-cyanoethoxy)(diisopropylamino)phosphino-3'-O-(4,4'-dimethoxytriphenyl)methyl-N^2-isobutyryl-(S)-9-(2,3-dihydroxypropyl)-guanine

This protocol describes the synthesis of N^2-isobutyryl-protected (S)-GNA guanosine nucleoside phosphoramidite S.12 from S.1 (Figure 2.5).

Materials

R-(+)-glycidol (S.1)

2-Amino-6-chloropurine

Potassium carbonate (K$_2$CO$_3$)

Dimethylformamide (DMF), anhydrous
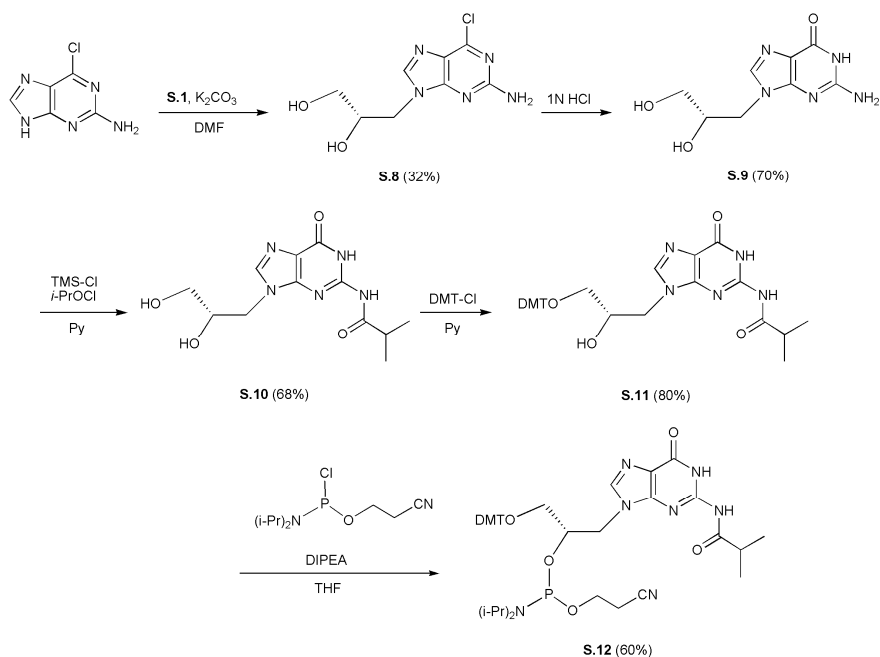
Figure 2.5. Preparation of S.11. Abbreviations: K$_2$CO$_3$, potassium carbonate; DMF, dimethylformamide; HCl, hydrogen chloride; TMSCl, trimethylsilyl chloride; i-PrOCl, isobutyryl chloride; Py, pyridine; DMT-Cl, 4,4'-dimethoxytritylchloride; DIPEA, diisopropylethylamine; THF, tetrahydrofuran. Reprinted with permission from Current Protocols in Nucleic Acid Chemistry. Copyright 2010 John Wiley & Sons, Inc.

1N Hydrochloric acid solution (1N HCl)

Pyridine, anhydrous

Trimethylsilyl chloride (TMS-Cl)

Isobutyryl chloride (i-PrOCl)

Ammonium hydroxide, concentrated (conc. NH$_4$OH)

Tetrahydrofuran (THF)

Pentane

Diisopropylethylamine (DIPEA)

Chloro(2-cyanoethoxy)(diisopropylamino)phosphine

Hexanes

Ethyl acetate (AcOEt)

Triethylamine (Et$_3$N)

Methanol (MeOH)

25-, 50-, 100-, and 250-ml round bottom flasks

Oil bath

Magnetic stir plate and stir bar

Gas balloon

Graham condenser

Rotary evaporator equipped with a vacuum pump

1.3 x 30-cm chromatography column

6.4 x 45-cm chromatography column

Silica gel (60 Å, 230-400 mesh)

TLC plate, EMD silica gel 60 F$_{254}$

UV lamp, 254 nm

Büchner funnel

Filter paper

Additional reagents and equipment for TLC and column chromatography

Prepare S.8

1. Add 0.96 ml (14.5 mmol) of R-(+)-glycidol, 2.4 g (14.2 mmol) of

   2-amino-6-chloropurine and 0.33 g (2.4 mmol) of K$_2$CO$_3$ to 45 ml of DMF

   in a 100-ml round bottom flask.

2. Attach a condenser to the flask and stir the reaction at 90°C for 14 hr.

3. Immediately remove the precipitates by vacuum filtration before the reaction mixture cools to r. t. Wash the residue with MeOH.

4. Remove the solvent using a rotary evaporator equipped with a vacuum pump and a cooling trap.

5. Resuspend the residue in MeOH and adsorb the mixture onto silica gel (sample/silica gel ratio 1:1.2, w/w).

6. Separate the mixture from step 5 by column chromatography using a step-wise gradient of AcOEt/MeOH 10:1 to 20:3 (v/v). Collect the product fractions and evaporate to dryness.

The resulting product, (S)-9-(2,3-dihydroxypropyl)-2-amino-6-chloro-purine (S.8) should be obtained in a 32% yield (1.16 g, 4.77 mmol) as a white solid. TLC (AcOEt/MeOH 20:3): $R_f = 0.40$. ^{1}H NMR (500 MHz, DMSO-d$_6$): δ=3.30 (m, 1H), 3.40 (m, 1H), 3.81(m, 1H), 3.90 (dd, J=8.8, 13.8 Hz, 1H), 4.18 (dd, J=3.2, 13.8 Hz, 1H), 4.81 (t, J=5.5 Hz, 1H), 5.09 (d, J=5.5 Hz, 1H), 6.88 (s br, 2H), 8.01 (s, 1H).

Prepare S.9

1. Dissolve 1.84 g (7.54 mmol) of S.8 in 66 ml of 1N HCl in a 250-ml round bottom flask.

2. Stir the reaction at 85°C for 3 h and cool the mixture to r. t.

3. Adjust the pH of the solution from step 2 using conc. NH$_4$OH to pH=9.0. The white product precipitates out of the solution.

4. Collect the product with vacuum filtration and remove the solvent under high vacuum.

The resulting product, (S)-9-(2,3-dihydroxypropyl)-guanine (S.9) should be obtained in a 70% yield (1.20 g, 5.28 mmol) as a white solid. TLC (AcOEt/MeOH 20:3): $R_f = 0.04$. ^{1}H NMR (500 MHz, DMSO-d$_6$): δ=3.28 (m, 1H), 3.40 (m, 1H), 3.76 (m, 1H), 3.80 (dd, J=8.0, 13.5 Hz, 1H), 4.08 (dd, J=3.5, 13.5 Hz, 1H), 4.78 (t, J=5.5 Hz, 1H), 5.06 (d, J=5.0 Hz, 1H), 6.46 (s, 2H), 7.59 (s, 1H), 10.54 (s, 1H).

Prepare S.10

1. Add 350 mg (1.55 mmol) of S.9 in 10 ml of anhydrous pyridine to a 50-ml round bottom flask in a water-ice bath at 0°C under an argon atmosphere.

2. Add 1.5 ml (11.8 mmol) of TMS-Cl to the solution and stir the reaction at 0°C for 30 min and thereafter at r. t. for another 2 hr.

3. Cool the mixture in a water-ice bath to 0°C and add 1.3 ml (7.75 mmol) of i-PrOCl in a drop-wise fashion to the reaction mixture.

4. Stir the reaction at 0°C for 30 min and thereafter at r. t. for another 4 hr.

5. Cool the solution to 0°C in a water-ice bath and quench the reaction by adding 2 ml of water to the solution, and stir for an additional 15 min.

6. Add 2 ml of conc. NH$_4$OH and incubate for 50 min at 0°C.

7. Evaporate the solvent to dryness using a rotary evaporator equipped with a vacuum pump and a cooling trap.

8. Resuspend the residue in MeOH, adsorb the mixture onto silica gel (sample/silica gel ratio 1:1.2, w/w).

9. Purify the sample by column chromatography using a step-wise gradient of AcOEt/MeOH 20:3 to 40:9 (v/v). Collect the product fractions and evaporate to dryness.

The resulting product, N^2-Isobutyryl-(S)-9-(2,3-dihydroxypropyl)-guanine (S.10) should be obtained in a 68% yield (283 mg, 0.96 mmol) as a white solid. TLC (AcOEt/MeOH 20:3): R_f = 0.15. ^{1}H NMR (300 MHz, DMSO-d$_6$): δ=1.08 (d, J=7.2 Hz, 6H), 2.75 (sept, J=7.1 Hz, 1H), 3.32 (m, 1H), 3.38 (m, 1H), 3.78 (m, 1H), 3.90 (dd, J=8.7, 13.8 Hz, 1H), 4.18 (dd, J=3.3, 13.8 Hz, 1H), 4.80 (t, J=5.4 Hz, 1H), 5.07 (d, J=5.7 Hz, 1H), 7.86 (s, 1H), 11.65 (s br, 1H), 12.02 (s br, 1H).

Prepare S.11

1. Dissolve 780 mg (2.65 mmol) of S.10 in 11.5 ml of anhydrous pyridine in a 50-ml round bottom flask.

2. Add 1.1 g (3.24 mmol) of DMT-Cl to the solution from step 1.

3. Stir the reaction at r. t. for 3 hr under an argon atmosphere.

4. Remove the solvent using a rotary evaporator equipped with a vacuum pump and a cooling trap.

5. Purify the residue from step 4 by column chromatography. Deactivate the column with hexanes/Et$_3$N 97:3 and elute using gradient elution (hexanes/AcOEt/Et$_3$N 1:2:0.01, then AcOEt/Et$_3$N 100:1, and finally

AcOEt/MeOH/Et$_3$N 25:1:0.01). Collect the product fractions and evaporate to dryness.

The resulting product, 3'-*O*-(4,4'-dimethoxytriphenyl)methyl-*N*^{2}-iso-butyryl-(S)-9-(2,3-dihydroxypropyl)-guanine (S.11) should be obtained in a 80% yield (127 mg, 2.12 mmol) as a white solid. TLC (AcOEt/MeOH 50:1): R$_f$ = 0.10. ^{1}H NMR (500 MHz, CDCl$_3$): δ=1.27 (t, J=7.0 Hz, 6H), 2.62 (sept, J=6.8 Hz, 1H), 3.18 (dd, J=5.8, 9.8 Hz, 1H), 3.27 (dd, J=5.0, 9.5 Hz, 1H), 3.80 (s, 6H), 4.03 (dd, J=8.0, 14.0 Hz, 1H), 4.25 (dd, J=2.5, 14.0 Hz, 1H), 4.36 (m, 1H), 4.81 (br, 1H), 6.83 (d, J=8.5 Hz, 4H), 7.22 (t, J=7.2 Hz, 1H), 7.26-7.34 (m, 6H), 7.45 (d, J=7.0 Hz, 2H), 7.55 (s, 1H), 8.43 (br, 1H), 11.79 (br, 1H).

Prepare S.12

1. Dry 200 mg (0.33 mmol) of S.11 overnight under vacuum of 0.05 mmHg in a 25-ml round bottom flask.

2. Dissolve S.11 from step 1 and 0.11 ml (0.6 mmol) of DIPEA in 3.5 ml of THF (freshly distilled over sodium) under an argon atmosphere.

3. Add 0.1 ml (0.45 mmol) of chloro(2-cyanoethoxy)(diisopropyl-amino)phosphine to the reaction mixture in a drop-wise fashion in a water-ice bath at 0°C.

4. Stir the reaction at r. t. After 2 hr, add another 0.01 ml (0.05 mmol) of chloro(2-cyanoethoxy)(diisopropylamino)phosphine (1/10 the volume from step 3) to the reaction mixture.

5. Allow the reaction to stir overnight at r. t under an argon atmosphere.

6. Evaporate the solvent and dissolve the residue in a minimal amount of DCM. Add the solution to 100 ml of vigorously stirring pentane. Collect the white precipitate by vacuum filtration.

7. Purify the residue by column chromatography. Deactivate the column with hexanes/Et$_3$N 97:3 and elute with a step-wise gradient of DCM/AcOEt/Et$_3$N 45:45:1 to AcOEt/Et$_3$N 100:1. Combine the product fractions and evaporate to dryness.

The resulting product, 2'-*O*-(2-cyanoethoxy)(diisopropylamino)-phosphino-3'-*O*-(4,4'-dimethoxytriphenyl)methyl-*N*^{2}-isobutyryl-(S)-9-(2,3-dihydroxypropyl)-guanine (S.11) should be obtained in a 60% yield (158 mg, 0.2 mmol) as a white foam. TLC (AcOEt/MeOH 50:1): R$_f$ = 0.32, 0.42. ^{1}H NMR (500 MHz, CDCl$_3$): δ=1.06 (t, J=7.0 Hz, 2H), 1.10-1.15 (m, 17H), 1.20-1.29 (m, 17H), 2.47 (t, J=6.0 Hz, 1H), 2.52-2.65 (m, 4H), 3.02 (dd, J=4.0, 9.5 Hz, 1H), 3.08 (dd, J=5.2, 9.8 Hz, 1H), 3.13-3.23 (m, 2H), 3.46-3.82 (m, 10H), 3.78 (s, 6H), 3.79 (s, 6H), 4.26-4.38 (m, 5H), 4.53 (m, 1H), 6.80 (m, 8H), 7.18-7.33 (m, 14H), 7.42-7.48 (m, 4H), 7.49 (s, 1H), 7.55 (s, 1H), 8.38 (br, 1H), 8.70 (br, 1H). ^{31}P NMR (202 MHz, CDCl$_3$): δ=148.20, 148.28.

2.8 Basic Protocol 5

Synthesis of 2'-*O*-(2-cyanoethoxy)(diisopropylamino)phosphino-3'-*O*-(4,4'-dimethoxytriphenyl)methyl-(S)-1-(2,3-dihydroxypropyl)thymine

This protocol describes the synthesis of (S)-GNA thymidine nucleoside phosphoramidite monomer S.14 from S.2 (Figure 2.6).
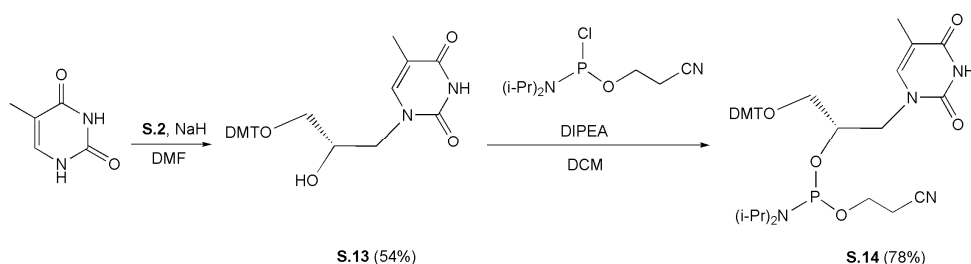
Figure 2.6. Preparation of S.14. Abbreviations: NaH, sodium hydride; DMF, dimethylformamide; DIPEA, diisopropylethylamine; DCM, dichloromethane. Reprinted with permission from Current Protocols in Nucleic Acid Chemistry. Copyright 2010 John Wiley & Sons, Inc.

Materials

DMT-O-(S)-glycidol (S.2)

Thymine

Sodium hydride (60% in mineral oil) (NaH)

Dimethylformamide (DMF), anhydrous

Dichloromethane (DCM)

Diisopropylethylamine (DIPEA)

Chloro(2-cyanoethoxy)(diisopropylamino)phosphine

Hexanes

Ethyl acetate (AcOEt)

Triethylamine (Et$_3$N)

Methanol (MeOH)

25- and 50-ml round bottom flasks

Oil bath

Magnetic stir plate and stir bar

Gas balloon

Graham condenser

Rotary evaporator equipped with a vacuum pump

1.3 x 30-cm chromatography column

6.4 x 45-cm chromatography column

Silica gel (60 Å, 230-400 mesh)

TLC plate, EMD silica gel 60 F_{254}

200-ml separatory funnel

UV lamp, 254 nm

Büchner funnel

Filter paper

Additional reagents and equipment for TLC and column chromatography

Prepare S.13

1. Add 0.95 g (7.54 mmol) of thymine and 64 mg (1.60 mmol) of NaH (60% in mineral oil) in 12 ml DMF to a 50-ml round bottom flask.

2. Stir the mixture at r. t. for 2 hr under an argon atmosphere.

3. Dissolve 2.68 g (7.13 mmol) of S.2 in 15 ml of DMF in a 50-ml round bottom flask.

4. Add the solution from step 3 to the 50-ml round bottom flask from step 1, attach a condenser to the flask, and stir the mixture at 110°C for 20 hr under an argon atmosphere.

5. Cool the reaction mixture to r. t. and remove the precipitate by vacuum filtration. Wash the residue with AcOEt.

6. Evaporate the organic solvent from step 5 to dryness using a rotary evaporator equipped with a vacuum pump and a cooling trap.

7. Purify the residue by column chromatography on 60 g of silica gel in a 6.4 x 45-cm column. Deactivate the column with DCM/Et$_3$N 97:3 and elute using DCM/MeOH/Et$_3$N 50:1:0.01. Collect the product fractions and evaporate to dryness.

The resulting product, 3'-O-(4,4'-dimethoxytriphenyl)methyl (S)-1-(2,3-dihydroxypropyl)thymine (S.13), should be obtained in a 54% yield (1.95 g, 3.88 mmol) as a colorless foam. TLC (DCM/MeOH 25:1): R$_f$ = 0.14. ^{1}H NMR (400 MHz, CDCl$_3$): δ= 1.73 (s, 3H), 3.07 (d, J=5.6 Hz, 2H), 3.55 (dd, J=7.2, 14.0 Hz, 1H), 3.68 (s, 6H), 3.89-3.96 (m, 2H), 6.73 (m, 4H), 6.94 (d, J=1.2 Hz, 1H), 7.10-7.22 (m, 7H), 7.30 (m, 2H).

Prepare S.14

1. Dry 1.1 g (2.2 mmol) of S.13 under vacuum of 0.05 mmHg in a 100-ml round bottom flask overnight.

2. Dissolve S.13 from step 1 and 2.2 ml (11.7 mmol) of DIPEA in 36 ml of DCM (freshly distilled over CaH$_2$) under an argon atmosphere.

3. Add 0.62 ml (2.8 mmol) of chloro(2-cyanoethoxy)(diisopropylamino)-phosphine to the reaction mixture in a drop-wise fashion.

4. After 2 hr, add another 0.06 ml (0.3 mmol) of chloro(2-cyanoethoxy)-(diisopropylamino)phosphine (1/10 of the amount from step 3) to the reaction mixture.

5. Stir the mixture at r. t. under an argon atmosphere for a total 3 hr.

Evaporate the solvent and purify the residue by column chromatography. Deactivate the column with hexanes/Et3N 97:3 and elute using hexanes/AcOEt/Et3N 1:2:0.01. Combine the product fractions and evaporate to dryness.

The resulting product, 2'-O-(2-cyanoethoxy)(diisopropylamino)-phosphino-3'-O-(4,4'-dimethoxytriphenyl)methyl-(S)-1-(2,3-dihydroxy-propyl)thymine (S.14) should be obtained in a 78% yield (1.21 g, 1.72 mmol) as a colorless foam. TLC (hexanes/AcOEt/Et$_3$N 1:2:0.01): R_f = 0.33, 0.42. ^{1}H NMR (300 MHz, CDCl$_3$): δ=1.12-1.18 (m, 24H), 1.84 (m, 6H), 2.41 (t, J=6.3 Hz, 2H), 2.59 (t, J=6.3 Hz, 2H), 3.16-3.32 (m, 4H), 3.49-3.88 (m, 10H), 3.78 (s, 6H), 3.79 (s, 6H), 4.03-4.23 (m, 4H), 6.83 (m, 8H), 7.05 (m, 2H), 7.18-7.36 (m, 14H), 7.44-7.46 (m, 4H). ^{31}P NMR (121 MHz, CDCl$_3$): δ=148.76, 148.95.

2.9 Basic Protocol 6

Synthesis, Isolation, and Characterization of Glycerol Nucleic Acid (GNA) Oligonucleotides

The glycerol nucleoside phosphoramidites are completely stable under standard oligonucleotide synthesis conditions on a DNA synthesizer. An

Applied Biosystems 3400 DNA synthesizer has been used successfully on both a 0.2- and 1.0-µmol scale synthesis. Optimal yields were achieved by extending the coupling time of the GNA phosphoramidite monomers to 300 s (5 min). Because the free hydroxyl termini of GNA nucleotides are subject to decomposition during the base-deprotection step (Tsai et al., 2007), we recommend capping the 5'- and 3'-ends of each GNA sequence with a single DNA residue, typically T (Yang et al., 2007). Several techniques have been reported for the isolation and purification of sequences containing GNA nucleotides, such as C18 reverse phase HPLC and denaturing polyacrylamide gel electrophoresis (PAGE). This protocol describes the synthesis of GNA oligonucleotides from GNA monomer phosphoramidite precursors.

Materials

GNA phosphoramidite monomers bearing nucleobases A, C, G, and T (S.5, S.7, S.12, and S.14, respecttively)

Anhydrous acetonitrile (MeCN) (Applied Biosystems)

Standard 5'-O-(4,4'-dimethoxytrityl)phosphoramidites (Glen Research)

Argon gas (Ultra high purity)

Ammonium hydroxide, concentrated (conc. NH_4OH)

n-Butanol

Nanopure water

25-ml round bottom flasks

4Å molecular sieves (freshly activated by heating at 300°C for at least 3 hr)

0.45-µm disposable syringe filter

Syringes and needles

Screw-capped eppendorf tube (2 ml)

Screw-capped falcon centrifuge tubes

Additional reagents and equipments for automated solid-phase

oligodeoxyribonucleotide synthesis, isolation and characterization of synthetic

nucleic acids

Synthesize GNA oligonucleotides (0.2 µmol scale synthesis)

1. Dissolve each GNA phosphoramidte (pre-dried under high vacuum for 48

   hr) in anhydrous MeCN in a 25-ml round bottom flask to make an

   approximate 0.05 M solution (150 mg, 150 mg, 150 mg, and 130 mg in 3

   ml MeCN for S.5 (A), S.7 (C), S.12 (G), and S.14 (T), respectively).

2. Dry the phosphoramidite solutions from step 1 over freshly activated

   molecular sieves overnight at r. t.

3. Filter the amidites solution using 0.45-µm disposable syringe filters

   directly into clean, dry bottles and immediately place the bottles on the

   corresponding amidite ports of the DNA synthesizer.

   Steps 1 through 3 are carried out under argon atmosphere. All the

glassware, syringes and needles are pre-dried in a drying oven and then cooled

to r. t. in a desiccator for later use.Start the automated solid-phase

oligonucleotide synthesis from CPG-column pre-charged with a

deoxynucleotide monomer or the universal support.

4. After completing the synthesis, remove the column and dry it for 2 min using vacuum pipeline.

5. Open the CPG-column container and transfer the CPG beads into a 2-ml screw-capped eppendorf tube.

Deprotect and purify GNA oligonucleotides

6. Add 1 ml of conc. NH$_4$OH, seal the tube, and incubate at 55°C for 48 hr.

7. Place the tube on ice and cool for at least 10 min prior to opening

8. Transfer the supernatant to a Falcon tube and add 10 ml of n-butanol.

9. Place the tube on ice for 5 min.

10. Centrifuge for 5 min at 4000 rpm, discard the supernatant and collect the oligonucleotide (white precipitate)

11. Re-suspend the oligonucleotide pellet in 500 µl of nanopure water.

12. Purify the crude product by denaturing PAGE and desalt the oligonucleotide by ethanol precipitation.

13. Quantify the oligonucleotide by UV absorbance using Beer's law at 254 nm.

Check molecular weight

14. Check the molecular weight of the desired GNA oligonucleotide by MALDI-TOF mass spectroscopy using aqueous saturated 3-hydroxypicolinc acid solution as matrix. Prepare matrix just before spotting the sample onto a MALDI-TOF mass spectrometry gold plate.

See Figure 2.7 for sample MALDI-TOF mass spectroscopy result.
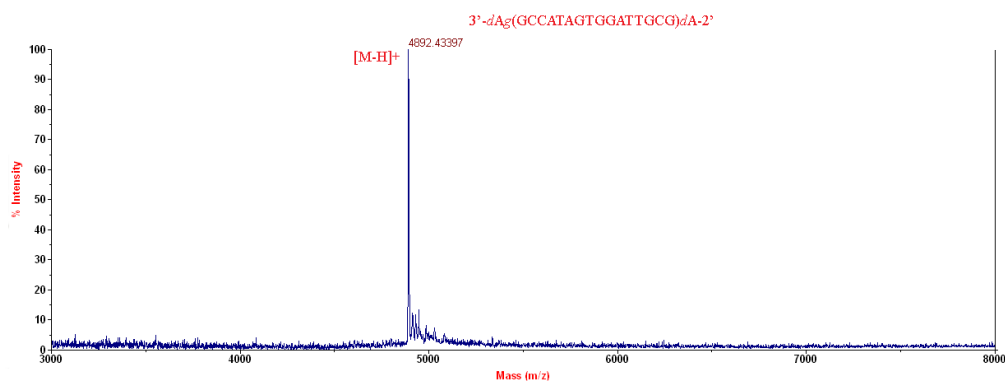
Figure 2.7. MALDI-TOF mass spectrum of GNA oligonucleotide. Reprinted with permission from Current Protocols in Nucleic Acid Chemistry. Copyright 2010 John Wiley & Sons, Inc.

2.10 Commentary

Background Information

Synthesis

The "spring-loaded" epoxide ring-opening procedure was first reported by Acevedo and Andrews to increase the structural diversity of acyclic libraries aimed at creating therapeutic agents to cellular and viral targets (Acevedo et al., 1996). This unit presents a modified version of this synthetic strategy. GNA nucleoside derivatives can also be synthesized using a nucleobase substitution approach with a tosylated isopropylidene-glycerol (Holý, 1975). However, this approach requires more synthetic steps and results in a lower overall yield. Another approach is to use modified base protecting groups (N-dimethylformamidine group for A and G, acetamide for C) (Schlegel et al., 2009b). Both of these protecting groups are compatible with faster oligonucleotide deprotection conditions, which can be advantageous for

minimizing GNA decomposition during the final deprotection step in ammonium hydroxide.

Properties of glycerol nucleic acid

GNA has an acyclic backbone composed of a three-carbon sugar connected by phosphodiester linkages. The repeating backbone unit contains one stereocenter, which results in two enantiomers, namely (S)- and (R)-GNA. Both (S)- and (R)-GNA form antiparallel duplex structures with complementary GNA strands of the same stereoconfiguration, and GNA duplexes are generally more stable than natural DNA and RNA (Zhang et al., 2005, 2006, Yang et al., 2007). For this reason, GNA represents the simplest solution to a stable nucleic acid structure based on phosphodiester bonds. Cross-pairing between complementary strands of GNA and RNA results in the formation of a dramatically weaker duplex than pure GNA or RNA alone. Cross-pairing between complementary sequences of GNA and DNA has not been observed, suggesting that the GNA helix is incompatible with the standard B-form helix of natural DNA (Zhang et al., 2005, Yang et al., 2007). This is also indicated by the fact that the incorporation of GNA residues into DNA strands strongly destabilizes duplex formation. Cross-pairing between (R)- and (S)-GNA is also highly destabilizing (Schlegel et al., 2007). The X-ray crystal structure of a self-complementary (S)-GNA double helix was recently solved to a resolution limit of 1.3 Å (Schlegel et al., 2008). The three-dimensional structure revealed that (S)-GNA adopts a right-handed helix

that is significantly different from the standard A- and B-form helical

structures observed for natural RNA and DNA, respectively. A striking feature

of the GNA duplex is the large average slide (3.4 Å) between adjacent base

pairs due to a large backbone-base inclination of the duplex (42-50º in GNA

versus almost 0º in the B-form DNA). This structural change results in

extensive interstrand base stacking and significantly less intrastrand base

stacking common to standard A- and B-form helical structures. Another

unusual feature is the formation of a loosely wrapped helical ribbon around the

central helix axis, which replaces the major groove by a convex surface and

leaves just one large minor groove (Schlegel et al., 2008, 2009a, and 2010).

GNA application

GNA, which is a structurally simplified artificial oligonucleotide, is

capable of forming stable duplex structures by Watson-Crick base-pairing

rules. This property indicates that GNA represents an interesting synthetic

material for constructing programmable nucleic acid nanostructures. Toward

this goal, we recently described the synthesis of a nanostructure composed

entirely of GNA. The GNA nanostructure, a 4-helix junction (4HJ) (Zhang et

al., 2008), mimics a fixed-sequence DNA Holliday junction described

previously by Seeman and co-workers (Kallenbach et al., 1983). Because the

GNA backbone contains only one stereocenter per repeating unit, it was

possible to synthesize two mirror-image nanostructures using (S)- and

(R)-GNA. A major finding was that the GNA 4HJ was significantly more

stable than the earlier 4HJ composed entirely of DNA (T_m 73°C for GNA

versus 37°C for DNA). This feature coupled with the ability to construct left-

and right-handed nanostructures provides new opportunities for building

highly stable nanostructures with topologies that are not readily available to

DNA.

Critical Parameters

For all reactions using anhydrous solvents, the glassware must be pre-dried

in a drying oven (80°C). To obtain optimal separations and to reduce the

purification time, we recommend using flash column chromatography for all

purification steps. For phosphoramidite synthesis, the reaction progress is

monitored by TLC and should be stopped as soon as the starting material is

consumed. The phosphitylation reaction affords two diastereomeric

phosphoramidite products, which are general visualized as two spots on the

TLC plate. However, phosphoramidites bearing cytosine migrate on the TLC

plate as a one spot, even though the products are actually two isomers. The

GNA phosphoramidite building blocks must be dried under high vacuum in a

desiccator for 24 hr to give a well-foamed solid before they can be used for

solid-phase synthesis.

Trouble shooting

The phosphoramidites containing the nucleobase, guanine, will decompose

when purified on a silical gel column (Zhang et al., 2006). In this protocol, a

modified method is described for column chromatography purification that

affords the desired intact compound verified by ^{1}H and ^{31}P NMR characterization.

Anticipated Results

The protocols described in this unit are useful for preparing GNA oligonucleotides. The GNA phosphoramidite monomers can be synthesized on a multi-gram scale. The oligonucleotide synthesis is similar to the standard DNA phosphoramidite protocol with the noted exception of elongated coupling times. The average stepwise yield is >95%, which is sufficient to produce GNA oligonucleotides in high yield.

Time Considerations

Each step in the chemical synthesis can be accomplished in 1-2 working days (including purification). The time for oligonucleotide synthesis varies slightly from the standard phosphoramidite method by longer coupling times. The deprotection, isolation and the analysis of the final synthetic oligonucleotide are accomplished in 2-3 days.

2.11 Reference

Acevedo, O.L., and R.S. Andrews. 1996. Synthesis of propane-2,3-diol combinatorial monomers. *Tetrahedron Lett.* 37: 3913-3934.

Holý, A. 1975. Aliphatic analogues of nucleosides, nucleotides, and oligonucledotides. *Collect. Czech. Chem. Commun.* 40: 187-214.

Kallenbach, N.R., R.-I. Ma, and N.C. Seeman. 1983. An immobile nucleic acid junction constructed from oligonucleotides. *Nature* 305: 829-831.

Schlegel, M.K., L.-O. Essen, and E. Meggers. 2008. Duplex formation of a minimal nucleic acid. *J. Am. Chem. Soc.* 130: 8158-8159.

Schlegel, M.K., L.-O. Essen, and E. Meggers. 2010. Atomic resolution duplex structure of the simplified nucleic acid GNA. *Chem. Commun.* 46: 1094-1096.

Schlegel, M.K., A.E. Peritz, K. Kittigowittana, L. Zhang, and E. Meggers. 2007. Duplex formation of the simplified nucleic acid GNA. *ChemBioChem* 8: 927-932.

Schlegel, M.K., X. Xie, L. Zhang, and E. Meggers. 2009a. Insight into the high duplex stability of the simplified nucleic acid GNA. *Angew. Chem. Int. Ed.* 48: 960-963.

Schlegel, M.K., and E. Meggers. 2009b. Improved phosphoramidite building blocks for the synthesis of the simplified nucleic acid GNA. *J. Org. Chem.* 74: 4615-4618.

Tsai, C.-H., J. Chen, and J.W. Szostak. 2007. Enzymatic synthesis of DNA on glycerol nucleic acid templates without stable duplex formation between product and template. *Proc. Natl. Acad. Sci. USA 104*: 14598-14603.

Yang, Y.-W., S. Zhang, E.O. McCullum, and J.C. Chaput. 2007. Experimental evidence that GNA and TNA were not sequential polymers in the prebiotic evolution of RNA. *J. Mol. Evol.* 65: 289-295.

Zhang, L., A.E. Peritz, P.J. Carroll, and E. Meggers. 2006. Synthesis of glycol nucleic acids. *Synthesis* 4: 645-653.

Zhang, L., A. Peritz, and E. Meggers. 2005. A simple glycol nucleic acid. *J. Am. Chem. Soc.* 127: 4174-4175.

Zhang, R.S., E.O. McCullum, and J.C. Chaput. 2008. Synthesis of two mirror image 4-helix junctions derived from glycerol nucleic acid. *J. Am. Chem. Soc.* 130: 5846-5647.

Zhang, S., C. Switzer, and J.C. Chaput. 2010. The resurgence of acyclic nucleic acids. *Chem. Biodivers.* 7: 245-258.

Chapter 3

SYNTHESIS OF THREOSE NUCLEIC ACID (TNA) PHOSPHORAMIDITE

MONOMERS AND OLIGONUCLEOTIDE POLYMERS

3.1 Contributions

The following chapter describes a published protocol for the synthesis of

TNA phosphoramidite monomers and oligonucleotide polymers. All

experiments were performed by Su Zhang under the guidance of Professor

John Chaput. Su Zhang and John Chaput wrote the manuscript. The results

from this work were submitted to *Current Protocols in Nucleic Acid*

Chemistry.

3.2 Abstract

This unit describes in detail the preparation of (3'→2')-α-L-threo-

furanosyl nucleic acid (TNA) phosphoramidite building blocks bearing

adenine (A), cytosine (C), diaminopurine (D), guanine (G), and thymine (T) as

the bases and the oligonucleotide thereof. Starting from commercially

available L-ascorbic acid, protected threofuranosyl sugars are obtained (Basic

Protocol 1) and subsequently converted to protected threonucleosides via

Vorbrüggen-Hilbert-Johnson reactions. The nucleosides are then converted to

their corresponding phosphoramidites (Basic Protocols 2 to 6), which can then

be used for the solid-phase synthesis of TNA oligonucleotides (Basic Protocol

7).

3.3 Introduction

TNA is a structural alternative of DNA with only a four-carbon

sugar-phosphate backbone repeating unit (Figure 3.1). Although containing

one less bond per repeating unit than DNA, TNA is capable of adapt itself to

Watson-Crick base-paring with DNA, RNA, and TNA complementary
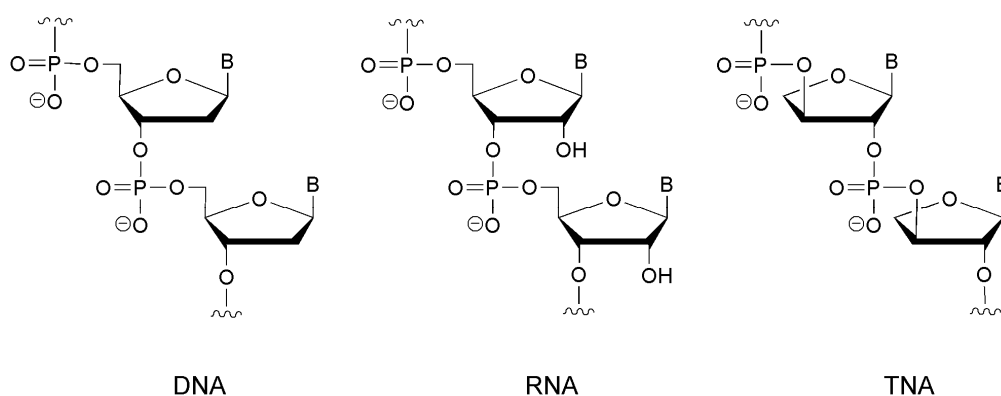
Figure 3.1. Chemical structures of DNA, RNA, and TNA.

sequence, respectively (Schöning et al. 2000; 2002). TNA has been considered

as an evolutionary progenitor of RNA due to the facts that TNA could transfer

genetic information as well as its simpler structure (Orgel 2000). The sequence

space of TNA has been recently explored to test the ability for TNA to possess

chemical and biological functions such as the ability to bind a desired target

with high affinity and specificity and the potential to catalyze reactions on a

particular substrate (Chaput et al. 2003a, 2003b; Horhota et al. 2005; Ichida et

al. 2005a, 2005b). These experiments might provide evidence to answer the

question whether TNA is a viable progenitor candidate of RNA, and hopefully,

shed light on the etiology and evolution of genetic system. This unit describes

the synthesis of TNA phosphoramidites monomers (Basic Protocol 2 through 6)

and the solid-phase synthesis and purification of TNA oligonucleotides using

these building blocks (Basic Protocol 7).

3.4 Basic Protocol 1

Synthesis of protected L-threofuranose, 1-*O*-Acetyl-2,3-di-*O*-benzoyl-L-

threofuranose and 1,2,3-tri-*O*-benzoyl-L-threofuranose

The synthesis of L-threose has been reported by several different groups,

although most of the methods were not suitable for large-scale preparation.

The present protocol is based on Eschenmoser's procedure (Schöning et al.

2000, 2002) and is illustrated in Figure 3.2. The synthesis is initiated from the

oxidative degradation of L-ascorbic acid (**1**) to afford L-threonic acid. A one-

pot lactonization and in-situ benzoylation affords the 2,3-di-*O*-benzoyl-L-

threonolactone (**3**). The lactone is then treated with DIBAL-H to furnish the

2,3-di-*O*-benzoylated lactol. 1-*O*-Acetyl-2,3-di-*O*-benzoyl-L-threofuranose (**5**)

obtained by acetylation of the lactol is used to prepare the C, D, G, and T

threose nucleoside phosphoramidites (see Basic Protocols 2-5, respectively).

The benzoylated lactol 1,2,3-Tri-*O*-benzoyl-L-threofuranose (**6**) is used to

make A threose nucleoside phosphoramidite (see Basic Protocol 6).

Materials

L-Ascorbic acid (**1**)

p-Toluenesulfonic acid (TsOH)

Benzoyl chloride (BzCl)

Diisobutylaluminium hydride (DIBAL-H)

4-Dimethylaminopyridine (DMAP)

Acetic anhydride (Ac$_2$O)

Acetonitrile, anhydrous (MeCN)

Pyridine, anhydrous

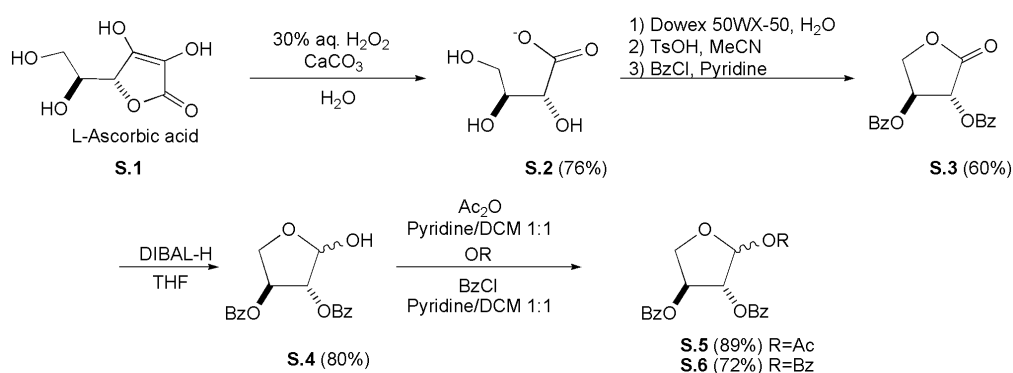Figure 3.2. Preparation of protected L-threofuranoses **S.5** and **S.6**. Abbreviations: TsOH, *p*-toluenesulfonic acid; MeCN, acetonitrile; BzCl, benzoyl chloride; DIBAL-H, diisobutylaluminium hydride; DCM, dichloromethane.

Tetrahydrofuran, anhydrous (THF)

Calcium carbonate (CaCO$_3$)

Hydrogen peroxide (H$_2$O$_2$)

Ultra pure water (H$_2$O)

Dichloromethane (DCM)

Ethyl acetate (EtOAc)

Hexanes

Methanol (MeOH)

Ethanol (EtOH)

1M aqueous hydrochloric acid solution (1M HCl)

Saturated aqueous sodium bicarbonate solution (sat. aq. NaHCO$_3$)

Brine (sat. aq. NaCl)

Saturated aqueous sodium potassium tartrate

Activated charcoal

Magnesium sulfate (MgSO$_4$)

Dowex 50WX4-50 resin

Argon gas

250-, 500-, 1000-, and 2000-ml round-bottom flask

Magnetic stir plate and stir bar

Gas balloon

Rotary evaporator equipped with a vacuum pump

Thin layer chromatography (TLC) plate, EMD silica gel 60 F$_{254}$

Silica gel (60 Å, 230-400 mesh) (Whatman Inc.)

250-, 500-, 1000-, and 2000-ml separatory funnel

Büchner funnel

Allihn condenser

Filter paper

UV lamp, 254 nm

Additional reagents and equipment for thin layer chromatography (TLC) and

column chromatography

Prepare S.2

1. Dissolve 88 g (0.5 mol) of L-ascorbic acid in 700 ml of H_2O in a 2000-ml round bottom flask and slowly add 100 g (1 mol) of $CaCO_3$ with stirring.

2. Cool the mixture to 15 °C and add 200 ml of 30% aq. H_2O_2 over a period of 3 hr using a dropping funnel.

3. Warm the reaction to r. t. and stir the reaction for 18 hr.

4. Add 20 g of activated charcoal to the solution and increase the temperature to 70 °C until no more peroxide.

5. Filter the hot reaction mixture and wash the filter cake with water.

6. Concentrate the filtrate to *ca.* 500 ml under reduced pressure.

7. Add 200 ml of MeOH dropwise using a dropping funnel over 5 hr at r. t. The product gradually precipitates out of the solution during this process.

8. Stir the mixture for an additional 5 hr at r. t., and then add another 200 ml of MeOH to the solution.

9. Isolate the precipitate by vacuum filtration and wash with cold MeOH.

10. Collect the filtrate and repeat step 5-9 to obtain additional precipitate. Dry the product under reduced pressure to a constant weight.

The resulting product, Calcium-L-threonate monohydrate (S.2), is obtained in a 76% yield (64g, 195 mmol) as a white powder. ^{1}H-NMR (300 MHz, D_2O): δ=3.63 (dd, J=11.6, 7.8, 1H), 3.70 (dd, J=11.4, 5.4, 1H), 3.98 (ddd, J=12.6, 5.4, 2.1, 1H), 4.04 (d, J=2.1, 1H).

Prepare S.3

1. Dissolve 30 g (91.3 mmol) of **S.2** in 640 ml of water at 40 °C, and then add 190 ml of Dowex 50WX4-50 resin (acidic form) to the solution.

2. Stir the reaction mixture at 70 °C for 30 min. Remove the resin by vacuum filtration and wash with water.

3. Concentrate the filtrate to dryness under reduced pressure to a constant weight at 50 °C.

4. Resuspend the residue in MeCN, co-evaporate it twice with MeCN, and resuspend the residue in 320 ml of anhydrous MeCN.

5. Add 0.64 g (3.2 mmol) of TsOH to the suspension and reflux the reaction for 24 hr under an argon atmosphere.

6. Allow the reaction to cool to r. t. Filter the solution and evaporate the solvent to dryness under reduced pressure.

7. Dissolve the residue in 128 ml of pyridine and cool the solution to 0 °C in a water-ice bath.

8. Add 26.4 ml (228 mmol) of BzCl dropwise to the solution from step 7.

9. Remove the water-ice bath and stir the reaction at r. t. for 16 hr.

10. Dilute the reaction mixture with 250 ml of DCM and wash the organic solvent with 160 ml of ice-cold 1M aq. HCl. Then wash the organic solvent sequentially with 200 ml of water, 200 ml of sat. aq. $NaHCO_3$ and 130 ml of brine.

11. Dry the organic phase over $MgSO_4$ and evaporate the solvent to dryness under reduced pressure.

12. Crystallize the residue from hexanes/EtOAc 6:1 (v/v).

The resulting product, 2,3-Di-*O*-benzoyl-L-threonolactone (**S.3**), is obtained in a 60% yield (17.9 g, 54.8 mmol) as colorless needles. ^{1}H-NMR (300 MHz, $CDCl_3$): δ=4.38 (m, 1H), 5.02 (m, 1H), 5.81 (m, 2H), 7.45-8.12 (m, 10H).

Prepare S.4

1. Dissolve 20 g (60 mmol) of **S.3** in 250 ml of anhydrous THF at -78 °C under an argon atmosphere, and then add 75 ml (225 mmol) of 1.5 M DIBAL-H solution in Toluene dropwise using a dropping funnel.

2. Stir the reaction at -78 °C until **S.3** is completely consumed.

3. Add 10 ml of MeOH over 3 min to the mixture from step 2 to quench the reaction. And then let the reaction mixture warm to r. t.

4. Dilute the mixture from step 3 with 400 ml of EtOAc and 200 ml of sat. aq. sodium potassium tartrate. Then vigorously stir the mixture at r. t. for 3 hr.

5. Separate the organic phase from step 4. Wash the organic solution with brine and dry it over $MgSO_4$.

6. Remove the solvent under reduced pressure to afford both diastereomers of the product without further purification.

The resulting product, 2,3-Di-*O*-benzoyl-L-threofuranose (**S.4**), is obtained in a 60% yield as colorless oil. TLC (hexanes/EtOAc 2:1): R_f 0.37. ^{1}H-NMR

(400 MHz, CDCl$_3$): δ=3.97 (dd, J=8.1, 3.0, 1H), 4.26 (dd, J=10.2, 4.0), 4.60

(m, 2H), 5.47 (t, J=4.0, 1H), 5.53 (m, 2H), 5.57 (s, 1H), 5.75 (m, 1H), 5.85 (d,

J=4.0, 1H), 7.42-7.47 (m, 8H), 7.53-7.59 (m, 4H), 8.01-8.10 (m, 8H).

Prepare S.5

1. Dissolve the crude **S.4** (19.7 g) in 200 ml of DCM/pyridine 1:1. Add 6.8

   ml (62 mmol) of Ac$_2$O and 50 mg (0.4 mmol) of DMAP to the solution.

2. Stir the reaction at r. t. for 18 hr.

3. Pour the reaction mixture into a stirring solution of ice-cold sat.

   NaHCO$_3$/DCM 1:1. Let the stirring continue at 0 °C for 1 hr.

4. Separate and wash the organic phase with sat. NaHCO$_3$, water, and brine.

   Dry the organic phase over MgSO$_4$.

5. Evaporate the organic solvent to dryness and purify the residue by column

   chromatography (SiO$_2$, hexanes/EtOAc 2:1 (v/v)) to afford both

   diastereomers of the product.

   The resulting product, 1-*O*-Acetyl-2,3-Di-*O*-benzoyl-L-threofuranose (**S.5**),

is obtained in an 89% yield (19.6 g, 53.4 mmol) as a colorless oil. TLC

(hexanes/EtOAc 2:1): R$_f$ 0.43. ^{1}H-NMR (400 MHz, CDCl$_3$): δ=2.04 (s, 3H),

2.16 (s, 3H), 4.05 (dd, J=10.4, 3.6, 1H), 4.22 (dd, J=10.6, 3.8, 1H), 4.62 (dd,

J=10.8, 6.4, 1H), 4.67 (dd, J=10.4, 6.0, 1H), 5.56 (m, 1H), 5.65 (s, 1H), 5.70 (t,

J=4.6, 1H), 5.76 (m, 1H), 6.38 (s, 1H), 6.64 (d, J=4.4, 1H), 7.26-7.50 (m, 8H),

7.58-7.64 (m, 4H), 8.05-8.10 (m, 8H).

Prepare S.6

1. Dissolve 4.1 g (12.5 mmol) of **S.4** in 42 ml of DCM/pyridine 1:1. Add 1.7

   ml (14 mmol) of BzCl to the solution at 0 °C in a water-ice bath.

2. Remove the water-ice bath and stir the reaction at r. t. for 18 hr.

3. Cool the reaction mixture to 0 °C and add 16 ml of 0.1 M aq. HCl.

4. Separate and wash the organic phase with sat. $NaHCO_3$, water, and brine.

   Dry the organic phase over $MgSO_4$.

5. Recrystallize the residue from EtOH to afford the β-isomer.

6. Concentrate the ethanolic mother liquor and run the residue through SiO_2

   using hexanes/EtOAc 2:1 (v/v) to yield the α-isomer.

   The resulting product, 1,2,3-Tri-*O*-benzoyl-L-threofuranose (**S.6**), is

obtained in a combined 72% yield (3.9 g, 9 mmol). ^{1}H-NMR α-isomer (300

MHz, $CDCl_3$): δ=4.15 (dd, J=11.6, 3.0, 1H), 4.70 (dd, J= 11.8, 6.0, 1H), 5.84

(m, 2H), 6.86 (d, J=4.0, 1H), 7.38-7.55 (m, 9H), 7.95-7.20 (m, 6H). β-isomer

(300 MHz, $CDCl_3$): δ=4.32 (dd, J=10.5, 3.0, 1H); 4.74 (dd, J=10.7, 5.9, 1H),

5.64 (dd, J=5.1, 3.0, 1H), 5.82 (s, 1H), 6.64 (s, 1H), 7.40-7.51 (m, 6H),

7.56-7.65 (m, 3H), 8.06-8.14 (m, 6H).

3.5 Basic Protocol 2

   Synthesis of 4-Benzoyl-1-{2'-*O*-[(2-cyanoethoxy)(diisopropylamino)-

phosphino]-3'-*O*-[(4,4'-dimethoxytriphenyl)methyl]-α-L-threofuranosyl}-

cytosine

This protocol describes the synthesis of N^4-benzoylated TNA cytidine nucleoside phosphoramidite **S.10** from **S.5** (Figure 3.3).
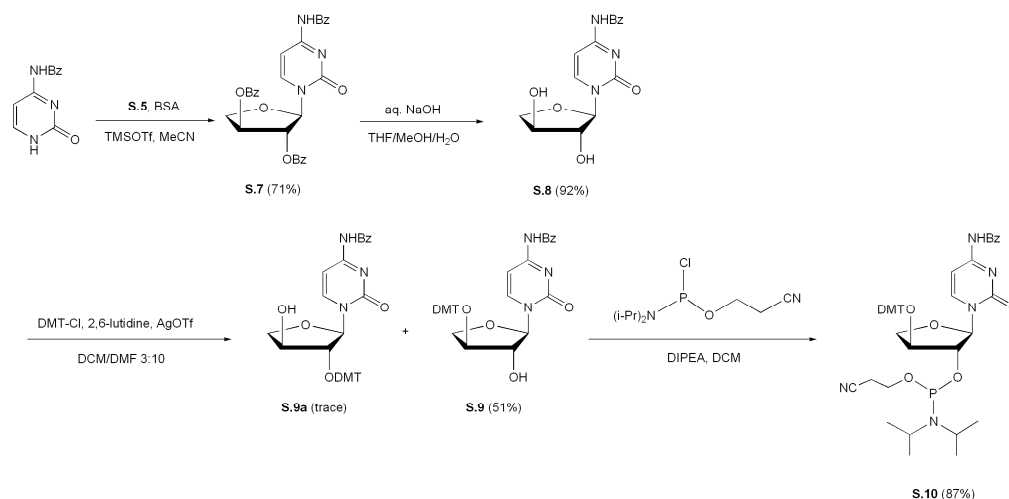

Figure 3.3. Preparation of **S.10**. Abbreviations: BSA, *N,O*-bis(trimethyl-silyl)acetamide; TMSOTf, trimethylsilyl trifluoromethanesulfonate; MeCN, acetonitrile; NaOH, sodium hydroxide; THF, tetrahydrofuran; MeOH, methanol; DMT-Cl, 4,4'-dimethoxytritylchloride; AgOTf, silver trifluoromethanesulfonate; DCM, dichloromethane; DMF, dimethylformamide; DIPEA, diisopropylethylamine.

Materials

N^4-benzoylcytosine (C^{Bz})

Sodium hydroxide (NaOH)

0.1 M aqueous hydrochloric acid solution (0.1 M HCl)

2,6-lutidine

N,O-bis(trimethylsilyl)acetamide (BSA)

4,4'-Dimethoxytriphenylmethyl chloride (DMT-Cl)

Trimethylsilyl trifluoromethanesulfonate (TMSOTf)

Silver trifluoromethanesulfonate (AgOTf)

Acetonitrile, anhydrous (MeCN)

Tetrahydrofuran (THF)

Ultra pure water (H$_2$O)

Dichloromethane (DCM)

Ethyl acetate (EtOAc)

Hexanes

Dichloromethane (DCM)

N,N-dimethylformamide (DMF)

Triethylamine (Et$_3$N)

Methanol (MeOH)

N,N-Diisopropylethylamine (DIPEA)

Chloro(2-cyanoethoxy)(diisopropylamino)phosphine

Saturated aqueous sodium bicarbonate solution (sat. aq. NaHCO$_3$)

Brine (sat. aq. NaCl)

Magnesium sulfate (MgSO$_4$)

Sodium sulfate (Na$_2$SO$_4$)

Argon gas

100- and 250-ml round-bottom flask

Magnetic stir plate and stir bar

Gas balloon

Rotary evaporator equipped with a vacuum pump

Thin layer chromatography (TLC) plate, EMD silica gel 60 F$_{254}$

Silica gel (60 Å, 230-400 mesh) (Whatman Inc.)

250-ml separatory funnel

Büchner funnel

Allihn condenser

Filter paper

UV lamp, 254 nm

Additional reagents and equipment for thin layer chromatography (TLC) and

column chromatography

Prepare S.7

1. Dissolve 2.4 g (6.5 mmol) of **S.5** in 19 ml of anhydrous MeCN in a 100-ml

   round bottom flask. To this solution, add 4 ml (16.2 mmol) of BSA and

   1.45 g (6.8 mmol) of N^4-benzoylcytosine.

2. Attach a condenser to the flask and stir the mixture at 70 °C for 1 hr.

3. Add 3.5 ml (19.4 mmol) of TMSOTf to the mixture from step 2 and stir

   the reaction at 70 °C for an addition *ca.* 1 hr until S.5 is completely

   consumed.

4. Cool the reaction to 24 °C, dilute the mixture with 60 ml of AcOEt and

   pour it into 76 ml of sat. NaHCO$_3$.

5. Stir the mixture for 15 min, isolate the product by vacuum filtration and

   wash the product with cold EtOAc.

6. Wash the organic phase of the filtrate with brine and dry it over MgSO$_4$.

   Evaporate the organic solvent under reduced pressure.

7. Purify the residue from step 6 by column chromatography. Elute the product from SiO_2 using a step-wise gradient of hexanes/EtOAc 4:1 to DCM/EtOAc 3:1 (v/v).

The resulting product, N^4-benzoyl-1-(2',3'-di-O-benzoyl-α-L-threo-furanosyl)cytosine (**S.7**), is obtained in a combined 71% yield (2.4 g, 4.6 mmol) as a white solid. TLC (DCM/EtOAc 1:1): R_f 0.48. ^{1}H-NMR (400 MHz, CDCl$_3$): δ=4.57 (d, J=11.2, 1H), 4.61 (dd, J=11.2, 3.6, 1H), 5.61 (d, J=3.2, 1H), 5.84 (s, 1H), 6.30 (s, 1H), 7.40-7.66 (m, 10H), 7.87 (m, 2H), 7.92 (d, J=7.6, 1H), 8.03 (d, J=7.2, 1H), 8.09 (m, 2H), 8.75 (br, s, 1H).

Prepare S.8

1. Suspend 2.6 g (5.1 mmol) of **S.7** in 83 ml of THF/MeOH/H$_2$O 5:4:1 and cool the mixture to 0 °C in a water-ice bath.

2. Add 8.3 ml of 2N NaOH and stir the reaction at 0 °C for 30 min.

3. Neutralize the mixture from step 2 with 2N HCl and evaporate the solvent to dryness under reduced pressure.

4. Re-suspend the residue in MeOH and co-evaporate it with toluene.

5. Resuspend the residue in MeOH, adsorb the mixture onto silica gel (sample/silica gel ratio 1:1.2, w/w)

6. Purify the residue by column chromatography. Elute the product using a step-wise gradient of DCM/MeOH 25:1 to 10:1 (v/v).

The resulting product, N^4-benzoyl-1-(α-L-threofuranosyl)cytosine (**S.8**), is obtained in a 92% yield (1.5 g, 4.7 mmol). TLC (DCM/MeOH 25:1): R_f 0.02.

^{1}H-NMR (400 MHz, DMSO-d$_6$): δ=4.03 (s, 1H), 4.08 (s, 1H), 4.19 (m, 2H), 5.66 (s, 1H), 7.30 (d, J=7.2, 1H), 7.52 (m, 2H), 7.63 (m, 1H), 8.00 (m, 2H), 8.11 (d, J=7.6, 1H).

Prepare S.9

1. Dissolve 1.5 g (4.7 mmol) of **S.8** in 9 ml of anhydrous DCM and 30 ml of anhydrous DMF.

2. To the solution from step 1, add 2.2 g (6.3 mmol) of DMT-Cl, 2.8 ml (24 mmol) of 2,6-lutidine and 1.44 g (0.6 mmol) of AgOTf.

3. After 5 hr of stirring at r. t., add another 0.15 g (0.45 mmol) of DMT-Cl and 0.14 g (0.6 mmol) of AgOTf and stir the reaction at r. t. overnight.

4. Quench the reaction with 6 ml of MeOH. Then add 1.5 ml of water and dilute the mixture with 150 ml of DCM.

5. Filter the mixture and wash the residue with DCM.

6. Wash the combined filtrate sequentially with ice-cold 0.1M aq. HCl solution (2x), water (2x), sat. aq. NaHCO$_3$ and brine.

7. Dry the solution over Na$_2$SO$_4$, and remove the organic solvent under reduced pressure.

8. Purify the residue by column chromatography. Deactivate the column with DCM/Et$_3$N 97:3 and elute the column using a step-wise gradient of 99:1 (v/v) DCM/Et$_3$N to DCM/MeOH/Et$_3$N 98:1:1.

The resulting product, N^4-benzoyl-1-{3'-O-[(4",4"'-dimethoxy-triphenyl)]-α-L-threofuranosyl}cytosine (**S.9**), is obtained in a 51.3% yield (1.5 g, 2.4 mmol). TLC (DCM/MeOH 16:1): R_f 0.26. ^{1}H-NMR (400 MHz, CDCl$_3$): δ=3.37 (d, J=9.6, 1H), 3.62 (dd, J=10.0, 3.2, 1H), 3.77 (2s, 6H), 3.95 (br, s, 1H), 4.26 (m, 2H), 5.64 (s, 1H), 6.80-6.84 (m, 4H), 7.20-7.36 (m, 9H), 7.53-7.66 (m, 4H), 7.91 (1H), 7.92 (d, J=7.6, 1H), 8.05 (d, J=7.6, 1H), 8.76 (br, s, 1H).

Prepare S.10

1. Dissolve 400 mg (0.68 mmol) of **S.9** in 3.2 ml of anhydrous DCM and 0.67 ml (3.8 mmol) of DIPEA in a 50 ml round bottom flask at r. t. under an argon atmosphere.

2. To the solution from step 1, add 180 μl (0.68 mmol) of chloro(2-cyanoethoxy)(diisopropylamino)phosphine in a dropwise manner.

3. Stir the reaction at r. t. for 1 hr. Then add an extra 20 μl (0.07 mmol) of chloro(2-cyanoethoxy)(diisopropylamino)phosphine to the reaction mixture.

4. Stir the reaction at r. t. for 18 hr.

5. Quench the reaction with 1 ml of MeOH.

6. Remove the solvent to dryness under reduced pressure and purify the residue by column chromatography. Deactivate the column with DCM/Et$_3$N 97:3, load the sample with benzene, and elute the product using hexanes/EtOAc/Et$_3$N 59:39:2.

The resulting product, N^4-benzoyl-1-{2'-O-[(2-cyanoethoxy)-(diisopropylamino)phosphino]-3'-O-[(4",4"'-dimethoxytriphenyl)]-α-L-threo-furanosyl}cytosine (**S.10**), is obtained in an 87% yield (466 mg, 0.59 mmol) as a mixture of diastereomers. TLC (DCM/MeOH 50:1): R_f 0.11, 0.17. ^{1}H-NMR (400 MHz, CDCl$_3$): δ=1.06-1.22 (m, 24H), 2.33, 2.68 (2m, 4H), 3.13 (d, J=10, 1H), 3.22 (d, J=9.6, 1H), 3.46-3.52 (m, 4H), 3.68, 3.69 (2s, 12H), 3.78-3.83 (m, 4H), 3.98-4.06 (m, 4H), 4.17 (d, J=3.6, 1H), 4.47 (d, J=8.4, 1H), 4.65 (d, J=9.2, 1H), 5.79 (s, 1H), 5.81 (s, 1H), 6.72-6.76 (m, 8H), 7.09-7.19 (m, 14H), 7.45-7.58 (m, 8H), 7.84-7.85 (m, 6H), 7.91 (d, J=7.6, 1H), 7.96 (d, J=7.2, 1H), 8.58 (br, 2H). ^{31}P-NMR (161 MHz, CDCl$_3$): δ=150.60, 152.58.

3.6 Basic Protocol 3

Synthesis of N^2,N^6-Dibenzoyl-9-{2'-O-[(2-cyanoethoxy)(diisopropyl-amino)-phosphino]-3'-O-[(4,4'-dimethoxytriphenyl)methyl]-α-L-threo-furanosyl}-2,6-diaminopurine

This protocol describes the synthesis of N^2,N^6-Dibenzoylated TNA diaminopurine nucleoside phosphoramidite **S.14** from **S.5** (Figure 3.4).

Materials

N^2,N^6-dibenzoyl-2,6-diaminopurine (D$^{(Bz)2}$)

Sodium hydroxide (NaOH)

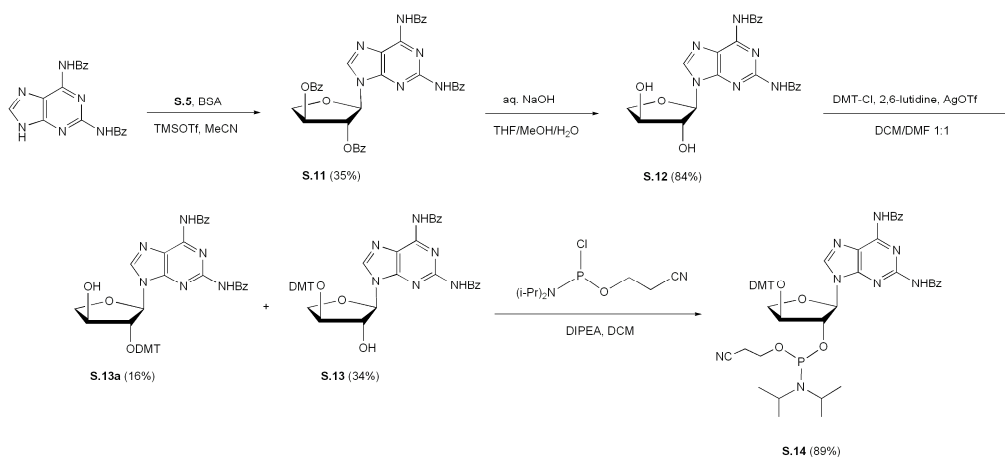0.1 M aqueous hydrochloric acid solution (0.1 M HCl)

Ammonium chloride (NH$_4$Cl)

Figure 3.4. Preparation of **S.14**. Abbreviations: BSA, N,O-bis(trimethylsilyl)-acetamide; TMSOTf, trimethylsilyl trifluoromethanesulfonate; MeCN, acetonitrile; NaOH, sodium hydroxide; THF, tetrahydrofuran; MeOH, methanol; DMT-Cl, 4,4'-dimethoxytritylchloride; AgOTf, silver trifluoromethanesulfonate; DCM, dichloromethane; DMF, dimethylformamide; DIPEA, diisopropylethylamine.

2,6-lutidine

N,O-bis(trimethylsilyl)acetamide (BSA)

4,4'-Dimethoxytriphenylmethyl chloride (DMT-Cl)

Trimethylsilyl trifluoromethanesulfonate (TMSOTf)

Silver trifluoromethanesulfonate (AgOTf)

Acetonitrile, anhydrous (MeCN)

Tetrahydrofuran (THF)

Ultra pure water (H$_2$O)

Dichloromethane (DCM)

Ethyl acetate (EtOAc)

Hexanes

Dichloromethane (DCM)

N,N-dimethylformamide (DMF)

Triethylamine (Et$_3$N)

Methanol (MeOH)

N,N-Diisopropylethylamine (DIPEA)

Chloro(2-cyanoethoxy)(diisopropylamino)phosphine

Benzene

Saturated aqueous sodium bicarbonate solution (sat. aq. NaHCO$_3$)

Brine (sat. aq. NaCl)

Magnesium sulfate (MgSO$_4$)

Sodium sulfate (Na$_2$SO$_4$)

Argon gas

Activated molecular sieves

100- and 250-ml round-bottom flask

Magnetic stir plate and stir bar

Gas balloon

Rotary evaporator equipped with a vacuum pump

Thin layer chromatography (TLC) plate, EMD silica gel 60 F$_{254}$

Silica gel (60 Å, 230-400 mesh) (Whatman Inc.)

250-ml separatory funnel

Büchner funnel

Allihn condenser

Filter paper

UV lamp, 254 nm

Additional reagents and equipment for thin layer chromatography (TLC) and

column chromatography

Prepare S.11

1. Suspend 3.2 g (8.9 mmol) of N^2,N^6-dibenzoyl-2,6-diaminopurine in 30 ml

    of anhydrous MeCN. Stir the mixture under an argon atmosphere at 65 °C.

2. To the suspension from step 1, add 5.9 ml (23.8 mmol) of BSA. Stir the

    mixture at 65 °C for an additional 30 min until a clear solution is formed.

3. Add 2.86 g (7.7 mmol) of **S.5** in 25.5 ml of anhydrous MeCN to the clear

    solution from step 2. Then add 2 ml (10 mmol) of TMSOTf to the solution.

4. Stir the reaction at 65 °C for 3 hr, and then cool it down to r. t.

5. Dilute with 190 ml of EtOAc and wash the reaction mixture with 250 ml

    of sat. aq. $NaHCO_3$.

6. Separate and filter the organic phase, wash it with sat. aq. $NaHCO_3$, then

    brine, and dry it over $MgSO_4$.

7. Remove the organic solvent under reduced pressure and purify the residue

    by column chromatography. Elute the column using a step-wise gradient of

    hexanes/EtOAc 3:2 to 1:3 (v/v).

    The resulting product, N^2,N^6-Dibenzoyl-9-(2',3'-di-O-benzoyl-α-L-threo-

furanosyl)-2,6-diaminopurine (**S.11**), is obtained in a 35% yield (1.8 g, 2.7

mmol) as a white foam. TLC (hexanes/EtOAc 1:3): R_f 0.48. ^{1}H-NMR (400

MHz, DMSO-d_6): δ=4.64 (dd, J=10.0, 6.4, 1H), 4.82 (dd, J=9.6, 4.8, 1H), 5.89

(m, 1H), 6.49 (m, 2H), 7.33-8.10 (m, 20H), 8.61 (s, 1H), 10.99 (s, 1H), 11.27
(s, 1H).

Prepare S.12

1. Dissolve 1.9 g (2.8 mmol) of **S.11** in 160 ml of ice-cold THF/MeOH/H$_2$O
   5:4:1. To this solution slowly add 11 ml (22.1 mmol) of 2N aq. NaOH over
   10 min.

2. Stir the reaction at 0 °C for an additional 5 min.

3. Quench the reaction with 1.25 (23.5 mmol) of solid NH$_4$Cl. Keep stirring
   the reaction at 0 °C for 30 min.

4. Remove the solvent under reduced pressure and co-evaporate the residue
   twice with MeOH.

5. Adsorb the residue from step 4 onto silica gel. Purify the product by
   column chromatography. Elute the column using a step-wise gradient of
   DCM/MeOH 100:1 to 40:1 (v/v).

   The resulting product, N^2,N^6-Dibenzoyl-9-(α-L-threofuranosyl)-

2,6-diaminopurine (**S.12**), is obtained in an 84% yield (1.1 g, 2.4 mmol) as a

white powder. TLC (DCM/MeOH 20:1): R$_f$ 0.26. ^{1}H-NMR (400 MHz,

DMSO-d$_6$): δ=4.07 (dd, J=9.2, 2.4, 1H), 4.14 (dd, J=9.2, 3.6, 1H), 4.18 (m,

1H), 4.41 (m, 1H), 5.61 (br, s, 1H), 5.90 (d, J=4.0, 1H), 5.95 (br, s, 1H),

7.46-7.59 (m, 6H), 7.98-8.05 (m, 4H), 8.05 (s, 1H), 11.00 (br, s, 2H).

Prepare S.13

1. Dissolve 3.44 g (7.5 mmol) of **S.12** in 63 ml of DCM/DMF 1:1 in a 500 ml round bottom flask at r. t. under an argon atmosphere and a water bath.

2. Add 3.46 g (10.2 mmol) of DMT-Cl to the solution from step 1 with stirring, then add 5.1 ml (44.2 mmol) of 2,6-lutidine.

3. Add *ca*. 3.5 g of activated molecular sieves, and then add 2.94 g (11.4 mmol) of AgOTf over a period of 10 min in 4 portions at r. t.

4. After stirring the reaction for an additional 4 hr, add 457 mg (1.4 mmol) of DMT-Cl and 330 mg (1.3 mmol) of AgOTf to the mixture from step 3.

5. Stir the reaction in the water bath at r. t. for an additional 20 hr.

6. Quench the reaction by 2.5 ml of MeOH, dilute with 290 ml of DCM, and then filter the resulting mixture.

7. Cool the filtrate to 0 °C and wash it with 290 ml of cold 0.1 M aq. HCl (2x), with sat. aq. $NaHCO_3$, and brine.

8. Dry the organic phase over Na_2SO_4, and then evaporate the solvent to dryness under reduced pressure.

9. Purify the residue by column chromatography. Deactivate the column with DCM/Et3N 97:3 and elute the column using a step-wise gradient of hexanes/benzene/DCM/EtOAc/Et_3N/MeOH 10:10: 2.5:2.5:0.2:0.1 to 10:10: 2.5:2.5:0.2:0.4 and finally DCM/ Et_3N 35:1 plus 1% Et_3N.

The resulting product, N^2,N^6-Dibenzoyl-9-{3'-O-[(4",4'"-dimethoxy-triphenyl)methyl]-α-L-threofuranosyl}-2,6-diaminopurine (**S.13**), and its 2'-O-regioisomer N^2,N^6-Dibenzoyl-9-{2'-O-[(4",4'"-dimethoxytriphenyl)-methyl]-α-L-threofuranosyl}-2,6-diaminopurine (**S.13a**), are obtained in a 33.7% (2.0 g, 2.5 mmol) and 16% (0.9 g, 1.2 mmol) yield, respectively. TLC (benzene/EtOAc/DCM/MeOH 8:1:1:0.2): R_f 0.40 (**S.13**), 0.35 (**S.13a**).

^{1}H-NMR **S.13** (400 MHz, CDCl$_3$): δ=3.05 (dd, J=9.4, 6.2, 1H), 3.36 (dd, J=9.2, 6.2, 1H), 3.77 (2s, 6H), 4.44 (m, 1H), 4.76 (m, 1H), 5.79 (d, J=4.4, 1H), 6.80-6.85 (m, 4H), 7.18-7.64 (m, 15H), 7.99-8.06 (m, 5H), 9.43 (br, s, 1H), 9.64 (br, s, 1H). **S.13a** (500 MHz, DMSO-d$_6$): δ=3.62 (2s, 6H), 3.76 (m, 1H), 3.97 (d, J=7.6, 1H), 4.08 (dd, J=7.6, 2.4, 1H), 4.44 (s, 1H), 5.36 (d, J=2.8, 1H), 5.92 (s, 1H), 6.70-6.77 (m, 4H), 7.14-7.65 (m, 16H), 7.98-8.20 (m, 4H), 11.01 (s, 1H), 11.19 (s, 1H).

Prepare S.14

1. Dissolve 400 mg (0.52 mmol) of **S.13** in 3.8 ml of anhydrous DCM and 0.34 ml of DIPEA in a 50 ml round bottom flask at r. t. under an argon atmosphere.

2. To the solution from step 1, add 840 μl (3.7 mmol) of chloro(2-cyanoethoxy)(diisopropylamino)phosphine in a dropwise manner.

3. Stir the reaction at r. t. for 1 hr. Then add an extra 84 μl (0.4 mmol) of chloro(2-cyanoethoxy)(diisopropylamino)phosphine to the reaction mixture.

4. Stir the reaction at r. t. for an additional 1 hr.

5. Remove the solvent to dryness under reduced pressure and purify the

   residue by column chromatography. Deactivate the column with

   DCM/Et$_3$N 97:3, load the sample with benzene, and elute the product

   using benzene/DCM/EtOAc/Et$_3$N/MeOH 10:1:1:0.2:0.2.

   The resulting product, N^2,N^6-Dibenzoyl-9-{2'-O-[(2-cyanoethoxy)-

(diisopropylamino)phosphino]-3'-O-[(4",4'''-dimethoxytriphenyl)methyl]-α-ʟ-

threofuranosyl}-2,6-diaminopurine (**S.14**) is obtained in a 89% (450 mg, 0.46

mmol) yield as a mixture of two diastereoisomers. TLC (DCM/MeOH 50:1):

R$_f$ 0.21. ^{1}H-NMR (400 MHz, CDCl$_3$): δ=1.08-1.30 (m, 24H), 2.39 (m, 2H),

2.60 (m, 2H), 3.36-3.39 (m, 1H), 3.54-3.90 (m, 8H), 3.72 (2s, 12H), 4.27 (m,

2H), 4.35 (m, 1H), 4.71 (d, J=8.0, 2H), 5.16 (d, J=9.6, 1H), 6.12 (s, 1H), 6.22

(s, 2H),6.78 (m, 8H), 7.12-7.35 (m, 17H), 7.46-7.50 (m, 13H), 7.94-8.04 (m,

8H), 8.26, 8.31 (2s, 2H), 8.90, 9.11, 9.23 (3br, 4H). ^{31}P-NMR (161 MHz,

CDCl$_3$): δ=150.5, 151.9. MOLDI-TOF MS (pos. 4-hydroxybenzylidene-

malononitrile): 985.63 (M+Na$^+$).

3.7 Basic Protocol 4

   Synthesis of N^2-Acetyl-O^6-diphenylcarbamoyl-9-{2'-O-[(2-cyanoethoxy)-

(diisopropylamino)phosphino]-3'-O-[(4,4'-dimethoxytriphenyl)methyl]-α-ʟ-th

reofuranosyl}guanine

This protocol describes the synthesis of base-protected TNA guanosine nucleoside phosphoramidite **S.18** from **S.5** (Figure 3.5).
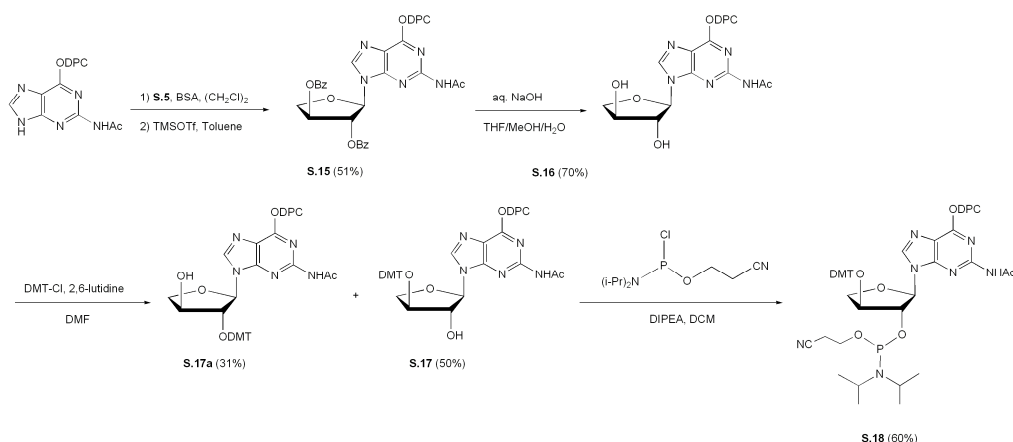

Figure 3.5. Preparation of **S.18**. Abbreviations: BSA, *N,O*-bis(trimethylsilyl)-acetamide; TMSOTf, trimethylsilyl trifluoromethanesulfonate; NaOH, sodium hydroxide; THF, tetrahydrofuran; MeOH, methanol; DMT-Cl, 4,4'-dimethoxytritylchloride; AgOTf, silver trifluoromethanesulfonate; DCM, dichloromethane; DMF, dimethylformamide; DIPEA, diisopropylethylamine.

Materials

N^2-acetyl-O^6-diphenylcarbamoylguanine (G^{PAC})

Sodium hydroxide (NaOH)

Ammonium chloride (NH_4Cl)

2,6-lutidine

N,O-bis(trimethylsilyl)acetamide (BSA)

4,4'-Dimethoxytriphenylmethyl chloride (DMT-Cl)

Trimethylsilyl trifluoromethanesulfonate (TMSOTf)

Silver trifluoromethanesulfonate (AgOTf)

Acetonitrile, anhydrous (MeCN)

Tetrahydrofuran (THF)

Ultra pure water (H$_2$O)

Dichloromethane (DCM)

1,2-Dichloroethane, anhydrous ((CH$_2$Cl)$_2$),

Toluene, anhydrous

Ethyl acetate (EtOAc)

Hexanes

Dichloromethane (DCM)

N,N-dimethylformamide (DMF)

Triethylamine (Et$_3$N)

Methanol (MeOH)

N,N-Diisopropylethylamine (DIPEA)

Chloro(2-cyanoethoxy)(diisopropylamino)phosphine

Saturated aqueous sodium bicarbonate solution (sat. aq. NaHCO$_3$)

Brine (sat. aq. NaCl)

Magnesium sulfate (MgSO$_4$)

Sodium sulfate (Na$_2$SO$_4$)

Argon gas

100- and 250-ml round-bottom flask

Magnetic stir plate and stir bar

Gas balloon

Rotary evaporator equipped with a vacuum pump

Thin layer chromatography (TLC) plate, EMD silica gel 60 F$_{254}$

Silica gel (60 Å, 230-400 mesh) (Whatman Inc.)

250-ml separatory funnel

Büchner funnel

Allihn condenser

Filter paper

UV lamp, 254 nm

Additional reagents and equipment for thin layer chromatography (TLC) and

column chromatography

Prepare S.15

1. Suspend 3.2 g (8.3 mmol) of N^2-acetyl-O^6-diphenylcarbamoylguanine in

   77 ml of anhydrous 1,2-Dichloroethane. To this suspension add 4ml (16.8

   mmol) of BSA. Stir the mixture under an argon atmosphere at 70 °C for

   1hr.

2. Evaporate the solvent to dryness and dissolve the residue in 35 ml of

   anhydrous toluene. To this solution, add 4 ml (16.8 mmol) of TMSOTf.

3. Add 2.8 g (7.6 mmol) of **S.5** in 35 ml of anhydrous toluene to the solution

   from step 2.

4. Stir the reaction at 70 °C for an additional 2.5 hr, and then cool it down to

   r.t.

5. Dilute with 70 ml of EtOAc and wash the reaction mixture with 100 ml of

   sat. aq. $NaHCO_3$, then brine, and dry it over $MgSO_4$.

6. Remove the organic solvent under reduced pressure and purify the residue by column chromatography. Elute the column using a step-wise gradient of DCM/EtOAc 18:1 plus 1% Et_3N to DCM/EtOAc/MeOH 13:1:0.05 plus 1% Et_3N (v/v).

The resulting product, N^2-acetyl-O^6-diphenylcarbamoyl-9-(2',3'-di-O-benzoyl-α-L-threofuranosyl)guanine (**S.15**), is obtained in a 51% yield (2.74 g, 4.2 mmol) as a colorless foam. TLC (hexanes/EtOAc 1:3): R_f 0.27. ^{1}H-NMR (400 MHz, $CDCl_3$): δ=2.48 (s, 3H), 4.57 (dd, J=11.0, 1.8, 1H), 4.66 (dd, J=11.0, 4.6, 1H), 5.74 (dd, J=2.8, 1.6, 1H), 6.21 (s, 1H), 6.32 (d, J=1.6, 1H), 7.34-8.09 (m, 20H), 8.23 (s, 1H).

Prepare S.16

1. Dissolve 2.74 g (3.9 mmol) of **S.15** in 230 ml of ice-cold THF/MeOH/H_2O 5:4:1. To this solution, slowly add 16 ml of 2N aq. NaOH over 5 min under an argon atmosphere.

2. Stir the reaction at 0 °C for an additional 15 min.

3. Quench the reaction with 1.8 (34 mmol) of solid NH_4Cl. Keep stirring the reaction at 0 °C for 30 min.

4. Remove the solvent under reduced pressure.

5. Adsorb the residue from step 4 onto silica gel. Purify the product by column chromatography. Elute the column using DCM/MeOH 15:1 (v/v).

The resulting product, N^2-acetyl-O^6-diphenylcarbamoyl-9-(α-L-threo-furanosyl)guanine (**S.16**), is obtained in a 70% yield (1.34 g, 2.7 mmol) as a white powder. TLC (DCM/MeOH 5:1): R_f 0.68. ^{1}H-NMR (400 MHz, DMSO-d_6): δ=4.08-4.16 (m, 3H), 4.42 (m, 1H), 5.53 (d, J=3.2, 1H), 5.88 (d, J=4.4, 1H), 5.92 (d, J=1.6, 1H), 7.29-7.48 (m, 10H), 8.52 (s, 1H), 10.73 (br, s, 1H).

Prepare S.17

1. Dissolve 1.0 g (2.0 mmol) of **S.16** in 30 ml of anhydrous DMF in a 100 ml round bottom flask at r. t. under an argon atmosphere.

2. Add 0.94 ml (8.2 mmol) of 2,6-lutidine and 1.4 g (4.1 mmol) of DMT-Cl to the solution from step 1 with stirring.

3. Stir the reaction at r. t. for 24 hr.

4. Add an additional 350 mg (1.0 mmol) of DMT-Cl and stir the reaction at r.t. for an additional 6 hr.

5. Quench the reaction by 8.5 ml of MeOH, dilute with 25 ml of DCM, and wash the organic solvent with sat. aq. $NaHCO_3$ and brine.

6. Dry the organic phase over $MgSO_4$, and evaporate the solvent to dryness under reduced pressure.

7. Purify the residue by column chromatography. Deactivate the column with DCM/Et3N 97:3 and elute the column using a step-wise gradient of DCM/Et_3N 99:1 to DCM/MeOH/Et_3N 98:1:1 (v/v).

The resulting product, N^2-acetyl-O^6-diphenylcarbamoyl-9-{3'-O-[(4",4"'-dimethoxytriphenyl)methyl]-α-L-threofuranosyl}guanine (**S.17**), and its 2'-O-regioisomer N^2-acetyl-O^6-diphenylcarbamoyl-9-{3'-O-[(4",4"'-dimethoxytriphenyl)methyl]-α-L-threofuranosyl}guanine (**S.17a**), are obtained in a 50% (0.79 g, 1 mmol) and 31% (0.5 g, 0.6 mmol), respectively. TLC (DCM/MeOH 25:1) R_f 0.36 (**S.17**), 0.52 (**S.17a**). ^{1}H-NMR **S.17** (500 MHz, DMSO-d_6): δ=2.20 (s, 3H), 3.44 (dd, J=7.4, 4.4, 1H), 3.58 (dd, J=7.6, 2.8, 1H), 3.68 (2s, 6H), 4.13 (m, 1H), 4.44 (m, 1H), 5.78 (m, 2H), 6.79-6.81 (m, 4H), 7.10-7.48 (m, 19H), 8.49 (s, 1H), 10.68 (s, 1H). **S.17a** (500 MHz, DMSO-d_6): δ=2.16 (s, 3H), 3.52 (s, 3H), 3.56 (s, 3H), 3.92 (dd, J=8.0, 2.8, 1H), 4.12 (m, 2H), 4.53 (s, 1H), 5.42 (d, J=7.2, 1H), 5.62 (s, 1H), 6.61-6.70 (m, 4H), 7.13-7.47 (m, 19H), 8.17 (s, 1H), 10.73 (s, 1H).

Prepare S.18

1. Dissolve 200 mg (0.25 mmol) of **S.17** and 0.27 ml (2.5 mmol) of DIPEA in 6.2 ml of anhydrous DCM.

2. To the solution from step 1, add 140 μl (0.62 mmol) of chloro(2-cyanoethoxy)(diisopropylamino)phosphine in a dropwise manner.

3. Stir the reaction at r. t. for 45 min. Then add an extra 15 μl (0.07 mmol) of chloro(2-cyanoethoxy)(diisopropylamino)phosphine to the reaction mixture.

4. Stir the reaction at r. t. for an additional 1hr.

5. Quench the reaction with 1 ml of MeOH.

6. Remove the solvent to dryness under reduced pressure and purify the residue by column chromatography. Deactivate the column with DCM/Et$_3$N 97:3 and elute the product using DCM/Et$_3$N 99:1 to give a syrup.

7. Dissolve the syrup from step 6 in 2 ml of DCM and add this solution slowly into 90 ml of rapidly stirred pentane. Collect the white precipitate by filtration.

The resulting product, N^2-acetyl-O^6-diphenylcarbamoyl-9-{2'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-3'-O-[(4",4'''-dimethoxytriphenyl)methyl]-α-L-threofuranosyl}guanine (**S.18**) is obtained in a 60% (150mg, 0.15 mmol) as a mixture of diastereomers. TLC (DCM/MeOH 50:1): R$_f$ 0.24, 0.33. ^{1}H-NMR (400 MHz, CDCl$_3$): δ=1.06-1.15 (m, 24H), 2.12-2.44 (m, 4H), 2.51, 2.60 (2s, 6H), 3.51-3.68 (m, 10H), 3.69-3.72 (4s, 12H), 4.30, 4.42 (m, 2H), 4.70 (d, J=8.4, 1H), 4.79 (d, J=8.0, 1H), 5.96, 6.03 (2s, 2H), 6.74-7.44 (m, 46H), 7.90, 8.02 (2br, 2H), 8.24, 8.25 (2s, 2H). ^{31}P-NMR (161 MHz, CDCl$_3$): 150.99, 151.34.

3.8 Basic Protocol 5

Synthesis of 1-{2'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-3'-O-[(4,4'-dimethoxytriphenyl)methyl]-α-L-threofuranosyl}thymine

This protocol describes the synthesis of TNA thymidine nucleoside phosphoramidite **S.22** from **S.5** (Figure 3.6).
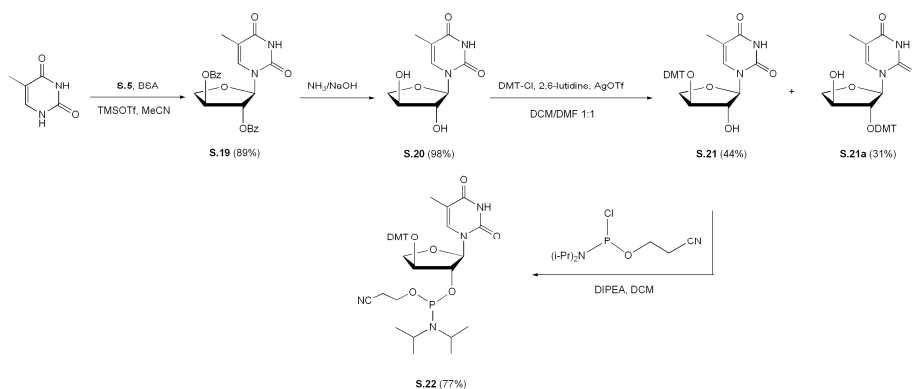
96

Figure 3.6. Preparation of **S.22**. Abbreviations: BSA, *N,O*-bis(trimethyl-silyl)acetamide; TMSOTf, trimethylsilyl trifluoromethanesulfonate; MeCN, acetonitrile; NaOH, sodium hydroxide; THF, tetrahydrofuran; MeOH, methanol; DMT-Cl, 4,4'-dimethoxytritylchloride; AgOTf, silver trifluoromethanesulfonate; DCM, dichloromethane; DMF, dimethylformamide; DIPEA, diisopropylethylamine.

Materials

Thymine (T)

2 M NH$_3$ in MeOH

0.1 M aqueous HCl solution (0.1 M HCl)

2,6-lutidine

N,O-bis(trimethylsilyl)acetamide (BSA)

4,4'-Dimethoxytriphenylmethyl chloride (DMT-Cl)

Trimethylsilyl trifluoromethanesulfonate (TMSOTf)

Silver trifluoromethanesulfonate (AgOTf)

Acetonitrile, anhydrous (MeCN)

Tetrahydrofuran (THF)

Ultra pure water (H$_2$O)

Dichloromethane (DCM)

Ethyl acetate (EtOAc)

Hexanes

Dichloromethane (DCM)

N,N-dimethylformamide (DMF)

Triethylamine (Et$_3$N)

Methanol (MeOH)

N,N-Diisopropylethylamine (DIPEA)

Chloro(2-cyanoethoxy)(diisopropylamino)phosphine

Saturated aqueous sodium bicarbonate solution (sat. aq. NaHCO$_3$)

Brine (sat. aq. NaCl)

Magnesium sulfate (MgSO$_4$)

Sodium sulfate (Na$_2$SO$_4$)

Argon gas

100- and 250-ml round-bottom flask

Magnetic stir plate and stir bar

Gas balloon

Rotary evaporator equipped with a vacuum pump

Thin layer chromatography (TLC) plate, EMD silica gel 60 F$_{254}$

Silica gel (60 Å, 230-400 mesh) (Whatman Inc.)

250-ml separatory funnel

Büchner funnel

Allihn condenser

Filter paper

UV lamp, 254 nm

Additional reagents and equipment for thin layer chromatography (TLC) and column chromatography

Prepare S.19

1. Suspend 4.4 g (12 mmol) of **S.5** and 1.5 g (12 mmol) of thymine in 35 ml of anhydrous MeCN in a 100-ml round bottom flask.

2. Attach a condenser to the flask and heat the mixture to 60 °C. Treat the hot reaction mixture with 5.8 ml (24 mmol) of BSA. Stir the reaction at 60 °C for an additional 1 hr under an argon atmosphere.

3. Add 6.4 ml (36 mmol) of TMSOTf to the mixture from step 2 and stir the reaction at 60 °C for an addition *ca.* 3 hr until S.5 is completely consumed.

4. Cool the reaction to 24 °C, pour the reaction mixture into an ice-cold stirred mixture of 120 ml of sat. aq. $NaHCO_3$ and 150 ml EtOAc.

5. Separate and wash the organic phase with brine and dry it over $MgSO_4$. Evaporate the organic solvent under reduced pressure.

6. Purify the residue from step 5 by column chromatography. Elute the product from SiO_2 using DCM/EtOAc 7:1 (v/v).

    The resulting product, 1-(2',3'-di-*O*-benzoyl-α-L-threofuranosyl)thymine (**S.19**), is obtained in an 89% yield (2.4 g, 4.6 mmol) as a colorless foam. TLC (hexanes/EtOAc 1:1): R_f 0.41. ^{1}H-NMR (400 MHz, $CDCl_3$): δ=1.88 (s, 3H),

4.48 (m, 2H), 5.60 (t, 1H), 5.67 (s, 1H), 6.21 (m, 1H), 7.27-8.09 (m, 11H),

8.25 (br, s, 1H).

Prepare S.20

1. Suspend 4.67 g (11.1 mmol) of **S.19** in 220 ml of 2 M NH_3 in MeOH and

   stir the reaction at r. t. for *ca*. 18 hr.

2. Concentrate the solution under reduced pressure and co-evaporate the

   residue three times with water.

3. Dissolve the residue in water and wash the solution with DCM.

4. Remove the solvent under reduced pressure to give the product.

   The resulting product, 1-(α-L-threofuranosyl)thymine (**S.20**), is obtained

in a 98% yield (2.5 g, 10.9 mmol). ^{1}H-NMR (400 MHz, DMSO-d_6): δ=1.74 (s,

3H), 3.32 (br, s, 2H), 4.00 (m, 4H), 5.64 (m, 1H), 7.56 (m, 1H), 11.27 (br, s,

1H).

Prepare S.21

1. Dissolve 2.5 g (11 mmol) of **S.20** in 150 ml of anhydrous DCM/DMF 1:1

   in a 500 ml round bottom flask at r. t. under an argon atmosphere.

2. Add 4.4 g (13 mmol) of DMT-Cl to the solution from step 1 with stirring,

   then add 6.4 ml (55 mmol) of 2,6-lutidine and 3.1 g (12 mmol) of AgOTf.

3. Stir the reaction at r. t. overnight under an argon atmosphere.

4. Quench the reaction by 15.2 ml of MeOH, dilute with 500 ml of DCM,

   and then filter the resulting mixture.

5. Wash the filtrate with ice-cold 0.1 M aq. HCl (2x), water (2x), sat. aq. NaHCO$_3$ (2x), and brine.

6. Dry the organic phase over Na$_2$SO$_4$, and then evaporate the solvent to dryness under reduced pressure.

7. Purify the residue by column chromatography. Deactivate the column with DCM/Et$_3$N 97:3 and elute the column using a step-wise gradient of DCM/Et$_3$N 100:1 to DCM/MeOH/Et$_3$N 98:1:1.

   The resulting product, 1-{3'-*O*-[(4",4"'-dimethoxytriphenyl)methyl]-α-L-threofuranosyl}thymine (**S.21**), and its 2'-*O*-regioisomer 1-{2'-*O*-[(4",4"'-dimethoxytriphenyl)methyl]-α-L-threofuranosyl}thymine (**S.21a**), are obtained in a 44% (2.55 g, 4.8 mmol) and 31% (1.79 g, 3.4 mmol), respectively. TLC (DCM/MeOH 25:1): R$_f$ 0.33 (**S.21**), 0.43 (**S.21a**). ^{1}H-NMR **S.21** (400 MHz, DMSO-d$_6$): δ=1.68 (s, 3H), 3.56-3.57 (m, 3H), 3.60 (s, 3H), 3.73 (s, 3H), 3.99 (d, J=1.6, 1H), 5.46 (d, J=1.6, 1H), 5.58 (d, J=4.4, 1H), 6.85-6.88 (m, 4H), 7.20-7.50 (m, 10H), 11.28 (s, 1H). **S.21a** (400 MHz, DMSO-d$_6$): δ=1.66 (s, 3H), 3.32 (s, 3H), 3.34 (s, 3H), 3.71 (d, J=2.8, 1H), 3.82 (m, 2H), 3.92 (d, J=1.6, 1H), 5.10 (d, J=2.8, 1H), 5.91 (d, J=1.6, 1H), 6.80-7.20 (m, 4H), 7.21-7.39 (m, 10H), 11.33 (s, 1H).

Prepare S.22

1. Dissolve 330 mg (0.63 mmol) of **S.21** and 0.43 ml (2.5 mmol) of DIPEA in 4.0 ml of anhydrous DCM.

2. To the solution from step 1, add 150 µl (0.7 mmol) of chloro(2-cyanoethoxy)(diisopropylamino)phosphine in a dropwise manner.

3. Stir the reaction at r. t. for 1 hr. Then add an extra 15 µl (0.07 mmol) of chloro(2-cyanoethoxy)(diisopropylamino)phosphine to the reaction mixture.

4. Stir the reaction at r. t. overnight.

5. Remove the solvent to dryness under reduced pressure and purify the residue by column chromatography. Deactivate the column with DCM/Et$_3$N 97:3 and elute the product using a step-wise gradient of hexanes/EtOAc/Et$_3$N 80:18:2 to 50:48:2.

The resulting product, 1-{2'-*O*-[(2-cyanoethoxy) (diisopropylamino)-phosphino]-3'-*O*-[(4",4"'-dimethoxytriphenyl)methyl]-α-L-threofuranosyl}-thymine (**S.22**), is obtained in a 77% (400 mg, 0.48 mmol) as a mixture of diastereomers. TLC (DCM/MeOH 25:2): R$_f$ 0.26, 0.36. ^{1}H-NMR (300 MHz, CDCl$_3$): δ=1.03, 1.13 (m, 24H), 1.79 (s, 6H), 2.64 (m, 4H), 3.13 (m, 2H), 3.55-3.80 (m, 18H), 4.25 (d, J=3.3, 1H), 4.47 (d, J=9.3, 1H), 5.77 (br, 2H), 6.81-6.86 (m, 8H), 7.22-7.40 (m, 20H). ^{31}P-NMR (202 MHz, CDCl$_3$): 153.4, 154.4.

3.9 Basic Protocol 6

Synthesis of N^6-Benzoyl-9-{2'-O-[(2-cyanoethoxy)(diisopropylamino)-

phosphino]-3'-O-[(4,4'-dimethoxytriphenyl)methyl]-α-L-threofuranosyl}-

adenine

This protocol describes the synthesis of N^6-benzoylated TNA adenosine

nucleoside phosphoramidite **S.26** from **S.6** (Figure 3.7).
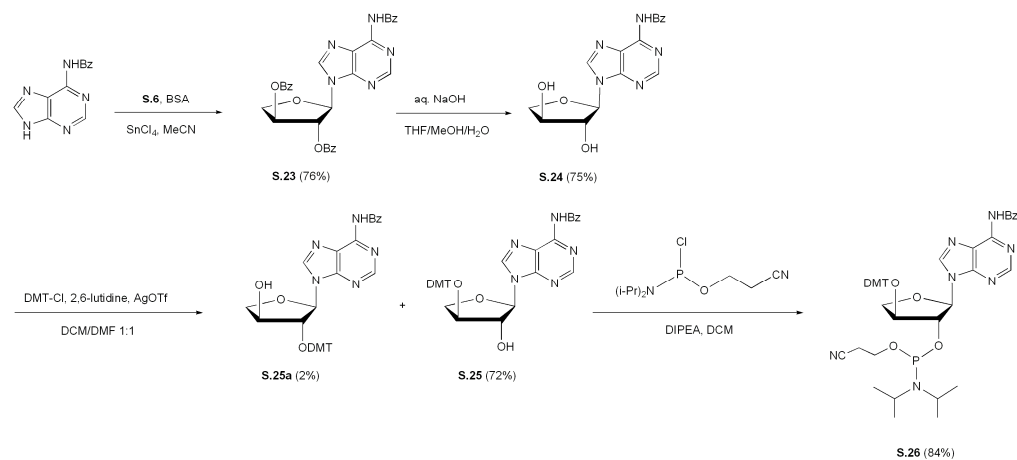
Figure 3.7. Preparation of **S.26**. Abbreviations: BSA, N,O-bis(trimethylsilyl)-
acetamide; SnCl$_4$, tin(IV) chloride; MeCN, acetonitrile; NaOH, sodium
hydroxide; THF, tetrahydrofuran; MeOH, methanol; DMT-Cl,
4,4'-dimethoxytritylchloride; AgOTf, silver trifluoromethanesulfonate; DCM,
dichloromethane; DMF, dimethylformamide; DIPEA, diisopropylethylamine.
Dichloromethane (DCM)

Materials

N^6-Benzoyladenine (A^{Bz})

Sodium hydroxide (NaOH)

2N aqueous HCl solution (2N HCl)

0.1 M aqueous HCl solution (0.1 M HCl)

2,6-lutidine

N,O-bis(trimethylsilyl)acetamide (BSA)

103

4,4'-Dimethoxytriphenylmethyl chloride (DMT-Cl)

Tin (IV) Chloride (SnCl$_4$)

Silver trifluoromethanesulfonate (AgOTf)

Acetonitrile, anhydrous (MeCN)

Tetrahydrofuran (THF)

Ultra pure water (H$_2$O)

Ethyl acetate (EtOAc)

Hexanes

Dichloromethane (DCM)

N,N-dimethylformamide (DMF)

Triethylamine (Et$_3$N)

Methanol (MeOH)

N,N-Diisopropylethylamine (DIPEA)

Chloro(2-cyanoethoxy)(diisopropylamino)phosphine

Saturated aqueous sodium bicarbonate solution (sat. aq. NaHCO$_3$)

Brine (sat. aq. NaCl)

Magnesium sulfate (MgSO$_4$)

Sodium sulfate (Na$_2$SO$_4$)

Argon gas

100- and 250-ml round-bottom flask

Magnetic stir plate and stir bar

Gas balloon

Rotary evaporator equipped with a vacuum pump

Thin layer chromatography (TLC) plate, EMD silica gel 60 F$_{254}$

Silica gel (60 Å, 230-400 mesh) (Whatman Inc.)

250-ml separatory funnel

Büchner funnel

Allihn condenser

Filter paper

UV lamp, 254 nm

Additional reagents and equipment for thin layer chromatography (TLC) and

column chromatography

Prepare S.23

1. Dissolve 1.33 g (3.1 mmol) of **S.6** in 8 ml of anhydrous MeCN in a 50 ml

   round bottom flask equipped with a stirring bar.

2. To the solution from step 1 add 830 mg (3.5 mmol) of N^6-benzoyladenine

   and 1.5 ml (6.2 mmol) of BSA.

3. Stir the reaction at 65 °C for *ca*. 30 min until a clear solution is formed.

4. To the solution from step 3, add 1.1 ml (5 mmol) of SnCl$_4$ in 3 portions

   over 2 min.

5. Stir the reaction at 65 °C for an additional 1 hr under an argon atmosphere.

6. Cool the reaction to r. t. and pour the reaction solution into an ice-cold

   mixture of EtOAc and sat. aq. NaHCO$_3$ 1:1.

7. Separate the organic phase, extract the aqueous phase with EtOAc, and combine the organic phase.

8. Wash the combined organic phase with sat. aq. $NaHCO_3$ and dry it over $MgSO_4$.

9. Evaporate the solvent to dryness under reduced pressure.

10. Purify the residue by column chromatography. Elute the product using DCM/EtOAc 5:2.

The resulting product, N^6-benzoyl-9-(2',3'-di-O-benzoyl-α-L-threo-furanosyl)adenine (**S.23**), is obtained in an 76% yield (1.29 g, 2.3 mmol) as a colorless foam. TLC (DCM/EtOAc 1:1): R_f 0.43. ^{1}H-NMR (300 MHz, $CDCl_3$): δ=4.58 (dd, J=11.1, 2.0, 1H), 4.68 (dd, J=11.1, 3.3, 1H), 5.77 (m, 1H), 6.30 (dd, J=2.0, 0.5, 1H), 6.47 (d, J=1.8, 1H), 7.43-7.67 (m, 9H), 7.86-7.89 (m, 2H), 8.03-8.12 (m, 4H),8.33 (s, 1H), 8.78 (s, 1H), 9.02 (br, 1H).

Prepare S.24

1. Dissolve 1.28 g (2.3 mmol) of **S.22** in 45 ml of THF/MeOH/H_2O 5:4:1 in a water-ice bath at 0 °C.

2. To the solution from step 1 slowly add 3 ml of 2N aq. NaOH.

3. Stir the reaction at 0 °C for 20 min.

4. Neutralize the solution with 3 ml of 2N aq. HCl.

5. Concentrate the mixture under reduced pressure until the desired product starts to precipitate.

6. Cool the flask to 10 °C and isolate the product by filtration.

7. Concentrate the filtrate and isolate a second precipitate by filtration.

8. Evaporate the filtrate from step 7 to dryness under reduced pressure and dissolve the residue in MeOH.

9. Co-evaporate the solution from step 8 with toluene and purify the residue by column chromatography. Elute the sample using DCM/EtOH 7:3 to give more product.

10. Recrystallize the combined products from EtOH.

The resulting product, N^6-benzoyl-9-(α-L-threofuranosyl)adenine (**S.24**), is obtained in a 76% yield (600 mg, 1.8 mmol) as a colorless crystal. TLC (DCM/AcOEt 1:1): R_f 0.03. ^{1}H-NMR (300 MHz, DMSO-d_6): δ=4.17 (m, 3H), 4.44 (m, 1H), 5.60 (br, 1H), 5.94 (d, J=4.5, 1H), 6.06 (d, J=2.0, 1H), 7.56 (m, 2H), 7.64 (m, 1H), 8.05 (m, 2H), 8.55 (s, 1H), 8.76 (s, 1H), 11.17 (br, 1H).

Prepare S.25

1. Dissolve 500 mg (1.45 mmol) of **S.24** in 9 ml of anhydrous DCM/DMF 1:1 in a 100 ml round bottom flask equipped with a stirring bar.

2. Cool the solution from step 1 to 10-15 °C.

3. Add 650 mg (1.9 mmol) of DMT-Cl, 0.9 ml (7.3 mmol) of 2,6-lutidine and 440 mg (1.7 mmol) of AgOTf to the round bottom flask.

4. Stir the reaction at 10 °C for 5 hr.

5. Add 65 mg (0.2 mmol) of DMT-Cl and 44 mg (0.2 mmol) of AgOTf to the reaction.

6. Stir the reaction at r. t. overnight.

7. Quench the reaction with 1.5 ml of MeOH/H$_2$O 1:4.

8. Dilute the mixture from step 7 with 75 ml of DCM.

9. Filter the mixture and wash the residue with DCM.

10. Wash the combined filtrate with ice-cold 0.1 M HCl (2x), water (2x), sat. aq. NaHCO$_3$ and brine.

11. Dry the organic solvent over Na$_2$SO$_4$.

12. Evaporate the solvent to dryness under reduced pressure.

13. Purify the residue by column chromatography. Deactivate the column with DCM/Et$_3$N 97:3 and elute the product using a step-wise gradient of DCM//Et$_3$N 99:1 to DCM/MeOH/Et$_3$N 98:1:1.

The resulting product, N^6-benzoyl-9-{3'-O-[(4",4"'-dimethoxytriphenyl)-methyl]-α-L-threofuranosyl}adenine (**S.25**), and its 2'-O-regioisomer N^6-benzoyl-9-{2'-O-[(4",4"'-dimethoxytriphenyl)methyl]-α-L-threofuranosyl}adenine (**S.25a**), are obtained in a 72% (679 mg, 1.05 mmol) and 2% (25 mg, 0.04 mmol), respectively. TLC (DCM/MeOH 16:1): R$_f$ 0.24 (**S.25**), 0.51 (**S.25a**). ^{1}H-NMR **S.25** (300 MHz, CDCl$_3$): δ=3.76 (s, 3H), 3.77 (s, 3H), 3.80 (m, 2H), 4.26 (m, 1H), 4.33 (m, 1H), 5.90 (d, J=2.1, 1H), 6.80 (m, 4H), 7.28-7.61 (m, 12H), 8.02 (d, J=7.2, 2H), 8.30 (s, 1H), 8.78 (s, 1H), 9.02 (br, 1H). **S.25a** (300 MHz, CDCl$_3$): δ=3.77 (s, 3H), 3.78 (s, 3H), 4.07 (dd, J=10.2, 1.2, 1H), 4.25 (dd, J=10.3, 4.2, 1H), 4.40 (m, 1H), 4.72 (m, 1H), 5.51 (d, J=1.5, 1H), 6.78 (m, 4H), 7.25-7.45 (m, 10H), 7.52-7.67 (m, 3H), 8.05 (m, 2H), 8.76 (s, 1H), 9.05 (br, 1H).

Prepare S.26

1. Dissolve 200 mg (0.31 mmol) of **S.25** and 0.3 ml of DIPEA in 2.0 ml of anhydrous DCM.

2. To the solution from step 1, add 80 μl (0.4 mmol) of chloro(2-cyanoethoxy)(diisopropylamino)phosphine in a dropwise manner.

3. Stir the reaction at r. t. for 1 hr. Then add an extra 10 μl (0.04 mmol) of chloro(2-cyanoethoxy)(diisopropylamino)phosphine to the reaction mixture.

4. Stir the reaction at r. t. overnight.

5. Quench the reaction with 2 drops of MeOH.

6. Remove the solvent to dryness under reduced pressure and purify the residue by column chromatography. Deactivate the column with DCM/Et$_3$N 97:3 and elute the product using hexanes/DCM/Et$_3$N 19:80:1. The resulting product, N^6-benzoyl-9-{2'-O-[(2-cyanoethoxy) (diisopropyl amino)phosphino]-3'-O-[(4",4'''-dimethoxytriphenyl)methyl]-α-L-threofurano syl}adenine (**S.26**), is obtained in an 84% (220 mg, 0.26 mmol) as a mixture of diastereomers. TLC (DCM/MeOH 50:1): R$_f$ 0.20, 0.25. ^{1}H-NMR (300 MHz, CDCl$_3$): δ=1.10-1.26 (m, 24H), 2.35-2.44 (m, 2H), 2.53-2.60 (m, 2H), 3.44 (m, 2H), 3.55-3.89 (m, 32H), 4.26 (m, 1H), 4.36 (m, 1H), 4.50 (m, 1H), 4.98 (m, 1H), 6.16 (s, 1H), 6.23 (s, 1H), 6.77-6.81 (m, 8H), 7.08-7.28 (m, 18H), 7.52-7.62 (m, 6H), 8.02-8.05 (m, 4H), 8.34 (s, 1H), 8.39 (s, 1H), 8.80 (s, 1H), 8.81 (s, 1H), 9.00 (br, 2H).

3.10 Basic Protocol 7

Synthesis, Isolation, and Characterization of Threose Nucleic Acid (TNA)

Oligonucleotides

The threofuranosyl nucleoside phosphoramidites are completely stable

under standard oligonucleotide synthesis conditions on a DNA synthesizer. An

Applied Biosystems 3400 DNA synthesizer has been used successfully on a

1.0-μmol scale synthesis (Yang et al. 2007). Alternative reagents in certain

steps are used in order to increase the reactivity of the phosphoramidites.

These include 1) the use of 0.1 M solution of threofuranosyl nucleoside

phosphoramidites in MeCN, 2) the use of a 6% Cl_2CHCO_2H solution in

$(ClCH_2)_2$ for detritylation, 3) the use of a 0.35 M solution of

(ethylthio)-1H-tetrazole for the coupling step. Meanwhile the modifications to

the synthesis cycle are employed to optimize the chemistry of the automated

synthesis (Figure 3.8). Several techniques have been reported for the isolation

and purification of sequences containing TNA nucleotides, such as C18

reverse phase HPLC and denaturing polyacrylamide gel electrophoresis

(PAGE). This protocol describes the synthesis of TNA oligonucleotides from

TNA monomer phosphoramidite precursors.

Materials

TNA phosphoramidite monomers bearing nucleobases A and T (**S22** and **S.26**,

respectively)
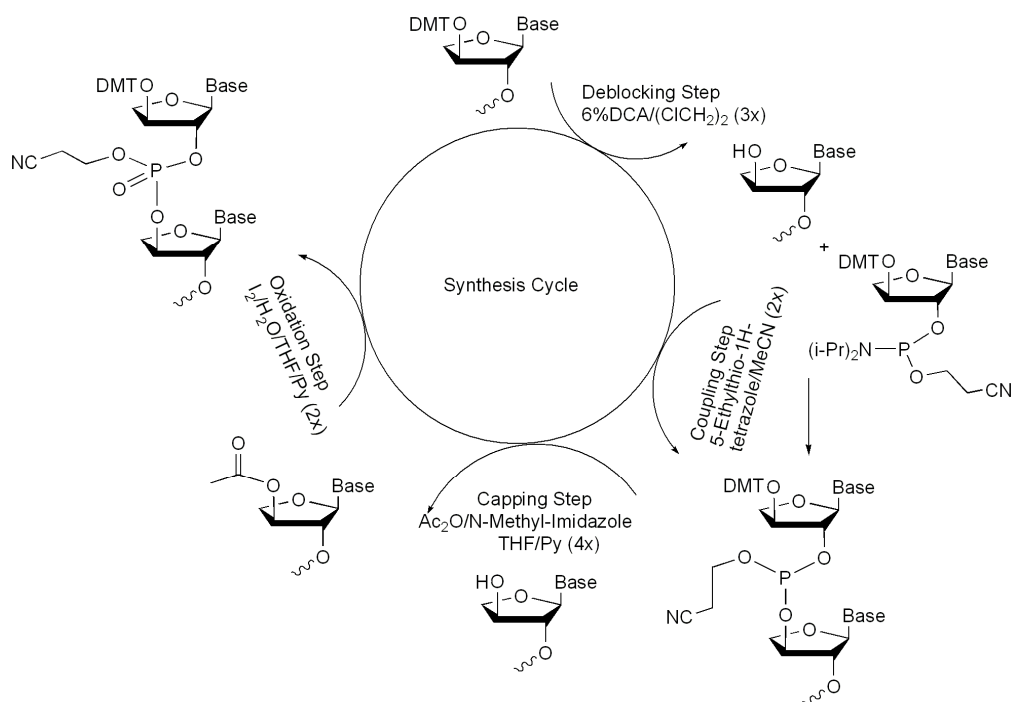
Figure 3.8. Synthesis cycle of automated solid phase TNA synthesis
Abbreviations: DCA, dichloroacetic acid; MeCN, acetonitrile; Ac_2O, acetic
anhydride; THF, tetrahydrofuran; Py, pyridine.

Anhydrous acetonitrile (MeCN) (Applied Biosystems)

6% Cl_2CHCO_2H solution in $(ClCH_2)_2$

0.35 M solution of (ethylthio)-1H-tetrazole

Universal support II (Glen Research)

Argon gas (Ultra high purity)

Ammonium hydroxide, concentrated (conc. NH_4OH)

n-Butanol

Nanopure water

25-ml round bottom flasks

4Å molecular sieves (freshly activated by heating at 300°C for at least 3 hr)

0.45-µm disposable syringe filter

Syringes and needles

Screw-capped eppendorf tube (2 ml)

Screw-capped falcon centrifuge tubes

Additional reagents and equipments for automated solid-phase oligodeoxyribonucleotide synthesis, isolation and characterization of synthetic nucleic acids

Synthesize TNA oligonucleotides (0.2 µmol scale synthesis)

1. Dissolve each TNA phosphoramidte (pre-dried under high vacuum for 48 hr) in anhydrous MeCN in a 25-ml round bottom flask to make an approximate 0.1 M solution (310 mg in 3.7 ml MeCN for **S.26** (A), 300 mg in 4.1 ml for **S.22** (T), respectively).

2. Dry the phosphoramidite solutions from step 1 over freshly activated molecular sieves overnight at r. t.

3. Filter the amidites solution using 0.45-µm disposable syringe filters directly into clean, dry bottles and immediately place the bottles on the corresponding amidite ports of the DNA synthesizer.

   Steps 1 through 3 are carried out under argon atmosphere. All the glassware, syringes and needles are pre-dried in a drying oven and then cooled to r. t. in a desiccator for later use.

4. Start the automated solid-phase oligonucleotide synthesis from CPG-column pre-charged with a universal support using the modified cycle.

112

5. After completing the synthesis, remove the column and dry it for 2 min using vacuum pipeline.

6. Open the CPG-column container and transfer the CPG beads into a 2-ml screw-capped eppendorf tube.

Deprotect and purify TNA oligonucleotides

7. Add 1 ml of conc. NH$_4$OH, seal the tube, and incubate at 55°C for 12 hr.

8. Place the tube on ice and cool for at least 10 min prior to opening

9. Transfer the supernatant to a Falcon tube and add 10 ml of n-butanol.

10. Place the tube on ice for 5 min.

11. Centrifuge for 5 min at 4000 rpm, discard the supernatant and collect the oligonucleotide (white precipitate).

12. Re-suspend the oligonucleotide pellet in 500 µl of nanopure water.

13. Purify the crude product by denaturing PAGE and desalt the oligonucleotide by ethanol precipitation.

14. Quantify the oligonucleotide by UV absorbance using Beer's law at 254 nm.

Check molecular weight

15. Check the molecular weight of the desired TNA oligonucleotide by MALDI-TOF mass spectroscopy using aqueous saturated 3-hydroxypicolinc acid solution as matrix. Prepare matrix just before spotting the sample onto a MALDI-TOF mass spectrometry gold plate.

3.11 Commentray

Background Information

Synthesis

The preparation of TNA building block monomers described in this unit is adopted from the method developed by Eschenmoser (Schöning et al. 2000, 2002). The α/β anomeric mixtures of the protected threofuranose undergo Vorbrüggen-Hilbert-Johnson reactions (Vorbrggen and Bennua 1981) with suitably protected nucleobases in the presence of a proper Lewis acid to afford the corresponding L-threofuranosyl nucleosides in exclusive α-configuration. The tritylation steps of these nucleosides show low regioselectivity and require careful chromatographic separations of the 3'-tritylated isomers from their 2' analogues.

Properties of threose nucleic acid system

TNA is discovered as an evolutionary progenitor of RNA as a result of systematic exploration of structural neighborhood of natural nucleic acids (Eschenmoser 1999; Schöning et al. 2000, 2002). Although the backbone displays a constitutional periodicity of five atoms with a phosphodiester connection between the 3'- and 2'-positions of the threose sugar, TNA is capable of intra- and intersystem base-pairing with itself and natural nucleic acids (RNA and DNA), respectively, to form stable duplexes in an antiparallel orientation (Schöning et al. 2000, 2002). Structural studies show that the intrasystem duplex of TNA forms a curved right-handed double helix in the

114

Watson-Crick pairing mode. Both the NMR and X-ray experiments implicate

that TNA remarkably resemble the structure of A-form RNA/DNA (Ebert et al.

2008; Pallan et al. 2003; Wilds et al. 2002). These findings corroborate the

experimental observations that TNA base pairs stronger with RNA than with

DNA and that RNA serves as a better template for ligating TNA fragments

than DNA (Wu et al. 2002).

Critical Parameters

For all reactions using anhydrous solvents, the glassware must be pre-dried

in a drying oven (80°C). To obtain optimal separations and to reduce the

purification time, we recommend using flash column chromatography for all

purification steps. For phosphoramidite synthesis, the reaction progress is

monitored by TLC and should be stopped as soon as the starting material is

consumed. The phosphitylation reaction affords two diastereomeric

phosphoramidite products, which are general visualized as two spots on the

TLC plate. The TNA phosphoramidite building blocks must be dried under

high vacuum in a desiccator for 24 hr to give a well-foamed solid before they

can be used for solid-phase synthesis.

Trouble shooting

The constitutional assignment of the 2'- and 3'-O-DMT isomers are based

on homo decoupling and/or H^1,H^1-COSY-NMR spectroscopy by correlating

the H-C(2') or H-C(3') to the OH atom. The TNA guanosine phosphoramidite

appears as a syrup after column purification. Thus, a pentane precipitation is

further performed to convert the syrup into a solid form, which would be conveniently used in the solid phase synthesis.

Anticipated Results

The protocols described in this unit are useful for preparing TNA oligonucleotides. The TNA phosphoramidite monomers can be synthesized on a multi-gram scale. The oligonucleotide synthesis is similar to the standard DNA phosphoramidite protocol with the noted exception of elongated coupling times. The average stepwise yield is >95%, which is sufficient to produce TNA oligonucleotides in high yield.

Time Considerations

Each step in the chemical synthesis can be accomplished in 1-2 working days (including purification). The time for oligonucleotide synthesis varies slightly from the standard phosphoramidite method by longer coupling times and multiple times of reagent delivery and wash. The deprotection, isolation and the analysis of the final synthetic oligonucleotide are accomplished in 2-3 days.

3.12 References

Chaput, J.C., J.K. Ichida, and J.W. Szostak. 2003a. DNA polymerase-mediated DNA synthesis on a TNA template. *J. Am. Chem. Soc.* 125: 856-857.

Chaput, J.C., and J.W. Szostak. 2003b. TNA synthesis by DNA polymerases. *J. Am. Chem. Soc.* 125: 9274-9275.

Ebert, M.-O., C. Mang, R. Krishnamurthy, A. Eschenmoser, and B. Jaun. 2008. The structure of a TNA-TNA complex in solution: NMR study of the octamer duplex derived from α-(L)-threofuranosyl-(3'-2')-CGAATTCG. *J. Am. Chem. Soc.* 130: 15105-15115.

Eschenmoser, A. 1999. Chemical etiology of nucleic acid structure. *Science* 284: 2118-2124.

Horhota, A., K. Zou, J.K. Ichida, B. Yu, L.W. McLaughlin, J.W. Szostak, and J.C. Chaput. 2005. Kinetic analysis of an efficient DNA-dependent TNA polymerase. *J. Am. Chem. Soc.* 127: 7427-7434.

J.K. Ichida, A. Horhota, K. Zou, L.W. McLaughlin, and J.W. Szostak. 2005a. High fidelity TNA synthesis by Therminator polymerase. *Nucleic Acids Res.* 33: 5219-5225.

Ichida, J.K., K. Zou, A. Horhota, B. Yu, L.W. McLaughlin, and J.W. Szostak. 2005b. An in vitro selection system for TNA. *J. Am. Chem. Soc.* 137: 2802-2803.

Orgel, L., 2000. A simpler nucleic acid. *Science* 17: 1306-1307.

Pallan, P.S., C.J. Wilds, Z. Wawrzak, R. Krishnamurthy, A. Eschenmoser, and M. Egli. 2003. Why does TNA cross-pair more strongly with RNA than with DNA? An answer from X-ray analysis. *Angew. Chem. Int. Ed.* 42: 5893-5895.

Schöning, K.-U., P. Scholz, S. Guntha, X. Wu, R. Krishnamurthy, and A. Eschenmoser. 2000. Chemical etiology of nucleic acid structure: the α-L-threofuranosyl-(3'→2')-oligonucleotide system. Science 290: 1347-1351.

Schöning, K.-U., P. Scholz, X. Wu, S. Guntha, G. Delgado, R. Krishnamurthy, and A. Eschenmoser. 2002. The α-L-threofuranosyl-(3'→2')-oligonucleotide system ('TNA'): synthesis and pairing properties. *Helv. Chim. Acta* 85: 4111-4153.

Vorbrggen, H., and B. Bennua. 1981. Nucleoside syntheses, XXV: a new simplified nucleoside synthesis. *Chem. Ber.* 114: 1279-1286.

Wilds, C.J., Z. Warwrzak, R. Krishnamurthy, A. Eschenmoser, and M. Egli. 2002. Crystal structure of a B-form DNA duplex containing (L)-α-threofuranosyl (3' →2') nucleosides: a four-carbon sugar is easily accommodated into the backbone of DNA. *J. Am. Chem. Soc.* 124: 13716-13721.

Wu, X., G. Delgado, R. Krishnamurthy, and A. Eschenmoser. 2002. 2,6-Diaminopurine in TNA: effects on duplex stabilities and on the efficiency of template-controlled ligations. *Org. Lett.* 4: 1283-1286.

Yang, Y.-W., S. Zhang, E.O. McCullum, and J.C. Chaput. 2007. Experimental evidence that GNA and TNA were not sequential polymers in the prebiotic evolution of RNA. *J. Mol. Evol.* 65: 289-295.

Chapter 4

EXPERIMENTAL EVIDENCE THAT GNA AND TNA WERE NOT

SEQUENTIAL POLYMERS IN THE PREBIOTIC EVOLUTION OF RNA

4.1 Contributions

The following chapter describes an experimental investigation into the

molecular recognition properties of GNA and TNA. The reported experiments

were performed by Ying-Wei Yang, Su Zhang and Elizabeth McCullum under

the guidance of Professor John Chaput. Ying-Wei Yang, Su Zhang and

Elizabeth McCullum synthesized the GNA and TNA phosphoramidites and

oligonucleotides. Ying-Wei Yang and Su Zhang measured the thermodynamic

properties of complementary GNA and TNA helices. John Chaput wrote the

manuscript with communications from Su Zhang. The results from this work

were published in *Journal of Molecular Evolution*.

4.2 Abstract

Systematic investigation into the chemical etiology of ribose has led to the

discovery of glycerol nucleic acid (GNA) and threose nucleic acid (TNA) as

possible progenitor candidates of RNA in the origins of life. Coupled with

their chemical simplicity, polymers for both systems are capable of forming

stable Watson-Crick antiparallel duplex structures with themselves and RNA,

thereby providing a mechanism for the transfer of genetic information between

successive genetic systems. Investigation into whether both polymers arose

independently or descended from a common evolutionary pathway would provide additional constraints on models that describe the emergence of a hypothetical RNA world. Here we show by thermal denaturation that complementary GNA and TNA mixed sequence polymers are unable, even after prolonged incubation times, to adopt stable helical structures by intersystem cross-pairing. This experimental observation suggests that GNA and TNA, whose structures derive from one another, were not consecutive polymers in the same evolutionary pathway to RNA.

4.3 Introduction

The belief that life on Earth arose during an ancestral period in which RNA stored genetic information and catalyzed reactions is commonly referred to as the RNA world hypothesis (Gilbert 1986; Joyce 1989). Although RNA-based organisms have yet to be discovered, presumably because any such organism would have predated all known geological record, evidence of their existence can be found in contemporary living systems (Joyce 2002). Principal among these are catalytic RNA, first discovered in RNA molecules derived from group I intervening sequences (Kruger et al. 1982), and the x-ray crystal structure of the ribosome, which shows that RNA is responsible for protein synthesis (Ban et al. 2000; Wimberly et al. 2000; Yusupov et al. 2001). Whether RNA was the first genetic material or an evolutionary intermediate that eventually gave rise to biotic life based on DNA and proteins remains less

120

clear. Limitations over the prebiotic synthesis of ribose and the problem of

enantiomeric cross-inhibition during nonenzymatic replication have led some

to speculate that RNA was preceded by a simpler genetic system that was

more prebiotically accessible and capable of synthesizing chiral polymers

from racemic mixtures of activated mononucleotides (Chaput and Switzer

2000; Joyce et al. 1987; Schneider and Benner 1990).

Determining the structures and conditions that may have given rise to an

RNA world is fundamental to understanding the origin and evolution of life.

The approach taken thus far has been to construct by chemical methods a

series of potentially natural nucleic acid systems whose structures derive from

a sugar-phosphate backbone, and to compare the capacity for each system to

transfer genetic information, in the form of Watson-Crick base pairing,

between itself and RNA (Eschenmoser 1999). Since each structure derives

from a natural sugar and nucleic acid-base pairing is a chemical property

considered fundamental to biology, nucleic acid systems capable of adopting

stable helical structures with themselves and RNA are considered potentially

natural progenitor candidates of RNA. Initial investigation into the chemical

etiology of ribofuranosyl nucleic acids by systematic exploration of the

structures related to RNA led to the observation that Watson-Crick base

pairing among the hexo- and pentopyranosyl analogues of RNA (in systems

where it occurs) is orthogonal to that of natural nucleic acids (Eschenmoser

1999). While the hexose and pentose systems were reportedly chosen for their

121

ability to maintain the same six-atom backbone repeat unit found in natural

nucleic acids, modeling studies combined with NMR data later revealed that

these systems did not adopt a standard RNA helix (Schlonvogt et al. 1996).

Further examination of other members of the aldose family of sugars, however,

led to the unexpected discovery that the tetrapyranosyl analogue,

(3′,2′)-α-L-threose nucleic acid (TNA), which has a five-atom backbone repeat

unit (one atom shorter than natural RNA), was able to form stable double

helical structures with itself and natural RNA and DNA (Schöning et al. 2000).

More recently, glycerol nucleic acid (GNA), an acyclic analogue of TNA

composed of a three-carbon propylene glycol backbone with a five-atom

repeat unit, was also found to base pair with itself and complementary RNA

(Zhang et al. 2005). Together, GNA and TNA (Figure 4.1) represent the only

known candidate progenitors of RNA that derive from a sugar-phosphate

backbone and exchange information with themselves and RNA.

4.4 Materials and Methods

4.4.1 Oligonucleotide Synthesis

GNA and TNA phosphoramidites were synthesized as previously described

(Acevedo and Andrews 1996; Holý 1975; Schöning et al. 2002).

Oligonucleotides were synthesized on an Applied Biosystems 3400 DNA

synthesizer using standard β-cyanoethyl phosphoramidite chemistry. GNA and

TNA oligonucleotides were synthesized on a Universal Support II CPG
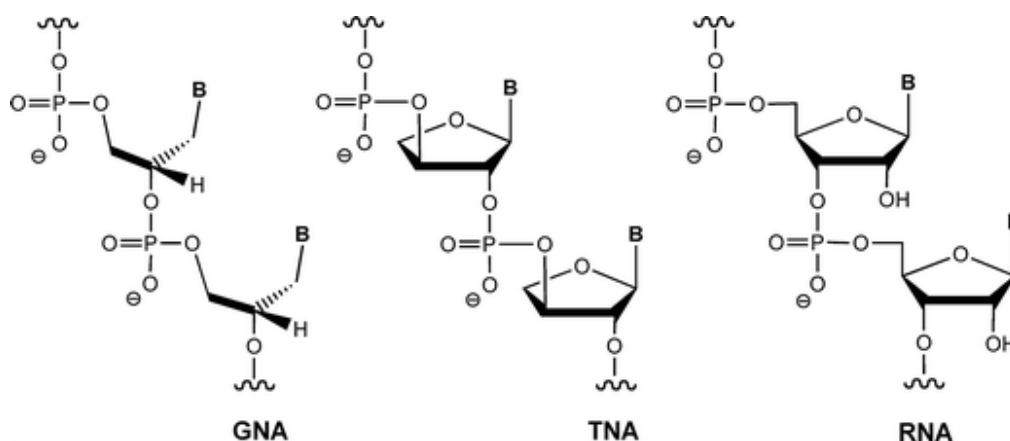
Figure 4.1. Possible candidate progenitors of RNA. Glycerol-derived nucleic acid (GNA) is composed of a three-carbon sugar with a five-atom backbone repeat unit. Threose nucleic acid (TNA) derives from a four-carbon sugar with a five-atom backbone repeat unit. Natural ribose nucleic acid (RNA) is a five-carbon sugar with a six-atom backbone repeat unit. GNA and TNA are both capable of forming stable duplexes with themselves and RNA, thereby providing a mechanism for the transfer of information between successive genetic systems. Reprinted with permission from Journal of Molecular Evolution. Copyright 2007 Springer.

column (1 μM scale; Glen Research). GNA oligonucleotides required a synthesis protocol with extended (5 min) coupling times. TNA oligonucleotides were synthesized under modified conditions as previously described (Schöning et al. 2000). Stepwise coupling yields were greater than 95% for all GNA and TNA oligonucleotides. Cleavage from the solid support and deprotection of the oligonucleotides was achieved in NH_4OH (33%) for 12–16 h at 55°C. RNA oligonucleotides were purchased from Dharmacon RNA technologies and deprotected using the manufacturer's recommended protocol. Oligonucleotides were purified by preparative denaturing polyacrylamide gel electrophoresis, isolated, ethanol precipitated, and desalted on a Sephadex G-25 ion-exchange resin. Pure salt-free oligonucleotides were

validated by MALDI-TOF mass spectroscopy and stored in H_2O at $-20°C$.

4.4.2 Thermal Melting Measurements

UV absorbance-versus-temperature profiles were measured on a Varian Cary 300 Bio UV-visible spectrophotometer in a temperature-controlled cell holder. All melts were performed under standard conditions with 1.0 M NaCl, 10 mM NaH_2PO_4, 0.1 mM NaEDTA, pH 7.0, and 1:1 oligonucleotide stoichiometry (c = 2.5 μM + 2.5 μM). Melting curves were obtained in the reverse and forward melting directions with a temperature gradient of 0 to 90°C and a ramping speed of 1°C per minute by monitoring the change in UV absorbance at 260 nm at each temperature. T_m values were determined as the first-order derivative using the Cary software package, and represent the average of three independent trials. Thermodynamic values for ΔG, ΔH, and ΔS were determined from plots of T_m^{-1} versus lnc (Marky and Breslauer 1987).

4.5 Results and Discussion

In the present study, we sought to investigate the evolutionary relationship between GNA and TNA as possible progenitor candidates of RNA. This investigation can be viewed as a functional selection for informational systems compatible with RNA. The observation that GNA pairs with RNA, but not DNA, while TNA pairs efficiently opposite both DNA and RNA led us to wonder whether TNA might be closer in evolutionary history to RNA than

124

GNA. This question stems from the belief that a nucleic acid polymer closer in evolutionary history to RNA would be expected to have greater affinity for DNA than a system closer to the primordial origin. Consistent with this notion is the general trend in structural complexity, which follows the order GNA << TNA < RNA due to intrinsic differences between the constitutional and the configuration isomers of the individual sugars. The same trend also applies to the observed base-pairing efficiency of GNA and TNA with RNA.

4.5.1 Chemical Synthesis of GNA and TNA Oligonucleotides

We began with the chemical synthesis of GNA and TNA oligonucleotides using previously established methodologies (Acevedo and Andrews 1996; Holý 1975; Schöning et al. 2002). GNA phosphoramidite monomers were synthesized from commercially available *R*-(+)-glycidol, **1** (Figure 4.2), by treating the protected chiral epoxide **2** with thymine or adenine in the presence of NaH to afford the acyclic glycol nucleosides **3** and **5**, respectively. The thymidine derivative **3** was directly converted to the phosphoramidite product **4**, while the adenosine derivative **5** required benzoyl protection of the exocyclic amine prior to phosphitylation of **5** to **7**. Synthesis of TNA phosphoramidite monomers (Figure 4.3) initiated with a multistep synthesis of
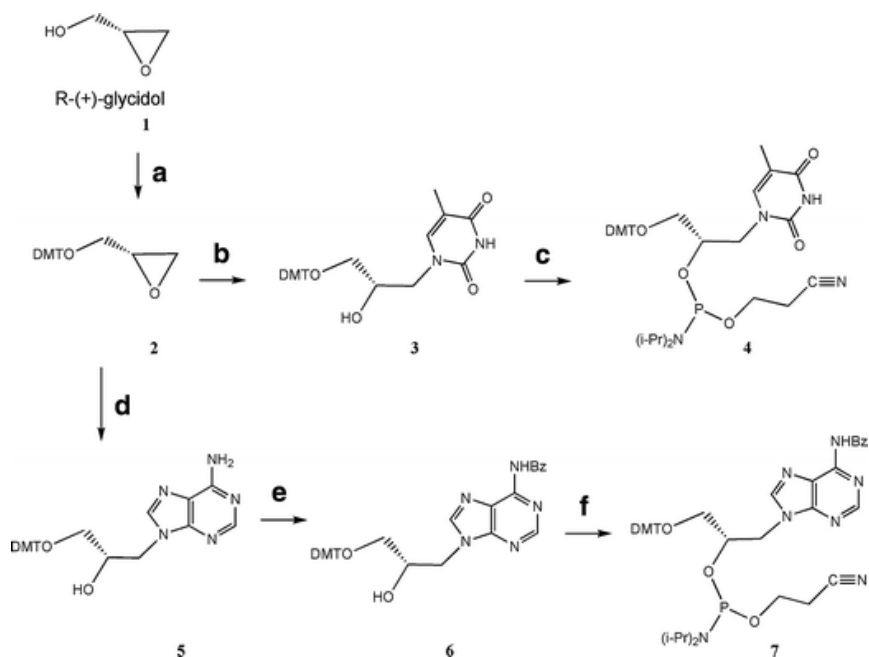
Figure 4.2. Chemical synthesis of (S)-glycol nucleic acid building blocks for automated solid phase oligonucleotide synthesis: (a) DMT-Cl, TEA, DCM (88%). (b) thymine, NaH, DMF (55%). (c) chloro(2-cyanoethoxy)-(diisopropylamino)phosphine, (i-Pr)$_2$NEt, DCM (78%). (d) adenine, NaH, DMF (51%). (e) TMSCl, Pyridine, BzCl, then concentrated ammonia (80%). (f) chloro(2-cyanoethoxy)-(diisopropylamino)phosphine, (i-Pr)$_2$NEt, DCM (70%). Reprinted with permission from Journal of Molecular Evolution. Copyright 2007 Springer.

triacylated L-threofuranose derivatives **9a** and **9b**, from L-ascorbic acid, **8**.

Vorbrüggen-Hilbert-Johnson glycosylation of the α/β-anomer mixtures of **9a** and **9b** with thymine and benzoyl protected adenosine gave nucleosides **10a** and **10b**, respectively, in high yield. Hydrolysis and tritylation afforded both the 2′ and the 3′ regioisomers for **12a** and **12b** that were separated by silica gel chromatography. The low tritylation yield (c.a. 25%) of the thymidine nucleoside **12a** was considered acceptable since the alternative strategy, which proceeds through an anhydrothymidine derivative, involves several additional steps, with an overall yield of c.a. 30%. Thymidine and adenosine nucleosides
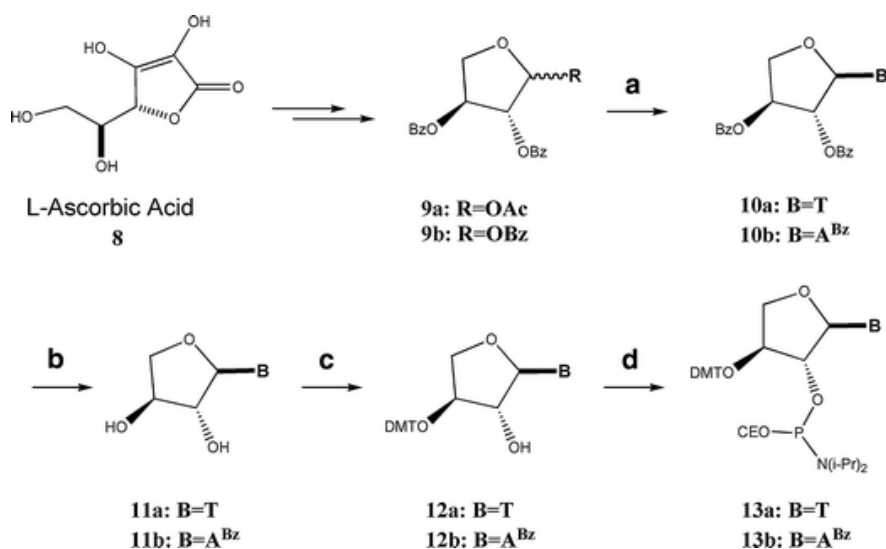
Figure 4.3. Chemical synthesis of (L)-α-threofuranosyl nucleic acid building blocks for automated solid phase oligonucleotide synthesis: (a) for **10a**: thymine, BSA, TMS-Tf, MeCN (89%); for **10b**: benzoyl adenine, BSA, TMS-Tf, MeCN (76%). (b) for **11a**: NH_3 in MeOH (99%); for **11b**: aq. NaOH, THF, MeOH, H_2O (85%). (c) DMT-Cl, 2,6-lutidine, Ag-Tf, DMF, DCM, followed by chromatographic separation (25% for **12a** and 25% for **12b**). (d) chloro(2-cyanoethoxy)-(diisopropylamino)phosphine, (i-Pr)$_2$NEt, DCM (70% for **13a** and 84% for **13b**). Reprinted with permission from Journal of Molecular Evolution. Copyright 2007 Springer.

12a and **12b** were then converted to phosphoramidite monomers **13a** and **13b** by standard phosphitylation chemistry. Thymidine and adenosine phosphoramidite monomers of GNA (**4**, **7**) and TNA (**13a**, **13b**) were used as building blocks for automated solid phase oligonucleotide synthesis.

4.5.2 Comparison of GNA, TNA, RNA, and DNA Self-Pairing Modes

We examined the base-pairing properties of GNA and TNA mixed sequence polymers by temperature-dependent ultraviolet (UV) spectroscopy. Melting temperature (T_m) values and thermodynamic parameters for all self-pairing and cross-pairing duplex structures are summarized in Tables 4.1 and 4.2. Self-pairing among the four genetic systems (Table 4.1, Figure 4.4a)

reveals that TNA base pairing occurs with similar thermal stability, but

markedly lower thermodynamic stability (ΔΔG ~ 4 kcal/mol), than duplexes

formed by RNA and DNA. By contrast, GNA base pairing gives rise to

duplexes with greater thermal stability (58°C versus 45–47°C, respectively)

but lower thermodynamic stability (ΔΔG ~ 2 kcal/mol) than RNA and DNA.

This observation is consistent with the earlier finding that the thermal and

thermodynamic stabilities of TNA self-pairing modes are not entirely

Table 4.1. Temperature-dependent UV melting data for all GNA, TNA, RNA, and DNA self-pairing and cross-pairing modes. Reprinted with permission from Journal of Molecular Evolution. Copyright 2007 Springer.

Entry	Oligonucleotide pairing system	T_m (°C)[a]	ΔG (kcal/mol)	ΔH (kcal/mol)	$T\Delta S$ (kcal/mol)
1	$d(A_4T_3AT_3AT_2A_2)+d(T_4A_3TA_3TA_2T_2)$	47	-14.6	-93.1	-78.5
2	$r(A_4U_3AU_3AU_2A_2)+r(U_4A_3UA_3UA_2U_2)$	45	-14.0	-94.4	-80.4
3	$d(A_4T_3AT_3AT_2A_2)+r(U_4A_3UA_3UA_2U_2)$	36	-8.6	-31.3	-22.7
4	$r(A_4U_3AU_3AU_2A_2)+d(T_4A_3TA_3TA_2T_2)$	30	-7.9	-31.1	-23.2
5	$g(A_4T_3AT_3AT_2A_2)+g(T_4A_3TA_3TA_2T_2)$	58	-12.3	-42.4	-30.1
6	$g(A_4T_3AT_3AT_2A_2)+r(T_4A_3TA_3TA_2T_2)$	36	-8.2	-16.9	-8.7
7	$r(A_4T_3AT_3AT_2A_2)+g(T_4A_3TA_3TA_2T_2)$	39	-8.6	-20.4	-11.8
8	$t(A_4T_3AT_3AT_2A_2)+t(T_4A_3TA_3TA_2T_2)$	42	-10.6	-47.1	-36.5
9	$t(A_4T_3AT_3AT_2A_2)+d(T_4A_3TA_3TA_2T_2)$	29	-8.7	-53.2	-44.5
10	$d(A_4T_3AT_3AT_2A_2)+t(T_4A_3TA_3TA_2T_2)$	45	-12.1	-65.4	-53.3
11	$t(A_4T_3AT_3AT_2A_2)+r(T_4A_3TA_3TA_2T_2)$	33	-9.7	-61.6	-51.9
12	$r(A_4T_3AT_3AT_2A_2)+t(T_4A_3TA_3TA_2T_2)$	42	-12.1	-74.9	-62.8

Table 4.2. T_m values for intra- and intersystem base pairing between GNA, TNA, RNA, and DNA. Reprinted with permission from Journal of Molecular Evolution. Copyright 2007 Springer.

		$A_4T_3AT_3AT_2A_2$			
		GNA	**TNA**	**RNA**	**DNA**
$T_2A_2TA_3TA_3T_4$	**GNA**	58	na	39	na
	TNA	na	42	42	45
	RNA	36	33	45	36
	DNA	na	29	31	47

comparable with natural systems (Schöning et al. 2000), possibly due to subtle

differences in the helical structure of TNA relative to RNA and DNA. A

similar comparison of GNA was not possible since the thermodynamic

properties of GNA were not previously reported.
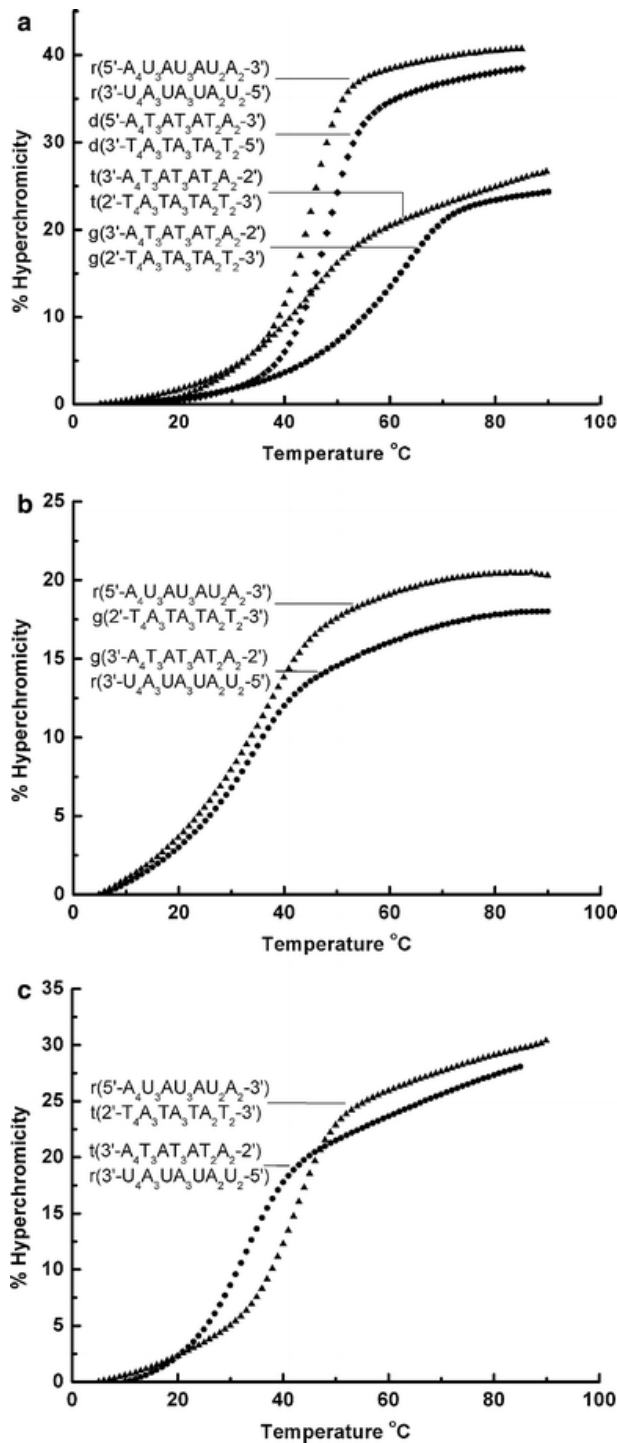
Figure 4.4. Temperature-dependent UV melting curves for all self-pairing and RNA cross-pairing modes. (A) Melting curves for GNA-GNA, TNA-TNA, RNA-RNA, and DNA-DNA self-pairing modes. (B-C) Melting curves for GNA and TNA cross-pairing modes opposite complementary RNA. Sigmoidal curves obtained for each melting plot show cooperative duplex formation. Reprinted with permission from Journal of Molecular Evolution. Copyright 2007 Springer.

4.5.3 Comparison of GNA and TNA Cross-Pairing with RNA

Cross-pairing of GNA and TNA oligonucleotides with natural RNA and DNA confirms that both GNA and TNA base pair opposite complementary RNA, while only TNA base pairs opposite complementary DNA. Among the cross-pairing data some differences occur in the melting behavior of the duplexes. Interstrand pairing between TNA and RNA was found to be more sequence dependent than interstrand pairing between GNA and RNA. GNA-RNA base pairing gives rise to two 16-mer duplexes (Table 4.1, Figure 4.4b) with comparable thermal and thermodynamic stabilities ($\Delta T_m \sim$ 3°C, $\Delta\Delta G \sim 0.4$ kcal/mol), while TNA-RNA base pairing leads to the formation of two duplexes (Figure 4.4c) with different thermal ($\Delta T_m \sim 9$°C) and thermodynamic ($\Delta\Delta G \sim 2.4$ kcal/mol) properties. Comparison of these differences to the differences found in duplexes formed by interstrand pairing between DNA and RNA, where the ΔT_m is 6°C and the $\Delta\Delta G$ value is 0.7 kcal/mol, suggest that TNA and DNA are more sequence dependent than GNA. This unusual property could be a consequence of the intrinsic flexibility of the acyclic backbone of GNA relative to the rigid backbones of all other nucleic acid systems that derive from natural sugars.

4.5.4 Comparison of GNA Cross-Pairing with TNA

We assessed the possibility of whether complementary strands of GNA and TNA would associate to form stable duplex structures by collecting

temperature-dependent UV absorbance profiles for both sets of

complementary strands. In sharp contrast to GNA and TNA self-pairing,

where thermal melting curves give sigmoidal plots with a single cooperative

transition between the native and the denatured states, melting curves obtained

for complementary GNA and TNA sequences (Figure 4.5) result in
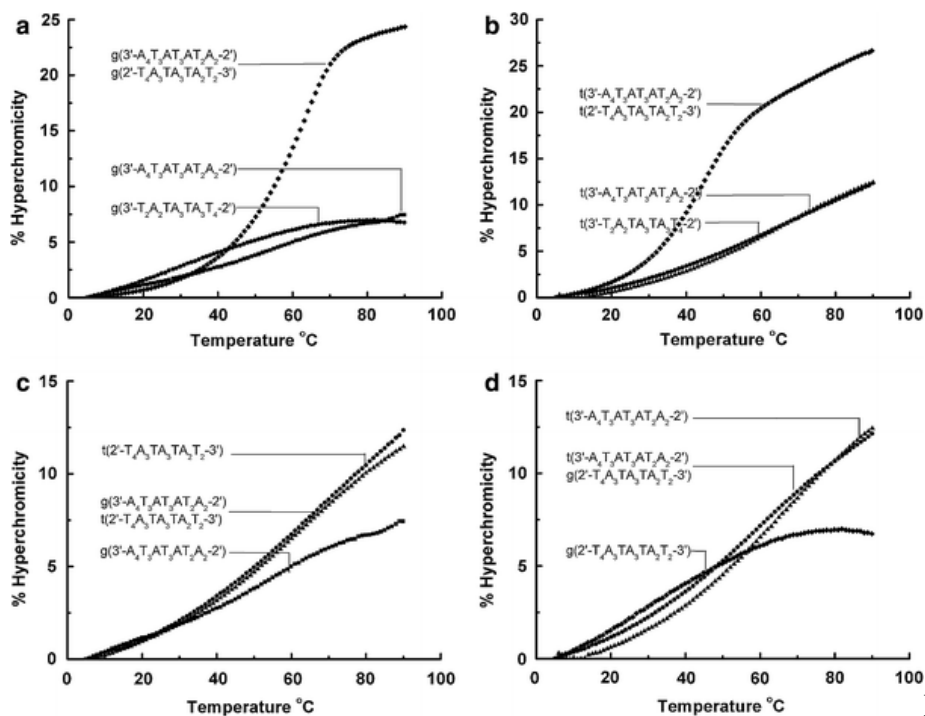
Figure 4.5. Temperature-dependent UV melting curves for GNA-TNA cross-pairing. (A) Melting curves for GNA duplex pairing relative to the individual GNA strands. (B) Melting curves for TNA duplex pairing relative to the individual TNA strands. (C-D) Melting curves for GNA-TNA heteroduplexes relative to their individual strands. The absence of a clear melting transition for GNA-TNA and TNA-GNA duplexes demonstrates that GNA and TNA are not capable of interstrand base pairing. Reprinted with permission from Journal of Molecular Evolution. Copyright 2007 Springer.

hyperchromicity changes that are indistinguishable from the single strands.

Attempts to overcome this problem by slowing the rate of temperature change

and sampling various salt and annealing conditions failed to induce duplex

formation. One possible reason for this could be that TNA adopts a backbone geometry that more closely resembles the B-form helix of DNA, a helical geometry that is incompatible with GNA base pairing (Zhang et al. 2005). Indirect evidence to support a B-form helix can be found in the high-resolution crystal structure of a self-complementary DNA duplex containing a single (L)-α-threofuranosyl thymidine residue in each strand (Wilds et al. 2002) and DNA polymerase experiments which show that threose nucleoside triphophates (tNTPs) are significantly more efficient substrates for template-directed polymerization by DNA polymerases than glycerol nucleotide triphosphates (gNTPs) (Horhota et al. 2005, 2006).

4.6 Conclusions

In summary, we find that GNA and TNA, two potentially primordial ancestors of RNA, are not capable of adopting stable helical structures by interstrand cross-pairing and thus were not consecutive polymers in the same evolutionary pathway to RNA. Given the potential for both systems to serve as functional progenitors of RNA and the need for a pre-RNA world to overcome inherent limitations in the synthesis and replication of RNA, we suggest three possible roles for GNA and TNA in the origin and evolution of RNA. One possibility is that an unknown adapter polymer may have emerged that bridged the two systems, thereby allowing RNA to evolve by a single evolutionary pathway from more prebiotically available materials. The second

133

possibility is that both systems independently converged on RNA through two separate evolutionary lineages. A significant implication of the second scenario is that RNA might possess unique chemical and physical properties not present in any of the ancestral systems that eventually led to the evolution of biotic life based on DNA and proteins. Although it is also possible that GNA and TNA were simply not progenitors of RNA, their unusual ability to transfer genetic information with RNA, chemical simplicity, and abundance on carbonaceous meteorites suggests that both molecules may have played an important role on the early Earth (Cooper et al. 2001). Discerning which evolutionary scenario most likely occurred will require comparing the fitness of each candidate system with respect to a given functional property. Such experiments are expected to challenge our ability to generate unnatural nucleic acid aptamers and catalysts by in vitro selection and directed evolution (Ichida et al. 2005).

4.7 References

Acevedo, O.L., and R.S. Andrews. 1996. Synthesis of propane-2,3-diol combinatorial monomers. *Tetrahedron Lett.* 37: 3931–3139.

Ban, N., P. Nissen, J. Hansen, P.B. Moore, and T.A. Steitz. 2000. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* 289:905–920.

Chaput, J.C., and C. Switzer. 2000. Nonenzymatic oligomerization on tempaltes containing phosphodiester-linked acyclic glycerol nucleic acid analogues. *J. Mol. Evol.* 51: 464–470.

Cooper, G., N. Kimmich, W. Belisle, J. Sarinana, K. Brabham, and L. Garrel. 2001. Carbonaceous meteorites as a source of sugar-related organic compounds for the early earth. *Nature* 414: 879–883.

Eschenmoser, A. 1999. Chemical etiology of nucleic acid structure. *Science* 284: 2118–2124.

Gilbert, W. 1986. The RNA world. *Nature* 319: 618.

Holý, A. 1975. Aliphatic analogues of nucleosides, nucleotides, and oligonucleotides. *Collect. Czech. Chem. Commun.* 40: 187–214.

Horhota, A., K. Zou, J.K. Ichida, B. Yu, L.W. McLaughlin, J.W. Szostak, and J.C. Chaput. 2005. Kinetic analysis of an efficient DNA-dependent TNA polymerase. *J. Am. Chem. Soc.* 127: 7427–7434.

Horhota, A.T., J.W. Szostak, and L.W. McLaughlin. 2006. Glycerol nucleoside triphosphates: Synthesis and polymerase substrate activities. *Org. Lett.* 8: 5345–5347.

Ichida, J.K., K. Zou, A.T. Horhota, B. Yu, L.W. McLaughlin, and J.W. Szostak. 2005. An in vitro selection system for TNA. *J. Am. Chem. Soc.* 127: 2802–2803.

Joyce, G.F. 1989. RNA evolution and the origins of life. *Nature* 338: 217–224.

Joyce, G.F. 2002. The antiquity of RNA-based evolution. *Nature* 418: 214–221.

Joyce, G.F., A.W. Schwartz, S.L. Miller, and L.E. Orgel. 1987. The case for an ancestral genetic system involving simple analogues of the nucleotides. *Proc. Natl. Acad. Sci. USA* 84: 4398–4402.

Kruger, K., P.J. Grabowski, A.J. Zaug, and T.R. Cech. 1982. Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence by Tetrahymena. *Cell* 31: 147–157.

Marky, L.A., and K.J. Breslauer. 1987. Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves. *Biopolymers* 26: 1601–1620.

Schlonvogt, I., S. Pitsch, C. Lesuer, A. Eschenmoser, B. Jaun, and R.M. Wolf. 1996. Pyranosyl-RNA ('p-RNA'): NMR and molecular dynamics study of the duplex formed by self-pairing of ribopyranosyl-(C-G-A-A-T-T-C-G). *Helv. Chim. Acta.* 79: 2316–2345.

Schneider, K.C., and S.A. Benner. 1990. Oligonucleotides containing flexible nucleoside analogues. *J. Am. Chem. Soc.* 112: 453–455.

Schöning K-U, P. Scholz, S. Guntha, X. Wu, R. Krishnamurthy, and A. Eschenmoser. 2000. Chemical etiology of nucleic acid structure: The α-L-threofuranosyl-(3′→2′) oligonucleotide system. *Science* 290: 1347–1351.

Schöning K-U, P. Scholz, X. Wu, S. Guntha, D. Guillermo, R. Krishnamurthy, and A. Eschenmoser. 2002. The α-L-threofuranosyl-(3′→2′)-oligonucleotide system('TNA'): Synthesis and pairing properties. *Helv. Chim. Acta.* 85: 4111–4153.

Wilds, C.J., Z. Wawrzak, R. Krishnamurthy, A. Eschenmoser, and M. Egli. 2002. Crystal structure of a B-form DNA duplex containing (L)-α-threofuranosyl (3′-2′) nucleosides: a four-carbon sugar is easily accommodated into the backbone of DNA. *J. Am. Chem. Soc.* 124: 13716–13721.

Wimberly, B.T., D.E. Brodersen, W.M. Clemons, R.J. Morgan-Warren, A.P. Carter, C. Vonrhein, T. Hartsch, and V. Ramakrishnan. 2000. Structure of the 30S ribosomal subunit. *Nature* 407: 327–339.

Yusupov, M.M., G.Z. Yusupova, A. Baucom, K. Lieberman, T.N. Earnest, J.H.D. Cate, and H.F. Noller. 2001. Crystal structure of the ribosome at 5.5 Å resolution. *Nature* 292: 883–896.

Zhang, L., A. Peritz, and E. Meggers. 2005. A simple glycol nucleic acid. *J. Am. Chem. Soc.* 127: 4174–4175.

Chapter 5

SYNTHESIS OF TWO MIRROR IMAGE 4-HELIX JUNCTIONS

DERIVED FROM GLYCEROL NUCLEIC ACID

5.1 Contributions

The following chapter describes the use of glycerol nucleic acid (GNA) for

assembling nucleic acid nanostructures. The reported experiments were

performed by Su Zhang and Elizabeth McCullum under the guidance of

Professor John Chaput. Su Zhang synthesized GNA phosphoramidites and

oligonucleotides. Su Zhang and Elizabeth McCullum characterized the

assembly of the GNA 4-helix junctions. John Chaput wrote the manuscript

with comments from Su Zhang and Elizabeth McCullum. The results from this

work were published in *Journal of the American Chemical Society*.

5.2 Abstract

Structural DNA nanotechnology relies on Watson−Crick base pairing rules

to assemble DNA motifs into diverse arrangements of geometric shapes and

patterns. While substantial effort has been devoted to expanding the

programmability of natural DNA, considerably less attention has been given to

the development of nucleic acid structures based on non-natural DNA

polymers. Here we describe the use of glycerol nucleic acid (GNA), a simple

polymer based on an acyclic repeating unit, as an alternative genetic material

for assembling nucleic acid nanostructures independent of RNA or DNA

recognition. We synthesized two 4-helix junctions based entirely on GNA self-pairing and showed that GNA provides easy access to highly stable nanostructures with left- and right-handed helical configurations.

5.3 Introduction

Structural DNA nanotechnology uses sequence minimization rules to construct micrometer scale objects with nanometer scale features (Seeman 1982). This is a "bottom-up" approach to nanotechnology that relies on Watson–Crick base pairing to assemble DNA motifs into different geometric shapes and patterns (For reviews, see Lin et al. 2006; Niemeyer 2000; Seeman 1998, 2005). While substantial effort has been devoted to expanding the programmability of standard B-form DNA (Feldkamp and Niemeyer 2006; Seeman 2003), considerably less attention has been given to the development of unnatural DNA nanostructures (Ng and Bergstrom 2005). By expanding DNA nanotechnology to include alternative polymers, it should be possible to create nanostructures with chemical and physical properties not found in natural DNA (For example, see Chaput and Switzer, 1999; Liu et al. 2003). Toward this goal, it would be useful to have a genetic system that was capable of forming nanostructures independent of DNA or RNA hybridization. Of the known orthogonal genetic systems (Eschenmoser 1999; Schöning et al. 2000), only glycerol nucleic acid (GNA) provides easy access to both left- and right-handed helical geometries (Yang et al. 2007; Zhang et al. 2005). This

138

unusual feature suggests that GNA could be used to construct nanostructures

with unique topologies. Here, we wish to report the first synthesis of a

nanostructure (Figure 5.1) based entirely on GNA self-pairing—two 4-helix

junctions (4HJ) with mirror image symmetry.
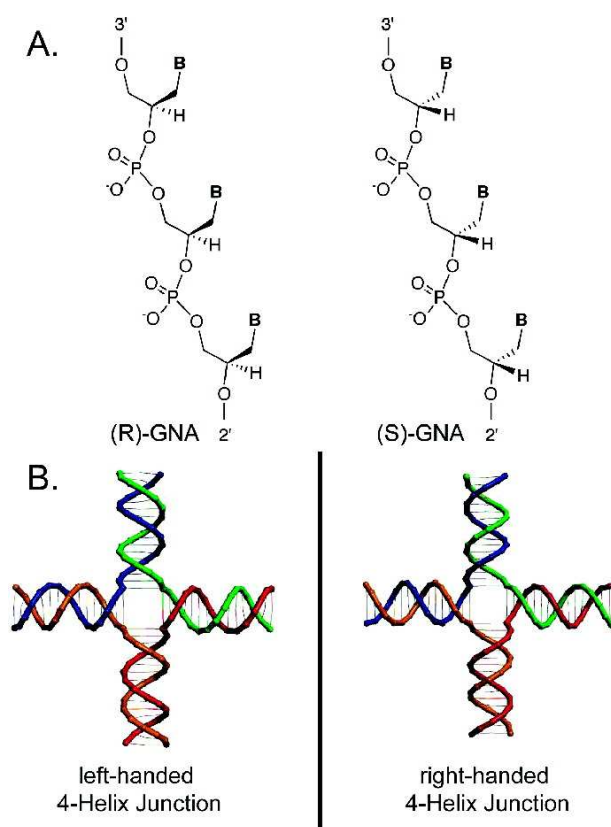
Figure 5.1. An immobile 4-helix junction composed of glycerol nucleic acid.
(A) Chemical structure of (R)- and (S)-GNA. (B) Cartoon image of left and
right-handed 4-helix junctions. Reprinted with permission from Journal of the
American Chemical Society. Copyright 2008 American Chemical Society.

5.4 Results

Using established methodology (Acevedo and Andrews 1996; Holý 1975;

Schlegel et al. 2007; Zhang et al. 2005, 2006; Yang et al. 2007), we

synthesized all four phosphoramidite monomers of (*R*)- and (*S*)-GNA. The

monomers were used as building blocks to construct mixed sequence

oligonucleotides of (*R*)- and (*S*)-GNA. To prevent unwanted base-catalyzed

decomposition during the deprotection step, we capped each GNA strand at

both ends with a single deoxyribonucleotide (Tsai et al. 2007). The sequence

for each strand (Figure 5.2A) was chosen based on a previous design by

Seeman and co-workers, where DNA was used to assemble a synthetic 4HJ

(Kallenbach et al. 1983). In contrast to natural 4HJs, which undergo branch

migration due to regions of symmetry within their sequence, synthetic 4HJs

are designed to be immobile structures (Holliday 1964). Using our GNA

oligonucleotides, we aimed to determine how well the sequence optimization

strategies used for DNA nanotechnology would apply to GNA, and whether

GNA could be used as a molecular scaffold for assembling nanostructures

with left- and right-handed helical geometries.

We began using nondenaturing gel electrophoresis to study the propensity for

GNA to self-assemble into an immobile 4HJ. In this assay, DNA complexes

are separated based on size and charge, with higher molecular weight

complexes having a slower electrophoretic mobility than smaller complexes.

On the basis of the design (Figure 5.2A), the desired 4HJ should result when

GNA strands 1−4 are present in equimolar ratios in solution. To identify the

intermediate forms of the 4HJ, we separately annealed strand 1 with strands 2,

2 + 3, and 2 + 3 + 4 and analyzed the samples by nondenaturing gel

electrophoresis. Our initial attempt at visualizing the gel by ethidium bromide

or SYBR-Gold staining methods failed to detect the presence of any GNA.
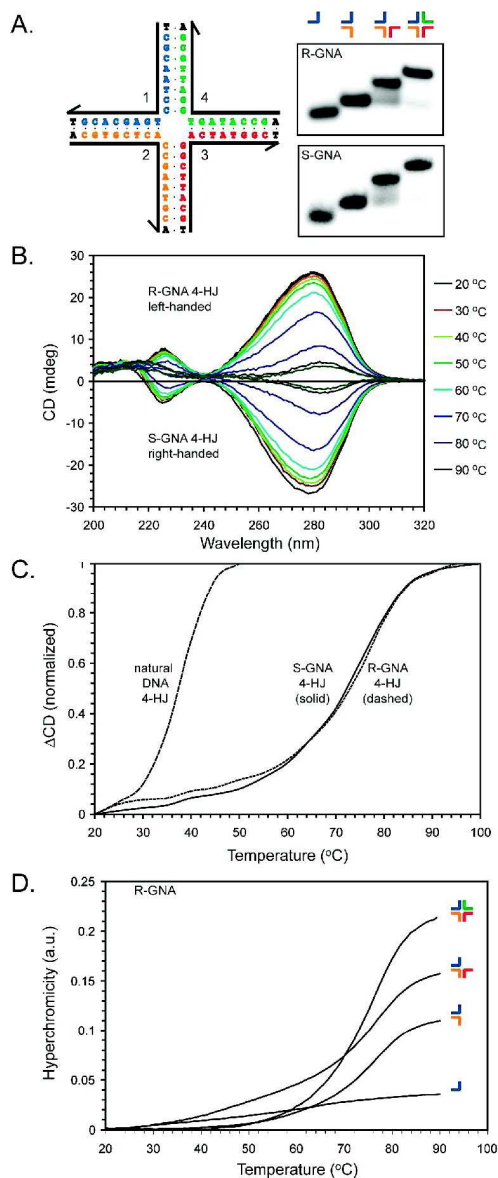
Figure 5.2. Characterization of two GNA 4-helix junctions. (A) The design used to assemble the GNA 4HJ (left). Nondenaturing gel electrophoresis mobility shift assays showing the monomer, dimer, trimer, and tetramer forms of the 4HJ (right). Drawn above each lane is a schematic view of the resulting complex. (B) CD analysis of the 4HJs derived from (R)- and (S)-GNA. (C and D) Thermal denaturing curves obtained from temperaturedependent CD and UV absorbance indicate that GNA forms a closed 4HJ that is thermally superior to the analogous DNA nanostructure. Reprinted with permission from Journal of the American Chemical Society. Copyright 2008 American Chemical Society.

Since this limitation could be due to structural differences between GNA and natural DNA and RNA, we decided to enzymatically radiolabel the DNA

residue at the 3′-end of each GNA strand with P^{32}-ATP. To our surprise, the

chimeric GNA strands proved to be remarkably good substrates for T4

polynucleotide kinase. Using P^{32}-labeled GNA, we then repeated the

strand-mixing assay and imaged the gel by phosphorimaging. The resulting

image (Figure 5.2A) shows that when adjacent strands are added to the

mixture higher order complexes emerge that exhibit mobility shifts consistent

with the formation of the dimer and trimer intermediates and final tetramer

product of the 4HJ. The same result was observed using (R)- and (S)-GNA

and natural DNA (Figure 5.3). We verified that all four strands associate in the
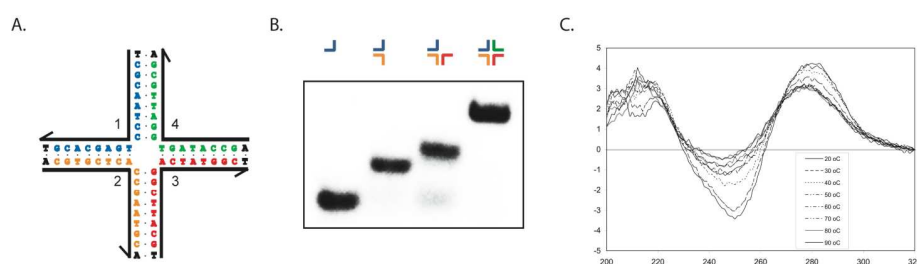
Figure 5.3. The DNA 4-Helix Junction. A) The sequence and design used to
assemble the DNA 4HJ. B) Non-denaturing gel electrophoresis image showing
complete assembly of the DNA 4HJ. C) Circular dichroism spectra showing
the thermal transition of the DNA 4HJ to single-stranded DNA. Reprinted
with permission from Journal of the American Chemical Society. Copyright
2008 American Chemical Society.

4HJ by comparing the gel mobility of each monomer, dimer, and trimer strand

combination (Figure 5.4) (Kallebach et al. 1983). Complete formation of the

tetramer structure demonstrates that GNA is capable of forming simple nucleic

acid nanostructures.

Next, we used circular dichroism (CD) spectroscopy to examine the

structural properties of our GNA nanostructures. The CD spectra of both 4HJs

are given in Figure 5.2B. The two nanostructures exhibit mirror image

symmetry with strong CD signals at 224 and 280 nm. This result is consistent

with previous CD data collected on double-stranded GNA duplexes (Zhang et

al. 2005) and supports the formation of short helices in each arm of the 4HJ.
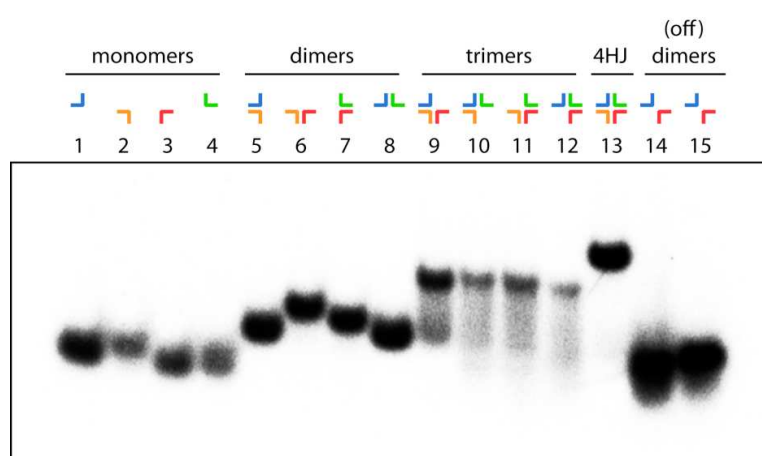
Figure 5.4. Non-denaturing polyacrylamide gel electrophoresis of GNA
strands and strand mixtures. Lanes 1-4 contain GNA strands 1-4, respectively;
Lanes 5-8 contain strand dimer combinations 1-2, 2-3, 3-4, and 1-4,
respectively; Lanes 9-12 contain strand trimer combinations 1-2-3, 4-1-2,
2-3-4, and 3-4-1, respectively; Lane 13 contains strands 1-2-3-4 that form the
complete 4-HJ; and Lanes 14 and 15 contain the non-hybridizing dimers 1-3
and 2-4. Reprinted with permission from Journal of the American Chemical
Society. Copyright 2008 American Chemical Society.

The CD spectra for both GNA nanostructures differ considerably from natural

DNA and RNA, indicating that GNA likely adopts a unique helical structure.

This hypothesis is supported by the observation that GNA cross-pairs weakly

with RNA but not all with DNA (Yang et al. 2007; Zhang et al. 2005).

Analysis of the CD spectra at different temperatures revealed a Cotton effect

that decreases with increasing temperature, as expected for a self-assembled

complex with a melting transition.

Last, we examined the thermal stability of the GNA 4HJ using

temperature-dependent spectroscopy. Quantitative values for the thermal

stability of the 4HJs were obtained by monitoring the change in the CD signal

at 280 nm as a function of increasing temperature. As illustrated in Figure

5.2C, 4HJs constructed from (R)- and (S)-GNA gave nearly identical curves

with a single transition between the folded and denatured states. The data from

both melting experiments were fit to a standard sigmoidal curve, which gave a

T_m of 76 °C for the two GNA structures. Comparison of the hyperchromicity

(Figure 5.2D) obtained for UV melts on strands 1, 1 + 2, 1 + 2 + 3, and 1 + 2 +

3 + 4 reveals a stepwise increase in absorbance consistent with a closed 4HJ

(Kallenbach et al. 1983)

To compare the thermal stability of a 4HJ constructed of GNA to one

constructed of DNA, we repeated the thermal denaturation experiment using

DNA oligonucleotides of identical sequence. As expected, the DNA version of

the 4HJ (Figure 5.2C) gave a melting transition of 37 °C, which is consistent

with the denaturing plot given in the original description of the 4HJ by

Seeman (Kallenbach et al. 1983). On the basis of these data, it appears that the

GNA 4HJ is significantly more stable than the DNA 4HJ. This is a remarkable

achievement given that the GNA backbone is acyclic and one atom shorter

than the backbone found in DNA and RNA. In the absence of detailed

144

structural information, it is difficult to rationalize the extreme thermal stability

of GNA. One possibility is that the flexible nature of the acyclic backbone

allows GNA to access base pairing and base stacking interactions that are

optimal for two complementary strands. DNA and RNA, though optimal in

many other ways, would then have pairing interactions that are weaker due to

structural constraints imposed by the furanose ring.

5.5 Conclusion

In summary, we describe the use of glycerol nucleic acid as a simple

building block for assembling nucleic acid nanostructures with left- and

right-handed helical geometries. We suggest that GNA could be used as a

general method for constructing chiral nanostructures with unique topologies

and extreme thermal stabilities. Once the helical dimensions of GNA are

known, we predict that the sequence optimization methods used to program

DNA will also apply to GNA since both polymers self-assemble using the

same Watson–Crick base pairing rules. The current study therefore represents

a first step toward the development of unnatural nucleic acid nanostructures.

5.6 Experimental Design

Materials and Methods

GNA Synthesis. (R)- and (S)-GNA phosphoramidite monomers were

synthesized for all eight nucleotides using standard literature methods

(Acevedo and Andrews 1996; Holý 1975; Schlegel et al. 2007; Zhang et al.

2005, 2006; Yang et al. 2007). GNA oligonucleotides were constructed using

an Applied Biosystems 3400 DNA Synthesizer. GNA synthesis was

performed in the trityl-off mode on 0.2 μmol scale with the coupling times

extended to 300 sec. The GNA oligonucleotides were deprotected in

concentrated NH4OH at 55°C for 72 h and purified by denaturing

polyacrylamide gel electrophoresis. The concentration of each strand was

determined by optical density using extinction coefficients taken from natural

DNA nucleotides. The identity of each GNA strand was confirmed by a

MALDI-TOF mass spectrometry (Table 5.1).

Table 5.1. MALDI-TOF MS Analysis of GNA Oligonucleotides. Reprinted
with permission from Journal of the American Chemical Society. Copyright
2008 American Chemical Society.

Entry 1	Sequence*	Calculated	Observed
Strand 1	3'-**T**CGCAATCCTGAGCACG**T**-2'	4787.80	(R)- 4787.48
			(S)- 4787.88
Strand 2	3'-**A**CGTGCTCACCGAATGC**A**-2'	4796.81	(R)- 4797.06
			(S)- 4797.39
Strand 3	3'-**T**GCATTCGGACTATGGC**T**-2'	4833.82	(R)- 4834.86
			(S)- 4834.33
Strand 4	3'-**A**GCCATAGTGGATTGCG**A**-2'	4891.86	(R)- 4892.80
			(S)- 4892.44

Circular Dichroism. CD spectra were collected between 320 to 200 nm using a

JASCO J-815 CD spectrophotometer. Each solution contained all four single-

stranded GNA oligonucleotides in a 1:1:1:1 ratio (2.5 μM each) in a buffer

containing 1 M NaCl, 0.5 M Na_2HPO_4, and 10 mM EDTA (pH 7.0). The

strands were annealed prior to each run by heated the solution to 90°C for 5

min and slowly cooled to 4°C. Thermally induced transitions of each mixture

146

were monitored at 280 nm for (R)-GNA and (S)-GNA 4HJs in the forward and

reverse directions. Thermal denaturation studies involving the DNA 4HJ were

performed by monitored the CD signal at 250 nm. The sample temperature

was increased at a rate of 0.5ºC per minute, and the data was collected for 1

sec at 1 nm intervals in a quartz crystal cuvette with 1 cm path-length.

UV Absorbance. UV absorbance versus temperature profiles were measured

on a Varian Cary 300 Bio UV-visible spectrophotometer in a temperature-

controlled cell holder. All melts were performed under standard conditions

with 1.0 M NaCl, 10 mM NaH_2PO_4, 0.1 mM NaEDTA, pH 7.0 and 1:1 strand

stoichiometry (2.5 μM per strand). Melting curves were obtained in the

reverse and forward melting directions with temperature gradient of 20º to

90ºC and a ramping speed of 1°C per min by monitoring the change in UV

absorbance at 260 nm at each temperature. T_m values were determined as the

first-order derivative using the Cary software package, and represent the

average of three independent trials.

Non-denaturing Polyacrylamide Gel Electrophoresis. Formation of the 4HJ

was analyzed by non-denaturing polyacrylamide gel electrophoresis. The

GNA Strands were P^{32}-radiolabeled by incubating γ-^{32}P-ATP with T4

polynucleotide kinase at 37ºC for 1 h. Individual complexes composed of

strand 1 alone, strands 1+2, 1+2+3 and 1+2+3+4 were prepared in equimolar

ratios and incubated in 20 μL of TAE-Mg buffer (20 mM Tris, pH 7.6, 2 mM

EDTA, 12.5 mM $MgCl_2$). The final concentrations for each strand was 5.0 μM.

Each complex was annealed, mixed with 2 μL of 50% glycerol, and analyzed

by non-denaturing gel electrophoresis and imaged by phosphorimaging.

5.7 Reference

Acevedo, O.L., and R.S. Andrews. 1996. Synthesis of propane-2,3-diol combinatorial monomers. *Tetrahedron Lett.* 37: 3913-3934.

Chaput, J.C., and C. Switzer. 1999. A DNA pentalex incorporating nucleobase quintets. *Proc. Natl. Acad. Sci. USA* 96: 10614-10619.

Eschenmose, A. 1999. Chemical etiology of nucleic acid structure. *Science* 284: 2118-2124.

Feldkamp, U., and C.M. Niemeyer. 2006. Rational design of DNA nanoarchitectures. *Angew. Chem. Int. Ed.* 45: 1856-1876.

Holliday, R. 1964. A mechanism for gene conversion in fungi. Genet. Res. 5: 282-304.

Holý, A. 1975. Aliphatic analogues of nucleosides, nucleotides, and oligonucledotides. *Collect. Czech. Chem. Commun.* 40: 187-214.

Kallenbach, N.R., R.-I. Ma, and N.C. Seeman. 1983. An immobile nucleic acid junction constructed from oligonucleotides. *Nature* 305: 829-831.

Lin, C., Y. Liu, S. Rinker, and H. Yan. 2006. DNA tile based self-assembly: building complex nanoarchitectures. *Chemphyschem* 7: 1641-1647.

Liu, H., J. Gao., S.R. Lynch, Y.D. Saito, L. Maynard, and E.T. Kool. 2003. A four-base paired genetic helix with expanded size. *Science* 302: 868-871.

Niemeyer, C.M. 2000. Self-assembled nanostructures based on DNA: towards the development of nanobiotechnology. *Curr. Opin. Chem. Biol.* 4: 609-618.

Ng, P.-S., and D.E. Bergstrom. 2005. Alternative nucleic acid analogues for programmable assembly: hybridization of LNA to PNA. *Nano Lett.* 5: 107-111.

Schlegel, M.K., A.E. Peritz, K. Kittigowittana, L. Zhang, and E. Meggers. 2007. Duplex formation of the simplified nucleic acid GNA. *ChemBioChem* 8: 927-932.

Schöning K-U, P. Scholz, S. Guntha, X. Wu, R. Krishnamurthy, and A. Eschenmoser. 2000. Chemical etiology of nucleic acid structure: The α-ʟ-threofuranosyl-(3′→2′) oligonucleotide system. *Science* 290: 1347–1351.

Seeman, N.C. 1982. Nucleic acid junctions and lattices. *J. Theor. Biol.* 99: 237-247.

Seeman, N.C. 1998. DNA nanotechnology: novel DNA constructions. *Annu. Rev. Biophys. Biomol. Struct.* 27: 225-248.

Seeman, N.C. 2003. At the crossroad of chemistry, biology, and materials: structural DNA nanotechnology. *Chem. Biol.* 10: 1151-1159.

Seeman, N.C. 2005. From genes to machines: DNA nanomechanical devices. *Trends Biochem. Sci.* 30: 119-125.

Tsai, C.-H., J. Chen, and J.W. Szostak. 2007. Enzymatic synthesis of DNA on glycerol nucleic acid templates without stable duplex formation between product and template. *Proc. Natl. Acad. Sci. USA 104*: 14598-14603.

Yang, Y.-W., S. Zhang, E.O. McCullum, and J.C. Chaput. 2007. Experimental evidence that GNA and TNA were not sequential polymers in the prebiotic evolution of RNA. *J. Mol. Evol.* 65: 289-295.

Zhang, L., A. Peritz, and E. Meggers. 2005. A simple glycol nucleic acid. *J. Am. Chem. Soc.* 127: 4174–4175.

Zhang, L., A.E. Peritz, P.J. Carroll, and E. Meggers. 2006. Synthesis of glycol nucleic acids. *Synthesis* 4: 645-653.

Chapter 6

IN VITRO EVOLUTION PROVIDES EVIDENCE FOR TNA AS A

FUNCTIONAL PROGENITOR OF RNA

6.1 Contributions

The following chapter describes an artificial genetic system based on

threose nucleic acid (TNA) capable of Darwinian evolution. The reported

experiments were performed by Su Zhang and Hanyang Yu under the guidance

of Professor John Chaput. Su Zhang synthesized TNA triphosphates. Hanyang

Yu and Su Zhang constructed the TNA library and designed the selection

strategy. Hanyang Yu completed the TNA aptamer selection. Hanyang Yu and

Su Zhang characterized the aptamers. John Chaput finished the manuscript

with comments from Su Zhang and Hanyang Yu. The results from this work

were submitted to *Nature Chemistry*.

6.2 Abstract

The RNA world hypothesis postulates that contemporary life evolved from

a simpler form of life that was based entirely on RNA, but it is not known

whether other genetic systems could perform this function. Presumably

whatever prebiotic chemistry gave rise to RNA would have produced other

RNA analogues, some of which may have preceded or competed directly with

RNA. Threose nucleic acid (TNA), an unnatural derivative of RNA, has

received considerable interest as a possible RNA progenitor due to its

chemical simplicity and ability to exchange genetic information with itself and RNA. Here, we demonstrate that TNA can fold into stable three-dimensional shapes that elicit a complex function. We applied the principles of Darwinian evolution to evolve a TNA receptor that binds to an arbitrary target with high affinity and specificity. This observation confirms that TNA could have functioned as an ancestral genetic system during an early stage of life.

6.3 Introduction

Systematic analysis of nucleic acid structure with regard to the chemical etiology of RNA reveals that several potentially natural RNA alternatives are capable of Watson-Crick base pairing (Eschenmose 1999). While this property demonstrates a capacity to store genetic information, primordial genetic systems would have needed to fold into stable tertiary structures with sophisticated functions like ligand binding and catalysis in order to establish a primitive metabolism (Benner et al. 1989). Determining which, if any, RNA analogues can fold into shapes that elicit a specific pre-defined function would constrain models that describe the origin and early evolution of life on Earth. Of those genetic systems studied thus far, (3',2')-α-L-threose nucleic acid (TNA, Figure 6.1) has generated considerable interest as a primordial genetic system because of the chemical simplicity of threose relative to ribose and the ability of TNA to form stable helical structures with complementary strands of itself and RNA (Ebert et al. 2008; Orgel 2000; Schöning et al. 2000). This

151

latter feature provides a plausible mechanism for the transfer of information

between successive genetic systems and suggests an evolutionary path where

life could have entered the RNA world.
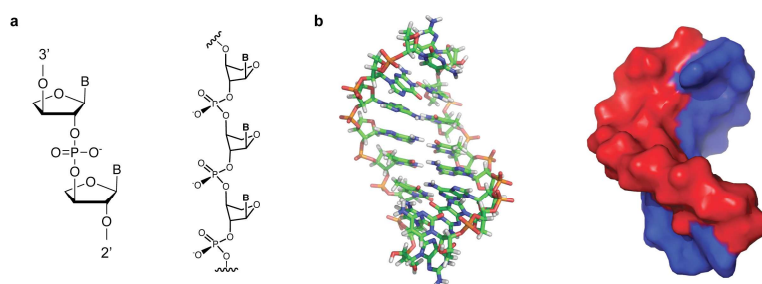
Figure 6.1. Structure of threose nucleic acid. a, Constitutional (left) and configurational (right) structures for the linearized backbone of an α-L-threofuranosyl-(3'→2')-oligonucleotide. TNA contains one less atom per backbone repeat unit than natural RNA and DNA. b, Solution NMR structure for the duplex formed from the self-pairing complementary sequence 3'-CGAATTCG-2' (Ebert et al. 2008). Duplex structures were rendered in PyMol with models showing ball and stick (left) and surface (right) images. TNA adopts a right-handed structure with helical parameters similar to A-form RNA.

Molecular evolution provides a powerful approach to investigate the

functional properties of nucleic acids (Joyce 2007; Wilson and Szostak 1999).

This method has traditionally been limited to DNA and RNA because these are

the only polymers with enzymes that are available to transcribe, reverse-

transcribe and amplify genetic information. Extending this approach to

artificial genetic systems like TNA requires (i) polymerases that can transcribe

a library of random DNA sequences into TNA; (ii) a method of selection that

allows functional members to be separated from the non-functional pool; and

(iii) a strategy to recover and amplify functional molecules that maintains the

integrity of the selected sequence.

Creating an artificial genetic system capable of Darwinian evolution is one of the grand challenges in synthetic biology, and an important step in the chemical path to synthetic life (Benner and Sismour 2005). Inspired by the potential to explore the functional properties of an alternative genetic system, studies were undertaken by our lab and others to identify DNA polymerases that could recognize TNA either as a template or as a nucleoside triphosphate (Chaput et al. 2003a, 2003b; Kempeneers et al. 2003). This work showed that several DNA polymerases can synthesize short segments of DNA on a TNA template (Chaput et al. 2003a), and other polymerases can synthesize limited stretches of TNA on a DNA template (Chaput et al. 2003b; Kempeneers et al. 2003). While the enzymes identified in these studies had activity that was too low to support in vitro selection; further screening of additional polymerases eventually led to the discovery of therminator DNA polymerase (Horhota et al. 2005; Ichida et al. 2005b). Therminator is an engineered variant of the 9°N DNA polymerase that functions as a remarkably efficient DNA-dependent TNA polymerase. Under optimal conditions, therminator can transcribe an 80-nt DNA template into TNA with high efficiency and fidelity (Ichida et al. 2005a). Since natural polymers of this length can fold into shapes with defined ligand binding sites, it was postulated that therminator could be used to generate pools of TNA molecules for in vitro selection (Ichida et al. 2005b). However, subsequent studies revealed that existing conditions did not support the synthesis of full-length TNA libraries (Ichida et al. 2005a).

153

6.4 Results

To identify conditions that were better suited for TNA synthesis, we

developed a program of research to overcome the limitations of

therminator-mediated TNA synthesis. Using chemical methods, we

synthesized all four TNA triphosphates (Figure 6.2a) and used these molecules

as substrates in primer-extension assays. A synthetic DNA primer was

Figure 6.2 Synthesis of TNA libraries by enzyme-mediated primer extension. a,
Chemical structure of TNA triphosphates: tDTP 1; tTTP 2; tGTP 3; and tCTP
4. b, Schematic diagram of the primer extension assay used to evaluate DNA
libraries L1-L3. Each library contains a central random region of 50-nts
flanked on either side by a 20-nt constant region. Library compositions: L1
equal distribution of A, C, G, and T; L2 equal distribution of A, C, and T; and
L3 one-half equivalent of G relative to A, C, and T. c, Therminator-mediated
TNA transcription assays analyzed by denaturing polyacrylamide gel
electrophoresis. Primer extension of L1 with tNTPs 1-4 yields only trace
amounts of full-length product (left panel). Primer extension of L1 using
defined combinations of dNTPs (black) and tNTPs (red) leads to full-length
product when tCTP is replaced with dCTP (center panel). Primer extension
across libraries L2 and L3 leads to full-length product in ~60% and ~30%
yield, respectively (right panel). M denotes DNA marker.

annealed to a synthetic DNA library (L1) that contained a central random

region of 50-nts followed by a second primer-binding site that could be used

154

for reverse transcription or PCR. Theminator DNA polymerase was challenged to copy the DNA library into TNA by extending the DNA primer with 70 sequential TNA residues (Figure 6.2b). The diaminopurine analogue of adenine was used in place of adenosine 3'-triphosphate, because this substitution is known to increase the thermo- dynamic stability of the TNA-DNA heteroduplex (Chaput et al. 2003a; Wu et al. 2002). As with earlier trials, attempts to copy the unbiased random- sequence DNA library into TNA failed under all conditions tested (Figure 6.2c, Ichida et al. 2005a). Varying the incubation time, temperature, and salt proved ineffective, suggesting that unbiased random-sequence DNA libraries likely contain sequence motifs such as repeating nucleotides or secondary structures that impede primer-extension activity.

To determine which sequence elements in the DNA library inhibit TNA synthesis, a set of primer extension assays was performed using defined mixtures of tNTP and dNTP substrates. This assay allowed TNA residues with elongation problems to be identified since substitution of the TNA substrate for the natural residue should favored strand elongation. The result from our screen of nearly all possible nucleotide combinations is shown in Figure 6.2c. Full-length primer extension was observed in all cases, except when tCTP was used in place of dCTP (Figure 6.2c). This finding suggests that repeating G nucleotides in the DNA library cause therminator to pause during TNA transcription. Since unbiased DNA libraries like L1 would by chance alone

155

have an abundance of short G-repeats, only a small fraction of DNA templates would promote full-length TNA synthesis. Indeed, analysis of representative clones from our starting library showed an abundance of G3 and G4 repeats.

To design DNA libraries that were better suited for TNA transcription, we synthesized DNA libraries L2 and L3 that minimized the occurrence of G-residues in the random region. L2 contained equal amounts of A, C, and T, while L3 limited the frequency of G nucleotides to 50% the occurrence of A, C, and T. Consistent with the design strategy, primer extension assays performed on the L2 and L3 libraries (Figure 6.2c) yield substantial amounts of the fully extended DNA primer (~60% and ~30%, respectively). In both cases, therminator DNA polymerase extended the DNA primer with 70 TNA nucleotides, demonstrating for the first time that combinatorial TNA libraries could be generated by enzyme-mediated extension of a DNA primer.

To evaluate the fitness of TNA as a primordial biopolymer, we employed an in vitro selection strategy to enrich for TNA molecules with receptor activity. The approach taken is analogous to other display technologies like mRNA display and DNA display that provide a covalent link between genotype (DNA) and phenotype (TNA) (Roberts and Szostak 1997; Rosenbaum and Liu 2003). The following strategy was devised by Ichida and Szostak but never implemented due to the inefficiency of TNA synthesis (Ichida et al. 2005b). In this scheme, a single-stranded DNA library is constructed with a stable stem-loop structure that serves as a DNA primer for

the extension of TNA across the random region (Figure 6.3a). Extension of the

primer leads to the formation of a chimeric TNA-DNA hairpin duplex. The

TNA portion of the duplex is displaced by extending a separate DNA primer

annealed to the stem-loop region of the hairpin with DNA. The product of the

strand displacement is a TNA molecule that is physically linked to its own

double-stranded DNA message. To enrich for TNA molecules with specific

ligand binding activity, the pool of TNA-DNA fusions is incubated with a
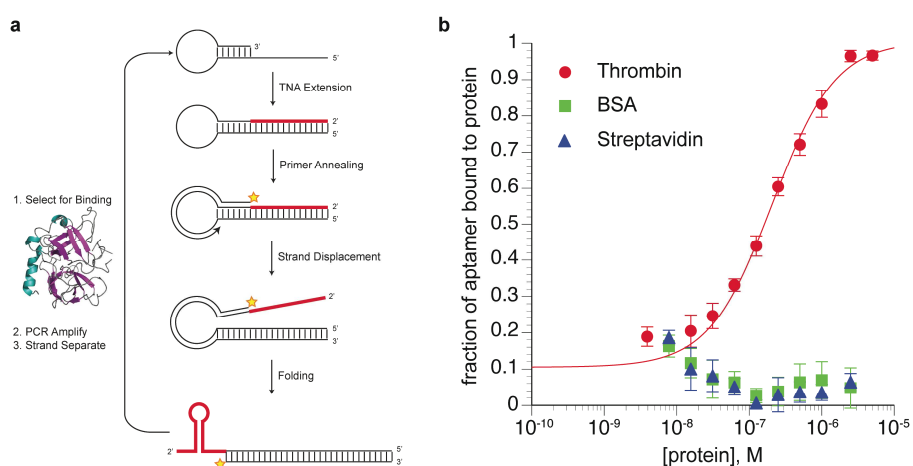
Figure 6.3. Evolution of TNA receptors in vitro. a, In vitro selection strategy
designed to isolate TNA aptamers with affinity to human thrombin. The DNA
library encodes a random region of 50-nt positions flanked on the 3' end with
a stem-loop structure that serves as a DNA primer and a fixed-sequence
primer-binding site located at the 5' end. The DNA primer is extended with
tNTPs to produce a chimeric TNA-DNA hairpin. A separate DNA primer
modified with 6-carboxy-fluorescein (star) is annealed to the stem-loop region
and extended with DNA to displace the TNA strand. The resulting pool of
TNA-DNA fusion molecules is incubated with the protein target. Bound
molecules are separated from the unbound pool by capillary electrophoresis
and amplified by PCR. The dsDNA is made single-stranded and the coding
strand is annealed to generate a new pool of DNA templates for the next
selection cycle. b, Equilibrium binding affinity measurement of the core
binding domain of TNA aptamer 3.12. The aptamer
(3'-TGTTDTDGDDDDDDTDDTGGTGGGGGGGTTTDGDTDDDGGGG-2')
binds to human thrombin with a K_d of 200 nM and shows no detectable
affinity for bovine serum albumin (BSA) or streptavidin. Error bars represent
the standard deviation of each data point (n = 3).

157

desired target and bound sequences are recovered and amplified by PCR. The DNA is then denatured into single-stranded form that folds into a stable stem-loop structure to facilitate another selection cycle.

Successive rounds of in vitro selection and amplification were performed using the L2 library, which gave the highest yield of full-length TNA product in our primer extension assay. Although the L2 library generates TNA polymers that lack cytidine, we reasoned that this was not a significant concern since cytidine was presumably not present in the first genetic material due to its tendency to undergo spontaneous deamination ($t_{1/2}$ = 340 yr at pH 7 and 25 °C) (Levy and Miller 1998). Furthermore, it has been shown that ribozymes missing cytidine can be generated by in vitro evolution, demonstrating that a three letter genetic alphabet retains the ability to fold and function (Rogers and Joyce 1999). Using the strategy described above, we selected for TNA molecules that could bind to human thrombin and therefore exhibit an arbitrary specific function. We chose capillary electrophoresis as our separation technique because it leads to a higher partitioning efficiency than is commonly observed for traditional gravity filtration methods (Lin et al. 2007; Williams et al. 2009). TNA-DNA fusion molecules carried the fluorescent label 6-carboxy-fluorescein to facilitate detection by laser-induced fluorescence. For each selection cycle, five 70 nL portions of library (~10^{11} TNA-DNA fusion molecules) were injected onto a neutral coated capillary. Functional sequences were recovered by allowing the unbound pool to pass

into a waste vial and applying pressure to the column to isolate the bound

molecules in a separate vial. After three rounds, a slower migrating peak

corresponding to TNA-DNA fusions bound to thrombin was visible (Figure

6.4), indicating that the pool had become enriched in TNA aptamers with

affinity to thrombin.
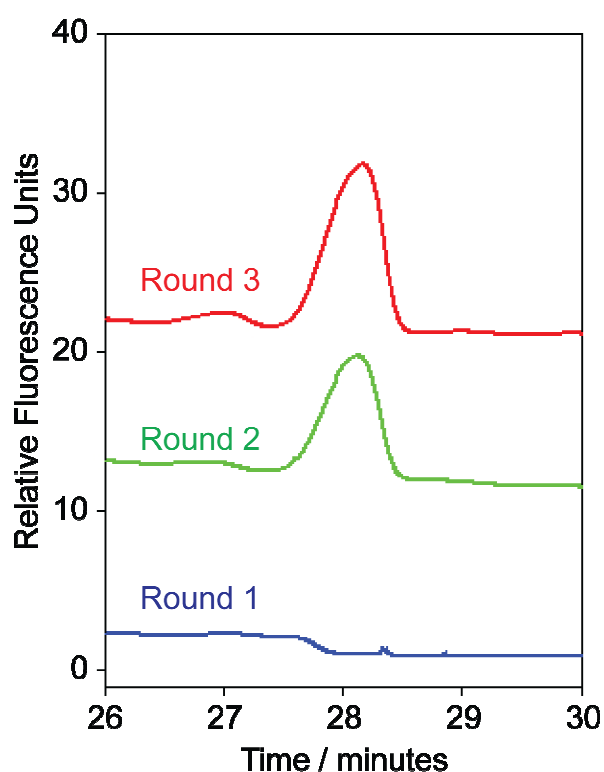
Figure 6.4. Enrichment of functional TNA molecules by in vitro selection. A peak corresponding to the bound thrombin-TNA aptamers became visible in capillary electrophoresis electropherogram after iterative rounds of in vitro selection and amplification.

We cloned and sequenced the DNA to determine the diversity of molecules

that remained in the pool. Several clones contained isolated G residues,

indicating that a low level of random mutagenesis had occurred during the

selection. It is difficult to determine whether these mutations arose during

159

PCR or strand displacement, but their presence indicates a further level of diversification. Ten representative sequences were chosen for individual analysis. The TNA molecules were synthesized by extending a DNA primer annealed to a synthetic DNA template with TNA and purifying the DNA-tagged TNA molecules by denaturing gel electrophoresis. We noticed that the representative sequences transcribed with much greater efficiency than the starting pool (1 versus 10 hours), indicating that the selection favored the enrichment of DNA templates that transcribe with high efficiency. The resulting TNA aptamers have equilibrium dissociation constants (Kd) for thrombin that fall within the range of 200 to 900 nM (Table 6.1), which is

Table 6.1. Sequence and Binding Affinity for TNA Aptamers.

Clone	Sequence (3'—2')	K_d (nM)
3.12	TGTTDTDGDDDDDDTDDTGGTGGGGGGTTTDGDTDDDGGGGTDTGDGDG	190
3.14	GGGGDGTGTGTCGGDDGTTGTGGGTTDGDTTDTTDGTGDDGTGDTGDDGG	250
3.9	DDGDDTTGGGGGDTTGGTDGGGDGDGDDGDGGTTDDDGTDDGGTDTDTD	320
3.11	TTGTTDDGTDTGDTGGTTTDTGTDDGGTGGDTDTTTTGDGTTGDGTGGGG	420
3.1	TTTTTGDTGTGGDGTCTGTTGGDTTGGDDGTTGTTTTTGTGDDGGGG	510
3.2	DGTGGDGTGTTDGTDDTGGTTTTDDGTGGGGTGDGDTGGGDDTTTGDTTT	590
3.10	GDTTGDGTGGDGDTGDTTTDDTTTGDGGTGTGTTGGDTTGTTGGG	590
3.5	DTGGTGDTTTGTTTDTGTGDDCDTGDTGTDTGDGDTTTTGGGGTGTTTTG	750
3.13	TGDGTDGDTTGGGDDTGTGTGDTTTGGGDTGTGTGGGGGTDDDTTDG	830
3.6	TGGGTDTDTGTTDGGDTDTDTGTTDDTTTDTDGGTTTGTGDGTGGTTTDG	880

similar to previously evolved DNA and RNA aptamers that bind to thrombin (Bock et al. 1992; Kubik et al. 1994). We determined the minimum binding region for the highest affinity aptamer (clone 3.12, Kd ~200 nM at 25 °C) by deletion analysis. TNA variants were generated by primer extension using DNA templates that contain incremental deletions from both ends of the

160

coding region. We compared the Kd for each deletion construct to the parent

sequence, which defined a core boundary of 41 nucleotides that retained high

affinity binding (Figure 6.3b, 6.5). The affinity of the aptamer does not change

when the DNA primer is exchanged for a different sequence and a randomly

generated TNA sequence has no affinity to thrombin (data not shown). To

Figure 6.5. Truncation analysis of TNA aptamer 3.12. Comparing the affinity
of six deletion constructs with the parent sequence defined a core region of 41
nucleotides (3.12-2t8) that retained high affinity binding to human thrombin.

demonstrate that clone 3.12 folds into a structure with a ligand binding site

that is complementary to the surface of human thrombin, we measured the

solution binding affinity of clone 3.12 to two common proteins, bovine serum

albumin (BSA) and streptavidin. In both cases, the TNA molecule shows no

detectable affinity for the off-target protein (Figure 6.3b), indicating that the

evolved aptamer binds to thrombin with high specificity.

The idea that life on Earth arose during an ancestral period in which RNA

stored genetic information and catalyzed metabolically relevant chemical

reactions was discussed extensively over 40 years ago (Crick 1968; Orgel

1968; Woese 1967), and later termed the RNA world hypothesis (Gilbert

1986). While the chemical plausibility of the RNA world remains strong (Cech

2011), it is not clear that RNA was the first genetic material (Joyce 2002).

Problems associated with the prebiotic synthesis of ribose and non-enzymatic replication of RNA have fueled interest in the possibility that RNA was preceded in the evolution of life by a genetic system that was more accessible on the primitive Earth (Joyce et al. 1987; Joyce 1991). Determining which genetic systems are both simpler than RNA and capable of Darwinian evolution is a daunting challenge. Chemical synthesis is required to build substrates that are not otherwise available and enzymes must be discovered that can make unnatural polymers with high efficiency and fidelity. In the case of TNA, for example, a decade-long effort was required to establish the methodology necessary to evolve TNA molecules in vitro. While new advances in protein engineering are making it easier to create polymerases that recognize unnatural substrates (Loakes and Holliger 2009), future examination of structurally diverse genetic systems by in vitro evolution will require similar advances in nucleic acid chemistry.

6.5 Conclusion

In summary, we have applied the principles of Darwinian evolution to evolve TNA molecules that fold into shapes that bind to an arbitrary protein target and therefore exhibit a complex function.   This observation coupled with the chemical simplicity of threose relative to ribose indicates that TNA could have functioned as an ancestral genetic system during a stage of life that preceded RNA. The best TNA aptamer identified in our selection binds to

162

human thrombin with activity similar to previously evolved RNA aptamers (Bock et al. 1992; Kubik et al. 1994), indicating that the chemical property of ligand binding is not unique to the DNA and RNA polymers. This fact was not clear at the beginning of our study and indeed one could imagine that purely chemical constraints, such as the shorter repeating backbone unit found in TNA, might preclude the ability of certain nucleic acid polymers to adopt folded structures that function in a desired pre-defined way. The fact that TNA is not limited in this regard suggests that it may be possible to isolate novel TNA enzymes from pools of random sequences using in vitro evolution. Such selections provide a rich opportunity to examine the fitness for TNA as an RNA progenitor in a hypothetical TNA world.

6.6 Methods

Genereal information

    Solvents (ACS reagent) for extraction and column chromatography were purchased from EMD chemicals. Solvents and reagents for reactions were purchased from Fluka and Sigma-Aldrich. All anhydrous solvents were dried and distilled freshly and stored over activated molecular sieves under Ar. Reactions were monitored by TLC on Silica G w/UV 254 aluminum plates (Sorbent Tech) visualized by UV absorption or by dipping into a solution of anisaldehyde/H_2SO_4/AcOH/EtOH 5:5:1:18 followed by heating. Solvents for NMR were purchased from Cambridge Isotope Laboratories. TMS was used

as internal standard in ^{1}H-NMR and 85% H_3PO_4 was used as external standard in ^{31}P-NMR. MALDI-TOF MS was performed on a Voyager-Elite mass spectrometry (Applied Biosystems) with THAP or DHB as the matrix. Ion exchange chromatography was performed using a customized DEAE Sephadex column (EssentialLife solutions) on a GE Pharmacia AKTA FPLC system with UV detection eluting with 0-1 M TEAB pH 8.0. Reverse HPLC purification was performed using a ZORBAX C-18 reversed-phase HPLC column (Agilent Technologies) on an Agilent 1100 HPLC system with UV detection eluting with 95% 100 mM TEAA pH 7.0 and a linear gradient from 5% methanol to 30% methanol.

N^4-Benzoyl-1-(2'-O-acetyl-α-L-threofuranosyl)cytosine

2.0 g (3 mmol) of N^4-Benzoyl-1-{2'-O-acetyl-3'-O-[(4",4"'-Dimethoxy-triphenyl)methyl]-α-L-threofuranosyl}cytosine was treated with 100ml of 3% TCA in CH_2Cl_2 for 3min. MeOH was added dropwise to quench DMT cation until the orange color disappeared. The acid was neutralized by ca. 40 ml of Et_3N, and the mixture was evaporated under reduced pressure immediately. The residue was adsorbed onto silica gel and purified by column chromatography (CH_2Cl_2/MeOH 50:1) to give the product (1g, 2.8 mmol) as a white powder in 93% yield. TLC (CH_2Cl_2/MeOH 25:1): R_f 0.23. ^{1}H-NMR (400 MHz, DMSO-d_6): δ=2.10 (s, 3H), 4.16 (m, 2H), 4.25 (d, J=9.2, 1H), 5.10 (s, 1H), 5.70 (d, J=2.4, 1H), 5.79 (s, 1H), 7.36 (d, J=7.2, 1H), 7.50 (m, 2H), 7.62 (m, 1H), 7.99 (m, 2H), 8.13 (d, J=8.0, 1H), 11.24 (s, 1H).

9-{2'-O-[(4",4"'-dimethoxytriphenyl)methyl]-α-L-threofuranosyl}guanine

1.5g (1.9 mmol) of N^2-Acetyl-O^6-diphenylcarbamoyl-9-{2'-O-[(4",4"'-dimethoxytriphenyl)methyl]-α-L-threofuranosyl}guanine was dissolved in 8M methylamine in ethanol: 12M methylamine in water (1:1) for 3 h at 50℃ followed by solvent removal. The residue was purified by column chromatography (column deactivation with CH_2Cl_2/Et_3N 100:1, gradient: CH_2Cl_2/Et_3N 100:1 to CH_2Cl_2/MeOH/Et_3N 100:4:1) to give the product (1g, 1.8 mmol) as a white powder in 95% yield. ^{1}H-NMR (500 MHz, $CDCl_3$): δ=3.71 (2s, 6H), 4.05 (d, J=9.5, 1H), 4.14 (m, 2H), 4.64 (s, 1H), 5.34 (s, 1H), 6.59 (br, 2H), 6.71-6.76 (m, 4H), 7.10 (br, s, 1H), 7.17-7.41 (m, 10H).

9-{2'-O-[(4",4"'-dimethoxytriphenyl)methyl]-α-L-threofuranosyl}-2,6-diaminopurine

1.5g (1.9 mmol) of N^2,N^6-Dibenzoyl-9-{2'-O-[(4",4"'-dimethoxy-triphenyl)methyl]-α-L-threofuranosyl}-2,6-diaminopurine was dissolved in 8M methylamine in ethanol: 12M methylamine in water (1:1) for 3 h at 50℃ followed by solvent removal. The residue was purified by column chromatography (column deactivation with CH_2Cl_2/Et_3N 100:1, gradient: CH_2Cl_2/Et_3N 100:1 to CH_2Cl_2/MeOH/Et_3N 100:4:1) to give the product (1g, 1.8 mmol) as a white powder in 95% yield. ^{1}H-NMR (400 MHz, $CDCl_3$): δ=3.77 (2s, 6H), 4.05 (d, J=9.6, 1H), 4.17 (dd, J=9.6, 3.6, 1H), 4.27 (m, 1H), 4.67 (d, J=5.6, 1H), 5.21 (d, J=1.2, 1H), 5.42 (br, s, 2H), 6.72 (br, 2H), 6.72-6.79 (m, 4H), 7.22-7.41 (m, 9H), 7.58 (s, 1H).

(α-ʟ-threofuranosyl)thymine-3'-triphosphate

50mg (95 μmol) of 1-{2'-*O*-[(4",4"'-Dimethoxytriphenyl)methyl]-α-ʟ-
threofuranosyl}thymine was co-evaporated twice with anhydrous pyridine and
stored under high vacuum over P_2O_5 for 1 h. The dry starting material
obtained was dissolved in 100 μl of anhydrous pyridine and 300 μl of
anhydrous dioxane and react with 95 μl of 2-chlor-4H-1,3,2-benzodioza-
phosphorin-4-one in anhydrous dioxane (1M stock). After 10 min a
well-vortexed mixture of 300 μl of tributylammonium pyrophosphate in
anhydrous DMF (0.5M stock, 1.5 mol eq.) and 100μl of tributylamine were
added and the reaction was allowed to stir for 15 min. 2.4 ml of I_2 solution in
pyridine/water 98:2 (1% stock, 1 mol eq.) was added to the reaction mixture.
The excess iodine was quenched after 10 min by adding 0.8 ml of 5% aq.
Na_2SO_3 solution followed by solvent removal. The residue was dissolved in
50% acetic acid for 20 min at 0℃. The solvent was removed and the residue
was then dissolved in water and washed twice with ethyl ether. The mixture
was lyophilized to dryness and purified by ion exchange chromatography and
reverse phase HPLC. The triethylammonium salt of the resulting triphosphate
was then converted to the sodium salt as a 5 mM solution for stock in 15%
yield. ^{31}P NMR (161.8 MHz, D_2O): δ=-10.3 (d, J=20.0), -12.4 (d, J=20.0),
-22.8 (t, J=19.7). MALDI-TOF MS (neg.) 466.77 (calc. 467.14).

(α-L-threofuranosyl)cytosine-3'-triphosphate

50mg (139 μmol) of N^4-Benzoyl-1-(2'-O-acetyl-α-L-threofuranosyl)-cytosine was co-evaporated twice with anhydrous pyridine and stored under high vacuum over P_2O_5 for 1 h. The dry starting material obtained was dissolved in 150 μl of anhydrous pyridine and 450 μl of anhydrous dioxane and react with 139 μl of 2-chlor-4H-1,3,2-benzodiozaphosphorin-4-one in anhydrous dioxane (1M stock). After 10 min a well-vortexed mixture of 420 μl of tributylammonium pyrophosphate in anhydrous DMF (0.5M stock) and 140μl of tributylamine were added and the reaction was allowed to stir for 15 min. 1.8 ml of I_2 solution in pyridine/water 98:2 (1% stock) was added to the reaction mixture. The excess iodine was quenched after 10 min by adding 0.6 ml of 5% aq. Na_2SO_3 solution followed by solvent removal. The residue was dissolved in 28% aq. NH_4OH solution at r. t. for 3 h. The solvent was removed and the residue was then dissolved in water and washed twice with CH_2Cl_2. The mixture was lyophilized to dryness and purified by ion exchange chromatography and reverse phase HPLC. The triethylammonium salt of the resulting triphosphate was then converted to the sodium salt as a 5 mM solution for stock in 12% yield. ^{31}P NMR (161.8 MHz, D_2O): δ=-10.4 (d, J=19.6), -12.6 (d, J=20.0), -22.9 (t, J=19.6). MALDI-TOF MS (neg.) 451.71 (calc. 452.13).

(α-L-threofuranosyl)guanine-3'-triphosphate

65mg (117 μmol) of 9-{2'-*O*-[(4",4"'-dimethoxytriphenyl)methyl]-α-L-threofuranosyl}guanine was co-evaporated twice with anhydrous pyridine and stored under high vacuum over P_2O_5 for 1 h. The dry compound obtained was dissolved in 195 μl of anhydrous pyridine and 780 μl of anhydrous DMF and react with 129 μl of 2-chlor-4H-1,3,2-benzodiozaphosphorin-4-one in anhydrous dioxane (1M stock). After 10 min a well-vortexed mixture of 351 μl of tributylammonium pyrophosphate in anhydrous DMF (0.5M stock) and 117 μl of tributylamine were added and the reaction was allowed to stir for 15 min. 3.0 ml of I_2 solution in pyridine/water 98:2 (1% stock) was added to the reaction mixture. The excess iodine was quenched after 10 min by adding 1.0 ml of 5% aq. Na_2SO_3 solution followed by solvent removal. The residue was dissolved in 50% acetic acid for 20 min at 0℃. The solvent was removed and the residue was then dissolved in water and washed twice with ethyl ether. The mixture was lyophilized to dryness and purified by ion exchange chromatography and reverse phase HPLC. The triethylammonium salt of the resulting triphosphate was then converted to the sodium salt as a 5 mM solution for stock in 8% yield. ^{31}P NMR (161.8 MHz, D_2O): δ=-8.57 (d, J=20.0), -12.2 (d, J=20.1), -22.6 (t, J=20.0). MALDI-TOF MS (neg.) 491.76 (calc. 491.97).

(α-ʟ-threofuranosyl)-2,6-diaminopurine-3'-triphosphate

53mg (95 μmol) of 9-{2'-*O*-[(4",4"'-dimethoxytriphenyl)methyl]-α-ʟ-threofuranosyl}-2,6-diaminopurine was co-evaporated twice with anhydrous pyridine and stored under high vacuum over P_2O_5 for 1 h. The dry starting material obtained was dissolved in 160 μl of anhydrous pyridine and 640 μl of anhydrous DMF and react with 104 μl of 2-chlor-4H-1,3,2-benzodiozaphosphorin-4-one in anhydrous dioxane (1M stock). After 10 min a well-vortexed mixture of 285 μl of tributylammonium pyrophosphate in anhydrous DMF (0.5M stock) and 95 μl of tributylamine were added and the reaction was allowed to stir for 15 min. 2.4 ml of I_2 solution in pyridine/water 98:2 (1% stock) was added to the reaction mixture. The excess iodine was quenched after 10 min by adding 0.8 ml of 5% aq. Na_2SO_3 solution followed by solvent removal. The residue was dissolved in 50% acetic acid for 20 min at 0℃. The solvent was removed and the residue was then dissolved in water and washed twice with ethyl ether. The mixture was lyophilized to dryness and purified by ion exchange chromatography and reverse phase HPLC. The triethylammonium salt of the resulting triphosphate was then converted to the sodium salt as a 5 mM solution for stock in 8% yield. ^{31}P NMR (161.8 MHz, D_2O): δ=-9.12 (d, J=20.0), -12.1 (d, J=20.0), -22.6 (t, J=18.4). MALDI-TOF MS (neg.) 490.76 (calc. 490.99).

Oligonucleotides and TNA triphosphates. DNA primers and templates were purchased from Integrated DNA Technologies. DNA library L1 was

169

purchased from the Keck Facilities at Yale University. DNA libraries L2 and L3 were synthesized on an automated ABI 3400 DNA synthesizer, deprotected in concentrated NH_4OH for 18 hours at 55 °C, butanol precipitated, and purified by denaturing polyacrylamide gel electrophoresis. TNA triphosphates were synthesized as previously described (Zou et al. 2005).

Primer extension assay. DNA primers were 5'-end labeled by incubating in the presence of [γ-^{32}P] ATP with T4 polynucleotide kinase for 1 hour at 37 °C. ^{32}P-labeled primer was annealed to the DNA template in 1x ThermoPol buffer [20 mM Tris-HCl, 10 mM $(NH_4)_2SO_4$, 10 mM KCl, 2 mM $MgSO_4$, 0.1% Triton X-100, pH 8.8] by heating at 95 °C for 5 minutes and cooling on ice. Primer extension reactions contained 100 mM tNTPs (or a combination of defined tNTPs and dNTPs mixtures), 500 nM primer-template complex, 1 mM DTT, 100 mg/ml BSA, 1.25 mM $MnCl_2$ and 0.1 U/ml Therminator DNA polymerase. Primer extension reactions were performed by adding the tNTP substrates to a solution containing all other reagents, and heating the mixture for 10 hours at 55 °C. Primer extension products were analyzed by 20% denaturing polyacrylamide gel electrophoresis.

In vitro selection. For each round of selection, the self-priming DNA library L2 was extended with tNTPs to generate a pool of chimeric TNA-DNA hairpin structures. The extension product was phenol/chloroform extracted, desalted, and subjected to a strand invasion step. A DNA primer (5'-/FAM/AAGGCTACTGCATACGAGTGTCACTACGTACCGTGGTACG

GCCGAGGTTG-3') annealed to the stem-loop region of the chimeric

TNA-DNA hairpin was extended with Therminator DNA polymerase using

2.5 mM dNTPs in 1x ThermoPol buffer for 30 minutes at 55 °C, followed by

an additional heating step of 90 minutes at 65 °C. The resulting pool of

double-stranded DNA-TNA fusion molecules was incubated with human

a-thrombin (Haematologic Technologies Inc.) at 25 °C for 1 hour. The ratio of

human thrombin to the TNA library decreased over the course of the selection

to favor the enrichment of high affinity aptamers. In round 1, the ratio was

10:1; in round 2 the ratio was 1:1; and in round 3, the ratio was 1:10. Bound

molecules were separated from the unbound pool using capillary

electrophoresis. Capillary electrophoresis was performed on Beckman

ProteomeLab PA 800 Protein Characterization System. Prior to use, the glass

capillary (inner diameter = 0.1 mm, total length = 60 cm) was rinsed with

water and equilibrated with selection buffer [50 mM Tris-HCl, 100 mM NaCl,

1 mM $MgCl_2$, pH 7.4]. A small portion (70 nL) of the library/thrombin mixture

was injected onto the capillary using pressure injection (0.5 psi for 5 seconds)

and electrophoresis was performed under a constant voltage of 10 kV at 20 °C

for 30 minutes. Laser-induced fluorescence (LIF) was used to monitor the

separation of 6-FAM labeled DNA-TNA fusion molecules (excitation = 488

nm; emission = 520 nm). Five injections were performed for each round of in

vitro selection. The thrombin bound TNA was collected and it encoding DNA

template was amplified by polymerase chain reaction (PCR), which was then

used to generate a new self-priming library as input for the next round of selection. After three rounds, the library was cloned and sequenced to examine the diversity of molecules that remained in the pool.

Filter binding assay. Representative TNA molecules were synthesized by primer extension and purified by denaturing polyacrylamide gel electrophoresis as described above with the exception that extension time was reduced to 1 hour. The DNA portion of the TNA molecule was labeled with ^{32}P, desalted, and annealed by heating for 5 minutes at 95 °C and cooling on ice. Folded structures were incubated with human thrombin poised at concentrations spanning the expected Kd (typically 1 nM - 1 mM) at 25 °C. After 1 hour, the protein bound TNA molecules were partitioned away from the unbound fractions using vacuum to pass the solution through a layer of nitrocellulose and nylon membranes. Both membranes were quantified by phosphorimaging, and dissociation constants were calculated using nonlinear least-squares regression analysis performed with DeltaGraph program (Lin et al. 2007; Rogers and Joyce 1999).

6.7 References

Benner, S.A., A.D. Ellington, and A. Tauer. 1989. Modern metabolism as a palimpsest of the RNA world. *Proc. Natl. Acad. Sci. USA* 86: 7054-7058.

Benner, S.A., and A.M. Sismour. 2005. Synthetic biology. *Nat. Rev. Genet.* 6: 533-543.

Bock, L.C., L.C. Griffin, J.A. Latham, E.H. Vermaas, and J.J. Toole. 1992. Selection of single-stranded DNA molecules that bind and inhibit human thrombin. *Nature* 355, 564-566.

Cech, T.R. 2011. The RNA Worlds in Context. *Cold Spring Harb. Perspect. Biol.*, doi: 10.1101/cshperspect.a006742.

Chaput, J.C., J.K. Ichida, and J.W. Szostak. 2003a. DNA polymerase-mediated DNA synthesis on a TNA template. *J. Am. Chem. Soc.* 125: 856-857.

Chaput, J.C., and J.W. Szostak. 2003b. TNA synthesis by DNA polymerases. *J. Am. Chem. Soc.* 125: 9274-9275.

Crick, F.H.C. 1968. The origin of the genetic code. *J. Mol. Biol.* 38: 367-369.

Ebert, M.-O., C. Mang, R. Krishnamurthy, A. Eschenmoser, and B. Jaun. 2008. The structure of a TNA-TNA complex in solution: NMR study of the octamer duplex derived from α-(L)-threofuranosyl-(3'-2')-CGAATTCG. *J. Am. Chem. Soc.* 130: 15105-15115.

Eschenmoser, A. 1999. Chemical etiology of nucleic acid structure. *Science* 284: 2118-2124.

Gilbert, W. 1986. The RNA world. *Nature* 319: 618.

Horhota, A., K. Zou, J.K. Ichida, B. Yu, L.W. McLaughlin, J.W. Szostak, and J.C. Chaput. 2005. Kinetic analysis of an efficient DNA-dependent TNA polymerase. *J. Am. Chem. Soc.* 127: 7427-7434.

Ichida, J.K., A. Horhota, K. Zou, L.W. McLaughlin, and J.W. Szostak. 2005a. High fidelity TNA synthesis by Therminator polymerase. *Nucleic Acids Res.* 33: 5219-5225.

Ichida, J.K., K. Zou, A. Horhota, B. Yu, L.W. McLaughlin, and J.W. Szostak. 2005b. An in vitro selection system for TNA. *J. Am. Chem. Soc.* 137: 2802-2803.

Joyce, G.F. 1991. The rise and fall of the RNA world. *The New Biologist* 3: 399-407.

Joyce, G.F. 2002. The antiquity of RNA-based evolution. *Nature* 418: 214-221.

Joyce, G.F. 2007. Forty years of in vitro evolution. *Angew. Chem. Int. Ed.* 46: 6420-6436.

Joyce, G.F., A.W. Schwartz, S.L. Miller, and L.E. Orgel. 1987. The case for an ancestral genetic system involving simple analogues of the nucleotides. *Proc. Natl. Acad. Sci. USA* 84: 4398-4402.

Kempeneers, V., K. Vastmans, J. Rozenski, and P. Herdewijn. 2003. Recognition of threosyl nucleotides by DNA and RNA polymerases. *Nucleic Acids Res.* 31: 6221-6226.

Kubik, M.F., A.W. Stephens, D. Schneider, R.A. Marlar, and D. Tasset. 1994. High-affinity RNA ligands to human a-thrombin. *Nucleic Acids Res.* 22: 2619-2626.

Levy, M., and S.L. Miller. 1998. The stability of RNA bases: implications for the origin of life. *Proc. Natl. Acad. Sci. USA* 95: 7933-7938.

Lin, L., D. Hom, S.M. Lindsay, and J.C. Chaput. 2007. In vitro selection of histone H4 aptamers for recognition imaging microscopy. *J. Am. Chem. Soc.* 129: 14568-14569.

Loakes, D., and P. Holliger. 2009. Polymerase engineering: towards the encoded synthesis of unnatural polymers. *Chem. Commun.* 4619-4631.

Mendonsa, S.D., and M.T. 2004. Bowser. In vitro evolution of functional DNA using capillary electrophoresis. *J. Am. Chem. Soc.* 126: 20-21.

Orgel, L.E. 1968. Evolution of the genetic apparatus. *J. Mol. Biol.* 38: 381-393.

Orgel, L.E. 2000. A simpler nucleic acid. *Science* 290: 1306-1307.

Roberts, R.W., and J.W. Szostak. 1997. RNA-peptide fusions for the in vitro selection of peptides and proteins. *Proc. Natl. Acad. Sci. USA* 94: 12297-12302.

Rogers, J., and G.F. Joyce. 1999. A ribozyme that lacks cytidine. *Nature* 402: 323-325.

Rosenbaum, D.M., and D.R. Liu. 2003. Efficient and sequence-specific DNA-templated polymerization of peptide nucleic acid aldehydes. *J. Am. Chem. Soc.* 125: 13924-13925.

Schöning, K.-U., P. Scholz, S. Guntha, X. Wu, R. Krishnamurthy, and A. Eschenmoser. 2000. Chemical etiology of nucleic acid structure: The α-L-threofuranosyl-(3′→2′) oligonucleotide system. *Science* 290: 1347–1351.

Williams, B.A.R., L. Lin, S.M. Lindsay, and J.C. Chaput. 2009. Evolution of a histone H4-K16 acetyl-specific DNA aptamer. *J. Am. Chem. Soc.* 131: 6330-6331.

Wilson, D.S., and J.W. Szostak. 1999. In vitro selection of functional nucleic acids. *Annu. Rev. Biochem.* 68: 611-647.

Woese, C.R. 1967. *The genetic code: The molecular basis for genetic expression.* New York: Harper & Row.

Wu, X., G. Delgado, R. Krishnamurthy, and A. Eschenmoser. 2002. 2,6-Diaminopurine in TNA: effects on duplex stabilities and on the efficiency of template-controlled ligations. *Org. Lett.* 4: 1283-1286.

Yarus, M. 2011. Getting past the RNA world: the initial Darwinian ancestor. *Cold Spring Harb. Perspect. Biol.*, doi: 10.1101/cshperspect.a003590.

Zou, K., A. Horhota, B. Yu, J.W. Szostak, and L.W. McLaughlin. 2005. Synthesis of α-L-Threofuranosyl Nucleoside Triphosphates (tNTPs). *Org. Lett.* 7: 1485-1487.

Chapter 7

CONCLUDING REMARKS

The work presented in the preceding chapters explores the application of unnatural nucleic acids in DNA nanotechnology and molecular evolution. As molecules structurally different from natural DNA/RNA, these unnatural nucleic acids have to be chemically synthesized. Chapter 2 and 3 described the synthetic procedures of phosphoramidites and oligonucleotides for GNA and TNA, respectively. In Chapter 5, a simple nanostructure exclusively composed of GNA was demonstrated. This structure had extraordinary thermalstability comparing with the same structure made up of DNA. The results indicated that GNA could be potentially useful in nanoarchitecture construction as an orthogonal building material. In Chapter 6, TNA triphosphates were synthesized and subjected to polymerase-mediated primer extension. The results had lead to an effective way of TNA random library construction. It has been of great interest to investigate the sequence space of TNA as a potential RNA progenitor candidate. The results in Chapter 6 made it possible to evolve functional molecules using a large pool of random TNA sequences. The very first TNA aptamer evolved from this pool indicated that ligand binding is not limited to natural biopolymers. In Chapter 4, the cross-pairing between complementary GNA and TNA were examined in order to find out whether there could be a consecutive relationship between these two polymer systems in an evolutionary perspective since both polymer systems are structurally

simpler and more accessible to prebiotic conditions than RNA. However, the observation that there is no hybridization between the two polymer systems suggested that GNA and TNA might not be sequential polymers in prebiotic evolution of RNA.

Future Directions

In Chapter 5, the successful construction of 4-helix junction using GNA as building block has shown promising application of GNA in structural DNA nanotechnology. It is intriguing and necessary to bring this work further to build more complex nanostructures. In order to achieve this goal, it is imperative to understand the structure of GNA in more detail, which will provide useful information such as helical pitch and flexibility of the GNA strands to direct the design of GNA nanostructures with higher complexity. The conformation study of GNA is another interesting direction. This study might shed light on general understanding of the duplex formation and stability.

In Chapter 6, TNA aptamers binding thrombin has been successfully isolated. Evolution of TNA catalysts will further enrich the functionality of TNA molecules. The next step of the research will be to understand the structural basis of these evolved TNA molecules. The binding site structures and catalytic mechanisms need to be elucidated and compared with that of RNA molecules. This will help address the question about the fitness

177

landscape of TNA comparing to RNA. This line of research may eventually

lead to the evolution of artificial genetic systems.

In summary, unnatural nucleic acids enrich the diversity of genetic

alphabet, provide us with useful tools for both basic research and applications

as well as expand the understanding to nucleic acid and evolution of life. A

major goal of synthetic biology is to recapitulate emergent properties of life, in

which genetic material plays an indispensable role. Unnatural nucleic acids

might have found their niche as novel materials, metabolite regulators, and

diagnostics and therapeutics. The evolution of unnatural nucleic acid

molecules may ultimatelt lead to the creation of an artifical genetic system

capable of Darwinian evolution.

References

Yang, Y.-W., S. Zhang, E.O. McCullum, and J.C. Chaput. 2007. Experimental evidence that GNA and TNA were not sequential polymers in the prebiotic evolution of RNA. *J. Mol. Evol.* 65: 289-295.

Yu, H., S. Zhang, and J.C. Chaput. Unpublished Data.

Zhang, S., and J.C. Chaput. 2010. Synthesis of Glycerol Nucleic Acid (GNA) Phosphoramidite Monomers and Oligonucleotide Polymers. *Curr. Protoc. Nucleic Acid Chem.* Unit 4.40.

Zhang, R.S., E.O. McCullum, and J.C. Chaput. 2008. Synthesis of two mirror image 4-helix junctions derived from glycerol nucleic acid. *J. Am. Chem. Soc.* 130: 5846-5847.

Zhang, S., C. Switzer, and J.C. Chaput. 2010. The resurgence of acyclic nucleic acids. *Chem. Biodiver.* 7: 245-258.

BIBLIOGRAPHY

Acevedo, O.L., and R.S. Andrews. 1996. Synthesis of propane-2,3-diol combinatorial monomers. *Tetrahedron Lett.* 37: 3913-3934.

Appella, D.H. 2009. Non-natural nucleic acids for synthetic biology. *Curr. Opin. Chem. Biol.* 13: 687-696.

Ban, N., P. Nissen, J. Hansen, P.B. Moore, and T.A. Steitz. 2000. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* 289:905–920.

Bartlett, J.M.S., and D. Stirling. 2003. A short history of the polymerase chain reaction. *Methods Mol. Biol.* 226: 3-6.

Benner, S.A., A.D. Ellington, and A. Tauer. 1989. Modern metabolism as a palimpsest of the RNA world. *Proc. Natl. Acad. Sci. USA* 86: 7054-7058.

Benner, S.A., and A.M. Sismour. 2005. Synthetic biology. *Nat. Rev. Genet.* 6: 533-543.

Bock, L.C., L.C. Griffin, J.A. Latham, E.H. Vermaas, and J.J. Toole. 1992. Selection of single-stranded DNA molecules that bind and inhibit human thrombin. *Nature* 355, 564-566.

Böhler, C., P.E. Nielsen, and L.E. Orgel. 1995. Templated switching between PNA and RNA oligonucleotides. *Nature* 376: 578-581.

Brudno, Y., M. Birnbaum, R.E. Kleiner, and D.R. Liu. 2010. An in vitro translation, selection and amplification system for peptide nucleic acids. *Nat. Chem. Biol.* 6: 148-155.

Brudno, Y., and D.R. Liu. 2009. Recent progress towards the templated synthesis and directed evolution of sequence-defined synthetic polymers. *Chem. Biol.* 16: 265-276.

Cech, T.R. 2011. The RNA Worlds in Context. *Cold Spring Harb. Perspect. Biol.*, doi: 10.1101/cshperspect.a006742.

Chaput, J.C., J.K. Ichida, and J.W. Szostak. 2003a. DNA polymerase-mediated DNA synthesis on a TNA template. *J. Am. Chem. Soc.* 125: 856-857.

Chaput, J.C., and C. Switzer. 1999. A DNA pentalex incorporating nucleobase quintets. *Proc. Natl. Acad. Sci. USA* 96: 10614-10619.

179

Chaput, J.C., and C. Switzer. 2000. Nonenzymatic oligomerization on templates containing phosphodiester-linked acyclic glycerol nucleic acid analogues. *J. Mol. Evol.* 51: 464-470.

Chaput, J.C., and J.W. Szostak. 2003b. TNA synthesis by DNA polymerases. *J. Am. Chem. Soc.* 125: 9274-9275.

Ciengshih, T., R. Sha, and N.C. Seeman. 2011. Automatic molecular weaving prototyped by using single-stranded DNA. *Angew. Chem. Int. Ed.* 50: 4419-4422.

Crick, F.H.C. 1968. The origin of the genetic code. *J. Mol. Biol.* 38: 367-369.

Cooper, G., N. Kimmich, W. Belisle, J. Sarinana, K. Brabham, and L. Garrel. 2001. Carbonaceous meteorites as a source of sugar-related organic compounds for the early earth. *Nature* 414: 879–883.

Croft, M. 2009. The role of TNF superfamily members in T-cell function and diseases. *Nat. Rev. Immunol.* 9: 271-285.

Declercq, R., A. Van Aerschot, R.J. Read, P. Herdewijn, and L. Van Meervelt. 2002. Crystal structure of double helical hexitol nucleic acids. *J. Am. Chem. Soc.* 124: 928-933.

Ebert, M.-O., C. Mang, R. Krishnamurthy, A. Eschenmoser, and B. Jaun. 2008. The structure of a TNA-TNA complex in solution: NMR study of the octamer duplex derived from α-(L)-threofuranosyl-(3'-2')-CGAATTCG. *J. Am. Chem. Soc.* 130: 15105-15115.

Egholm, M., O. Buchardt, L. Christensen, C. Behrens, S.M. Freier, D.A. Driver, R.H. Berg, S.K. Kim, P.E. Norden, and P.E. Nielsen. 1993. PNA hybridizes to complementary oligonucleotides obeying the Watson-Crick hydrogen-bonding rules. *Nature* 365: 566-568.

Eguchi, Y., T. Itoh, and J. Tomizawa. 1991. Antisense RNA. *Annu. Rev. Biochem.* 60: 631-652.

Eschenmose, A. 1999. Chemical etiology of nucleic acid structure. *Science* 284: 2118-2124.

Famulok, M. 2009. Exploring chemical space with aptamers. *J. Med. Chem.* 52: 6951-6957.

Feldkamp, U., and C.M. Niemeyer. 2006. Rational design of DNA nanoarchitectures. *Angew. Chem. Int. Ed.* 45: 1856-1876.

Gardner, A.F., and W.E. Jack. 2002. Acyclic and dideoxy terminator preferences denote divergent sugar recognition by archaeon and *Taq* DNA polymerases. *Nucleic Acid Res.* 30: 605-613.

Gartner, Z.J., B.N. Tse, R. Grubina, J.B. Doyon, T.M. Snyder, and D.R. Liu. 2004. DNA-templated organic synthesis and selection of a library of macrocycles. *Science* 305: 1601-1605.

Gilbert, W. 1986. The RNA world. *Nature* 319: 618.

Goodwin, J.T., and D.G. Lynn. 1992. Template-directed synthesis: use of a reversible reaction. *J. Am. Chem. Soc.* 114: 9197-9198.

Hansen, M.H., P. Blakskjær, L.K. Peterson, T.H. Hansen, J.W. Højfeldt, K.V. Gothelf, and N.J.V. Hansen. 2009. A yoctoliter-scale DNA reactor for small-molecule evolution. *J. Am. Chem. Soc.* 131: 1322-1327.

Hendrix, C., H. Rosemeyer, I. Verheggen, F. Seela, A. Van Aerschot, and P. Herdewijn. 1997. 1',5'-Anhydrohexitol oligonucleotides: synthesis, base pairing and recognition by regular oligodeoxyribonucleotides and oligoribonucleotides. Chem. Eur. J. 3: 110-120.

Heuberger, B.D., and C. Switzer. 2006. Nonenzymatic oligomerization of RNA by TNA templates. *Org. Lett.* 8: 5809-5811.

Heuberger, B.D., and C. Switzer. 2008. A pre-RNA candidate revisited: both enantiomers of flexible nucleoside triphosphates are DNA polymerase substrates. *J. Am. Chem. Soc.* 130: 412-413.

Holliday, R. 1964. A mechanism for gene conversion in fungi. Genet. Res. 5: 282-304.

Holý, A. 1975. Aliphatic analogues of nucleosides, nucleotides, and oligonucledotides. *Collect. Czech. Chem. Commun.* 40: 187-214.

Hornota, A.T., J.W. Szostak, and L.W. McLaughlin. 2006. Glycerol nucleoside triphosphates: synthesis and polymerase substrate activities. *Org. Lett.* 8: 5345-5347.

Horhota, A., K. Zou, J.K. Ichida, B. Yu, L.W. McLaughlin, J.W. Szostak, and J.C. Chaput. 2005. Kinetics analysis of an efficient DNA-dependent TNA polymerase. *J. Am. Chem. Soc.* 127: 7427-7434.

Ichida, J.K., A. Horhota, K. Zou, L.W. McLaughlin, and J.W. Szostak. 2005a. High fidelity TNA synthesis by Therminator polymerase. *Nucleic Acids Res.* 33: 5219-5225.

Ichida, J.K., K. Zou, A. Horhota, B. Yu, L.W. McLaughlin, and J.W. Szostak. 2005b. An in vitro selection system for TNA. *J. Am. Chem. Soc.* 127: 2802-2803.

Johnson, A.T., M.K. Schlegel, E. Meggers, L.-O. Essen, and O. Wiest. 2011. On the structure and dynamics of the duplex GNA. *J. Org. Chem.* 76: 7964-7974.

Jovin, T.M., and D.M. Soumpasis. 1983. The transition between B-DNA and Z-DNA. *Annu. Rev. Phys. Chem.* 38: 521-560.

Joyce, G.F. 1989. RNA evolution and the origins of life. *Nature* 338: 217–224.

Joyce, G.F. 1991. The rise and fall of the RNA world. *The New Biologist* 3: 399-407.

Joyce, G.F. 2002. The antiquity of RNA-based evolution. *Nature* 418: 214-221.

Joyce, G.F. 2007. Forty years of in vitro evolution. *Angew. Chem. Int. Ed.* 46: 6420-6436.

Joyce, G.F., A.W. Schwartz, S.L. Miller, and L.E. Orgel. 1987. The case for an ancestral genetic system involving simple analogues of the nucleotides. *Proc. Natl. Acad. Sci. USA* 84: 4398-4402.

Joyce, G.F., G.M. Visser, C.A.A. van Boeckel, J.H. van Boom, L.E. Orgel, and J. van Westrenen. 1984. Chiral selection in poly(C)-directed synthesis of oligo(G). *Nature* 310: 602-604.

Kallenbach, N.R., R.-I. Ma, and N.C. Seeman. 1983. An immobile nucleic acid junction constructed from oligonucleotides. *Nature* 305: 829-831.

Kang, H., M.H. Fisher, D. Xu, Y.J. Miyamoto, A. Marchand, A. Van Aerschot, P. Herdewijn, and R.L. Juliano. 2004. Inhibition of MDR1 gene expression by chimeric HNA antisense oligonucleotides. *Nucleic Acid Res.* 32: 4411-4419.

Keefe, A.D., and S.T. Cload. 2008. SELEX with modified nucleotides. *Curr. Opin. Chem. Biol.* 12: 448-456.

Kempeneers, V., K. Vastmans, J. Rozenski, and P. Herdewijn. 2003. Recognition of threosyl nucleotides by DNA and RNA polymerases. *Nucleic Acids Res.* 31: 6221-6226.

Kerremans, L., G. Schepers, J. Rozenski, R. Busson, A. Van Aerschot, and P. Herdewijn. 2001. Hybridization between "six-membered" nucleic acids: RNA as a universal information system. *Org. Lett.* 3: 4129-4132.

Khudyakov, Y.E., and H.A. Fields. eds. 2002. *Artificial DNA: methods and applications*. Boca Raton, Florida: CRC Press.

Knudsen, S.M., M.P. Robertson, and A.D. Ellington. 2002. In vitro selection using modified or unnatural nucleotides. *Curr. Protoc. Nucleic. Acid Chem.* Chapter 9: unit 9.6.

Kolb, G., S. Reigadas, C. Boiziau, A. Van Aerschot, A. Arzumanov, M.J. Gait, P. Herdewijn, and J.-J. Toulme. 2005. Hexitol nucleic acid-containing aptamers are efficient ligands of HIV-1 TAR RNA. *Biochemistry* 44: 2926-2933.

Kruger, K., P.J. Grabowski, A.J. Zaug, and T.R. Cech. 1982. Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence by Tetrahymena. *Cell* 31: 147–157.

Kubik, M.F., A.W. Stephens, D. Schneider, R.A. Marlar, and D. Tasset. 1994. High-affinity RNA ligands to human a-thrombin. *Nucleic Acids Res.* 22: 2619-2626.

Kuwahara, M., S. Obika, J.-I. Nagashima, Y. Ohta, Y. Suto, H. Ozaki, H. Sawai, and T. Imanishi. 2008. Systematic analysis of enzymatic DNA polymerization using oligo-DNA templates and triphosphate analogs involving 2',4'-bridged nucleosides. *Nucleic Acid Res.* 36: 4257-4265.

Lebars, I., T. Richard, C. Di Primo, and J.J. Toulme. 2007. LNA derivatives of a kissing aptamer targeted to the *trans*-activating responsive RNA element of HIV-1. *Blood Cells Mol. Dis.* 38: 204-209.

Levy, M., and S.L. Miller. 1998. The stability of RNA bases: implications for the origin of life. *Proc. Natl. Acad. Sci. USA* 95: 7933-7938.

Li, X., and D.R. Liu. 2004. DNA-templated organic synthesis: nature's strategy for controlling chemical reactivity applied to synthetic molecules. *Angew. Chem. Int. Ed.* 43: 4848-4870.

Lin, L., D. Hom, S.M. Lindsay, and J.C. Chaput. 2007. In vitro selection of histone H4 aptamers for recognition imaging microscopy. *J. Am. Chem. Soc.* 129: 14568-14569.

Lin, C., Y. Ke, Z. Li, J.H. Wang, Y. Liu, and H. Yan. 2009 Mirror image DNA nanostructures for chiral supramolecular assemblies. *Nano Lett.* 9: 433-436.

Lin, C., Y. Liu, S. Rinker, and H. Yan. 2006. DNA tile based self-assembly: building complex nanoarchitectures. *Chemphyschem* 7: 1641-1647.

Lin, C., Y. Liu, and H. Yan. 2009. Designer DNA nanoarchitectures. *Biochemistry* 48: 1663-1674.

Liu, H., J. Gao., S.R. Lynch, Y.D. Saito, L. Maynard, and E.T. Kool. 2003. A four-base paired genetic helix with expanded size. *Science* 302: 868-871.

Loakes, D., and P. Holliger. 2009. Polymerase engineer: towards the encoded synthesis of unnatural biopolymers. *Chem. Commun.* 4619-4631.

Lukeman, P.S., A.C. Mittal, and N.C. Seeman. 2004. Two dimensional PNA/DNA arrays: estimating the helicity of unusual nucleic acid polymers. *Chem. Commun.* 1694-1695.

Maier, T., I. Przylas, N. Strater, P. Herdewijn, and W. Saenger. 2005. Reinforced HNA backbone hydration in the crystal structure of a decameric HNA/RNA hybrid. *J. Am. Chem. Soc.* 127: 2937-2934.

Marky, L.A., and K.J. Breslauer. 1987. Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves. *Biopolymers* 26: 1601–1620.

Megger, E., and L. Zhang. 2010. Synthesis and properties of the simplified nucleic acid GNA. *Acc. Chem. Res.* 43: 1092-1102.

Mendonsa, S.D., and M.T. 2004. Bowser. In vitro evolution of functional DNA using capillary electrophoresis. *J. Am. Chem. Soc.* 126: 20-21.

Milligan, J.F., and O.C. Uhlenbeck. 1989. Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol.* 180: 51-62.

Muller, S., J. Wolf, and S.A. Ivanov. 2004. Current strategies for the synthesis of RNA. *Curr. Org. Syn.* 1: 293-307.

Nahvi, A., N. Sudarsan, M.S. Ebert, X. Zou, K.L. Brown, and R.R. Breaker. 2002. Genetic control by a metabolite binding mRNA. *Chem. Biol.* 9: 1043-1049.

Nangreave, J., Y. Liu, and H. Yan. 2011. DNA nanostructures as models for evaluating the role of enthalpy and entropy in polyvalent binding. *J. Am. Chem. Soc.* 133: 4490-4497.

Ng, P.-S., and D.E. Bergstrom. 2005. Alternative nucleic acid analogues for programmable assembly: hybridization of LNA to PNA. *Nano Lett.* 5: 107-111.

Nielsen, P.E. 1995. DNA analogues with nonphosphodiester backbones. *Annu. Rev. Biophys. Biomol. Struct.* 24: 167-183.

Nielsen, K.E., J. Rasmussen, R. Kumar, J. Wengel, J.P. Jacobsen, and M. Peterson. 2004. NMR studies of fully modified locked nucleic acid (LNA) hybrids: solution structure of an LNA: RNA hybrid and characterization of an LNA: DNA hybrid. *Bioconjug. Chem.* 15: 449-457.

Niemeyer, C.M. 2000. Self-assembled nanostructures based on DNA: towards the development of nanobiotechnology. *Curr. Opin. Chem. Biol.* 4: 609-618.

Obika, S., M. Dorio, D. Naubu, and T. Imanishi. 1997. Synthesis and conformation of 3'-O,4'-C-methyleneribonucleotides, novelbicyclic nucleoside analogues for 2',5'-linked oligonucleotide modification. *Chem. Commun.* 1643-1644.

Orgel, L.E. 1968. Evolution of the genetic apparatus. *J. Mol. Biol.* 38: 381-393.

Orgel, L. 2000. A simpler nucleic acid. *Science* 17: 1306-1307.

Pallan, P.S., C.J. Wilds, Z. Wawrzak, R. Krishnamurthy, A. Eschenmoser, and M. Egli. 2003. Why does TNA cross-pair more strongly with RNA than with DNA? An answer from X-ray analysis. *Angew. Chem. Int. Ed.* 42: 5893-5895.

Rich, A., A. Nordheim, and A.H. Wang. 1984. The chemistry and biology of left-handed Z-DNA. *Annu. Rev. Biochem.* 53: 791-846.

Rinker, S., Y. Liu, and H. Yan. 2006. Two dimensional LNA/DNA arrays: estimating the helicity of LNA/DNA hybrid duplex. *Chem. Commun.* 2675-2677.

Roberts, R.W., and J.W. Szostak. 1997. RNA-peptide fusions for the in vitro selection of peptides and proteins. *Proc. Natl. Acad. Sci. USA* 94: 12297-12302.

Rogers, J., and G.F. Joyce. 1999. A ribozyme that lacks cytidine. *Nature* 402: 323-325.

Rosenbaum, D.M., and D.R. Liu. 2003. Efficient amd sequence-specific DNA-templated polymerization of peptide nucleic acid aldehydes. *J. Am. Chem. Soc.* 125: 13924-13925.

Sacca, B., R. Meyer, U. Feldkamp, H. Schroeder, and C.M. Niemeyer. 2008. High-throughput, real-time monitoring of the self-assembly of DNA nanostructures by FRET spectroscopy. *Angew. Chem. Int. Ed.* 47: 2135-2137.

Schlegel, M.K., L.-O. Essen, and E. Meggers. 2008. Duplex formation of a minimal nucleic acid. *J. Am. Chem. Soc.* 130: 8158-8159.

Schlegel, M.K., L.-O. Essen, and E. Meggers. 2010. Atomic resolution duplex structure of the simplified nucleic acid GNA. *Chem. Commun.* 46: 1094-1096.

Schlegel, M.K., A.E. Peritz, K. Kittigowittana, L. Zhang, and E. Meggers. 2007. Duplex formation of the simplified nucleic acid GNA. *ChemBioChem* 8: 927-932.

Schlegel, M.K., X. Xie, L. Zhang, and E. Meggers. 2009a. Insight into the high duplex stability of the simplified nucleic acid GNA. *Angew. Chem. Int. Ed.* 48: 960-963.

Schlegel, M.K., and E. Meggers. 2009b. Improved phosphoramidite building blocks for the synthesis of the simplified nucleic acid GNA. *J. Org. Chem.* 74: 4615-4618.

Schlonvogt, I., S. Pitsch, C. Lesuer, A. Eschenmoser, B. Jaun, and R.M. Wolf. 1996. Pyranosyl-RNA ('p-RNA'): NMR and molecular dynamics study of the duplex formed by self-pairing of ribopyranosyl-(C-G-A-A-T-T-C-G). *Helv. Chim. Acta.* 79: 2316–2345.

Schmidt, J.G., P.E. Nielsen, and L.E. Orgel. 1997a. Information transfer from peptide nucleic acids to RNA by template-directed syntheses. *Nucleic Acid Res.* 25: 4797-4802.

Schmidt, J.G., L. Christensen, P.E. Nielsen, and L.E. Orgel. 1997b. Information transfer from DNA to peptide nucleic acids by template-directed syntheses. *Nucleic Acid Res.* 25: 4792-4796.

Schneider, K.C., and S.A. Benner. 1990. Oligonucleotides containing flexible nucleoside analogs. *J. Am. Chem. Soc.* 112: 453-455.

Schöning, K.-U., P. Scholz, S. Guntha, X. Wu, R. Krishnamurthy, and A. Eschenmoser. 2000. Chemical etiology of nucleic acid structure: the α-threofuranosyl-(3'→2') oligonucleotide system. *Science* 290: 1347-1351.

Schöning, K.-U., P. Scholz, X. Wu, S. Guntha, G. Delgado, R. Krishnamurthy, and A. Eschenmoser. 2002. The α-L-threofuranosyl-(3'→2')-oligonucleotide system ('TNA'): synthesis and pairing properties. *Helv. Chim. Acta* 85: 4111-4153.

Schwartz, A.W., and L.E. Orgel. 1985. Template-directed synthesis of novel, nucleic acid like structures. *Science*. 228: 585-587

Seeman, N.C. 1982. Nucleic acid junctions and lattices. *J. Theor. Biol.* 99: 237-247.

Seeman, N.C. 1998. DNA nanotechnology: novel DNA constructions. *Annu. Rev. Biophys. Biomol. Struct.* 27: 225-248.

Seeman, N.C. 2003. At the crossroad of chemistry, biology, and materials: structural DNA nanotechnology. *Chem. Biol.* 10: 1151-1159.

Seeman, N.C. 2005. From genes to machines: DNA nanomechanical devices. *Trends Biochem. Sci.* 30: 119-125.

Singh, S.K., P. Nielson, A.A. Koshkin, and J. Wengel. 1998. LNA (locked nucleic acid): synthesis and high-affinity nucleic acid recognition *Chem. Commun.* 455-456.

Tsai, C.H., J. Chen, and J.W. Szostak. 2007. Enzymatic synthesis of DNA on glycerol nucleic acid templates without stable duplex formation between product and template. *Proc. Natl. Acad. Sci. USA* 104: 14598-14603.

Urata, H., E. Ogura, K. Shinohara, Y. Ueda, and M. Akagi. 1992. Synthesis and properties of mirror-image DNA. *Nucleic Acid Res.* 20: 3325-3332.

Urata, H., K. Shinohara, E. Ogura, Y. Ueda, and M. Akagi. 1991. Mirror-image DNA. *J. Am. Chem. Soc.* 113: 8174-8175.

Vastmans, K., S. Pochet, A. Peys, L. Kerremans, A. Van Aerschot, C. Hendrix, P. Marliere, and P. Herdewijn. 2000. Enzymatic incorporation in DNA of 1,5-anhydrohexitol nucleotides. *Biochemistry* 39: 12757-12765.

Vaught, J.D., C. Bock, J. Carter, T. Fitzwater, M. Otis, D. Schneider, J. Rolando, S. Waugh, S.K. Wilcox, and B.E. Eaton. 2010. Expanding the chemistry of DNA for in vitro selection. *J. Am. Chem. Soc.* 132: 4141-4151.

Vester, B., and J. Wengel. 2004. LNA (locked nucleic acid): high-affinity targeting of complementary RNA and DNA. *Biochemistry* 43: 13233-13241.

Vorbrggen, H., and B. Bennua. 1981. Nucleoside syntheses, XXV: a new simplified nucleoside synthesis. *Chem. Ber.* 114: 1279-1286.

Wagenknecht, H.-A. 2008. Fluorescent DNA base modifications and substitutes: multiple fluorophore labeling and the DETEQ concept. *Ann. N.Y. Acad. Sci.* 1130: 122-130.

Wang, L., J. Xie, and P.G. Schultz. 2006. Expanding the genetic code. *Annu. Rev. Biophys. Biomol. Struct.* 35: 225-249.

Watson, J.D., and F.H.C. Crick. 1953. Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature* 171: 737-738.

Wilds, C.J., Z. Warwrzak, R. Krishnamurthy, A. Eschenmoser, and M. Egli. 2002. Crystal structure of a B-form DNA duplex containing (L)-α-threofuranosyl (3' →2') nucleosides: a four-carbon sugar is easily accommodated into the backbone of DNA. *J. Am. Chem. Soc.* 124: 13716-13721.

Williams, B.A.R., L. Lin, S.M. Lindsay, and J.C. Chaput. 2009. Evolution of a histone H4-K16 acetyl-specific DNA aptamer. *J. Am. Chem. Soc.* 131: 6330-6331.

Wilson, D.S., and J.W. Szostak. 1999. In vitro selection of functional nucleic acids. *Annu. Rev. Biochem.* 68: 611-647.

Wimberly, B.T., D.E. Brodersen, W.M. Clemons, R.J. Morgan-Warren, A.P. Carter, C. Vonrhein, T. Hartsch, and V. Ramakrishnan. 2000. Structure of the 30S ribosomal subunit. *Nature* 407: 327–339.

Woese, C.R. 1967. *The genetic code: The molecular basis for genetic expression.* p.179. New York: Harper & Row.

Wu, X., G. Delgado, R. Krishnamurthy, and A. Eschenmoser. 2002. 2,6-Diaminopurine in TNA: effects on duplex stabilities and on the efficiency of template-controlled ligations. *Org. Lett.* 4: 1283-1286.

Yan, H. 2004. Nucleic acid nanotechnology. *Science* 17: 2048-2049.

Yang, Y.-W., S. Zhang, E.O. McCullum, and J.C. Chaput. 2007. Experimental evidence that GNA and TNA were not sequential polymers in the prebiotic evolution of RNA. *J. Mol. Evol.* 65: 289-295.

Yarus, M. 2011. Getting past the RNA world: the initial Darwinian ancestor. *Cold Spring Harb. Perspect. Biol.*, doi: 10.1101/cshperspect.a003590.

Yu, H., S. Zhang, and J.C. Chaput. Unpublished Data.

Yusupov, M.M., G.Z. Yusupova, A. Baucom, K. Lieberman, T.N. Earnest, J.H.D. Cate, and H.F. Noller. 2001. Crystal structure of the ribosome at 5.5 Å resolution. *Nature* 292: 883–896.

Zhang, L., A.E. Peritz, P.J. Carroll, and E. Meggers. 2006. Synthesis of glycol nucleic acids. *Synthesis* 4: 645-653.

Zhang, L., A. Peritz, and E. Meggers. 2005. A simple glycol nucleic acid. *J. Am. Chem. Soc.* 127: 4174-4175.

Zhang, R.S., E.O. McCullum, and J.C. Chaput. 2008. Synthesis of two mirror image 4-helix junctions derived from glycerol nucleic acid. *J. Am. Chem. Soc.* 130: 5846-5847.

Zhang, S., C. Switzer, and J.C. Chaput. 2010. The resurgence of acyclic nucleic acids. *Chem. Biodivers.* 7: 245-258.

Zou, K., A. Horhota, B. Yu, J.W. Szostak, and L.W. McLaughlin. 2005. Synthesis of α–L-threofuranosyl nucleoside triphosphates (tNTPs). *Org. Lett.* 7: 1485-1487.

APPENDIX A

SUPPLEMENTAL NMR SPECTROSCOPY

δ=2.61 (dd, J=2.4, 5.1 Hz, 1 H), 2.76 (m, 1 H), 3.12 (m, 2 H), 3.29 (m, 1 H), 3.79 (s, 6 H), 6.83 (m, 4 H), 7.05-7.40 (m, 7 H), 7.45 (m, 2 H).

Tritylated R-Glycidol for S-GNA ZS_C_

Sampleexphloromethylated polyethersulfone
File: exp
Pulse Sequence: s2pul
Pulse Sequence: s2pul

Solvent: cdcl3
Temp. 26.0 C / 299.1 K
Operator: walkup
VNMRS-400 "400MR"

Relax. delay 1.000 sec
Pulse 45.0 degrees
Acq. time 2.049 sec
Width 6410.3 Hz
8 repetitions
OBSERVE H1, 399.6496696 MHz
DATA PROCESSING
Resol. enhancement -0.0 Hz
FT size 65536
Total time 0 min, 31 sec

δ=3.02 (dd, J=6.3, 9.6 Hz, 1H), 3.25 (dd, J=5.7, 9.3 Hz, 1H), 3.78 (s, 6H), 4.18 (m, 1H), 4.29 (dd, J=6.6, 14.4 Hz, 1H), 4.40 (dd, J=2.4, 14.2 Hz, 1H), 5.75 (br s, 2H), 6.80 (m, 4H), 7.20-7.30 (m, 7H), 7.37-7.40 (m, 2H), 7.72 (s, 1H), 8.26 (s, 1H).

ZS-A-76

DMT_Adenine_SGNA

File: exp

Pulse Sequence: s2pul

Solvent: cdcl3
Temp. 26.0 C / 299.1 K
Operator: walkup
VNMRS-400 "400MR"

Relax. delay 1.000 sec
Pulse 45.0 degrees
Acq. time 2.049 sec
Width 6410.3 Hz
8 repetitions
OBSERVE H1, 399.6496984 MHz
DATA PROCESSING
Resol. enhancement -0.0 Hz
FT size 65536
Total time 0 min, 30 sec

ZS_B_84

Automation directory: /home/walkup/vnmrsys/data/auto_2007.06.4419
File : exp
Sample id : tmpstudy
Sample : ZS_B_84

Pulse Sequence: s2pul

δ=3.14 (dd, J=6.0, 9.5 Hz, 1H), 3.21 (dd, J=6.0, 9.5 Hz, 1H), 3.78 (s, 6H), 4.34 (dd, J=7.2, 14.2 Hz, 1H), 4.48 (dd, J=2.8, 14.2 Hz, 1H), 4.19 (m, 1H), 6.81 (d, J=9.0 Hz, 4H), 7.21 (t, J=7.2 Hz, 1H), 7.25-7.27 (m, 6H), 7.39 (d, J=7.0 Hz, 2H), 7.53 (t, J=7.8 Hz, 2H), 7.61 (t, J=7.5 Hz, 1H), 8.00 (s, 1H), 8.04 (d, J=7.5 Hz, 2H), 8.74 (s, 1H), 9.03 (br s, 1H).

193

^{1}H NMR (500 MHz, CDCl$_3$): δ=1.00 (d, J=6.5 Hz, 6H), 1.06 (d, J=7.0 Hz, 6H), 1.10 (d, J=7.0 Hz, 12H), 2.39 (t, J=6.2 Hz, 2H), 2.47 (q, J=6.2 Hz, 2H), 3.15 (dd, J=6.0, 10.0 Hz, 1H), 3.21-3.26 (m, 2H), 3.30 (dd, J=4.0, 12.5 Hz, 1H), 3.45-3.69 (m, 8H), 3.78 (s, 6H), 3.79 (s, 6H), 4.37 (m, 2H), 4.47-4.58 (m, 4H), 6.82 (m, 8H), 7.17-7.33 (m, 14H), 7.45 (m, 4H), 7.53 (m, 4H), 7.60 (m, 2H), 8.04 (m, 4H), 8.05 (s, 1H), 8.06 (s, 1H), 8.78 (s, 1H), 8.79 (s, 1H), 9.03 (br, 2H).

^{31}P NMR (202 MHz, CDCl$_3$): δ=148.98, 149.59.

P31_GNA_phosphoramidite_A

Automation directory: /home/walkup/vnmrsys/data/auto_2007.06.04
File : exp
Sample id : tmpstudy
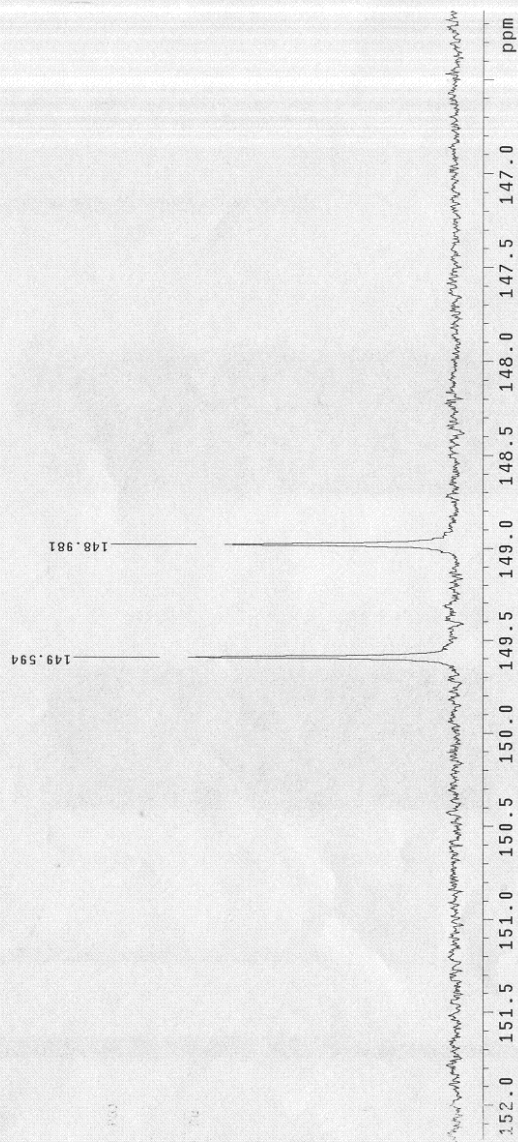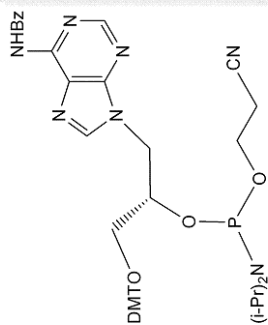
Pulse Sequence: s2pul

148.981

149.594

152.0 151.5 151.0 150.5 150.0 149.5 149.0 148.5 148.0 147.5 147.0 ppm

δ= 3.12 (dd, J=5.7, 9.6 Hz, 1H), 3.26 (dd, J=5.6, 9.4 Hz, 1H), 3.78 (s, 6H), 3.84 (dd, J=6.8, 13.8 Hz, 1H), 4.22 (m, 1H), 4.36 (dd, J=2.7, 13.8 Hz, 1H), 6.84 (m, 4H), 7.20-7.64 (m, 14H), 7.89 (s, 1H), 7.90 (s, 1H), 8.67 (s br, 1H).

ZS_C_56

DMT-C for S-GNA

File: exp

Pulse Sequence: s2pul
Solvent: cdcl3
Temp. 26.0 C / 299.1 K
Operator: walkup
VNMRS-400 "400MR"

Relax. delay 1.000 sec
Pulse 45.0 degrees
Acq. time 2.049 sec
Width 6410.3 Hz
8 repetitions
OBSERVE H1, 399.6497091 MHz
DATA PROCESSING
Resol. enhancement -0.0 Hz
FT size 65536
Total time 0 min, 31 sec

196

δ=1.12-1.16 (m, 16H), 1.27 (t, J=7.2 Hz, 8H), 2.43 (t, J=6.4 Hz, 2H), 2.55-2.77 (m, 3H), 3.15 (dd, J=3.0, 10.0 Hz, 1H), 3.22 (dd, J=4.5, 10.0 Hz, 1H), 3.28 (dd, J=4.0, 10.0 Hz, 1H), 3.42-3.90 (m, 10H), 3.78 (s, 6H), 3.79 (s, 6H), 4.31-4.45 (m, 4H), 6.82-6.85 (m, 8H), 7.19-7.36 (m, 16 H), 7.45-7.54 (m, 8H), 7.58 (m, 4H), 7.90 (d, J=7.5 Hz, 4H), 8.60 (br, 2H).

^{1}H NMR (500 MHz, DMSO-d_6): δ=3.30 (m, 1H), 3.40 (m, 1H), 3.81(m, 1H), 3.90 (dd, J=8.8, 13.8 Hz, 1H), 4.18 (dd, J=3.2, 13.8 Hz, 1H), 4.81 (t, J=5.5 Hz, 1H), 5.09 (d, J=5.5 Hz, 1H), 6.88 (s br, 2H), 8.01 (s, 1H).

Automation directory: /home/walkup/vnmrsys/data/auto_2007.09.05_03
File : exp
Sample id : tmpstudy
Sample : ZS_B_96
Pulse Sequence: s2pul

^{1}H NMR (500 MHz, DMSO-d_6): δ=3.28 (m, 1H), 3.40 (m, 1H), 3.76 (m, 1H), 3.80 (dd, J=8.0, 13.5 Hz, 1H), 4.08 (dd, J=3.5, 13.5 Hz, 1H), 4.78 (t, J=5.5 Hz, 1H), 5.06 (d, J=5.0 Hz, 1H), 6.46 (s, 2H), 7.59 (s, 1H), 10.54 (s, 1H).

ZS_B_116
Automation directory: /home/walkup/vnmrsys/data/auto_2007.09.05_03
File : exp
Sample id : tmpstudy
Sample : ZS_B_116

Pulse Sequence: s2pul

^1H NMR (300 MHz, DMSO-d_6): δ=1.08 (d, J=7.2 Hz, 6H), 2.75 (sept, J=7.1 Hz, 1H), 3.32 (m, 1H), 3.38 (m, 1H), 3.78 (m, 1H), 3.90 (dd, J=8.7, 13.8 Hz, 1H), 4.18 (dd, J=3.3, 13.8 Hz, 1H), 4.80 (t, J=5.4 Hz, 1H), 5.07 (d, J=5.7 Hz, 1H), 7.86 (s, 1H), 11.65 (s br, 1H), 12.02 (s br, 1H).

^1H NMR (500 MHz, CDCl$_3$): δ=1.27 (t, J=7.0 Hz, 6H), 2.62 (sept, J=6.8 Hz, 1H), 3.18 (dd, J=5.8, 9.8 Hz, 1H), 3.27 (dd, J=5.0, 9.5 Hz, 1H), 3.80 (s, 6H), 4.03 (dd, J=8.0, 14.0 Hz, 1H), 4.25 (dd, J=2.5, 14.0 Hz, 1H), 4.36 (m, 1H), 4.81 (br, 1H), 6.83 (d, J=8.5 Hz, 4H), 7.22 (t, J=7.2 Hz, 1H), 7.26-7.34 (m, 6H), 7.45 (d, J=7.0 Hz, 2H), 7.55 (s, 1H), 8.43 (br, 1H), 11.79 (br, 1H).

201
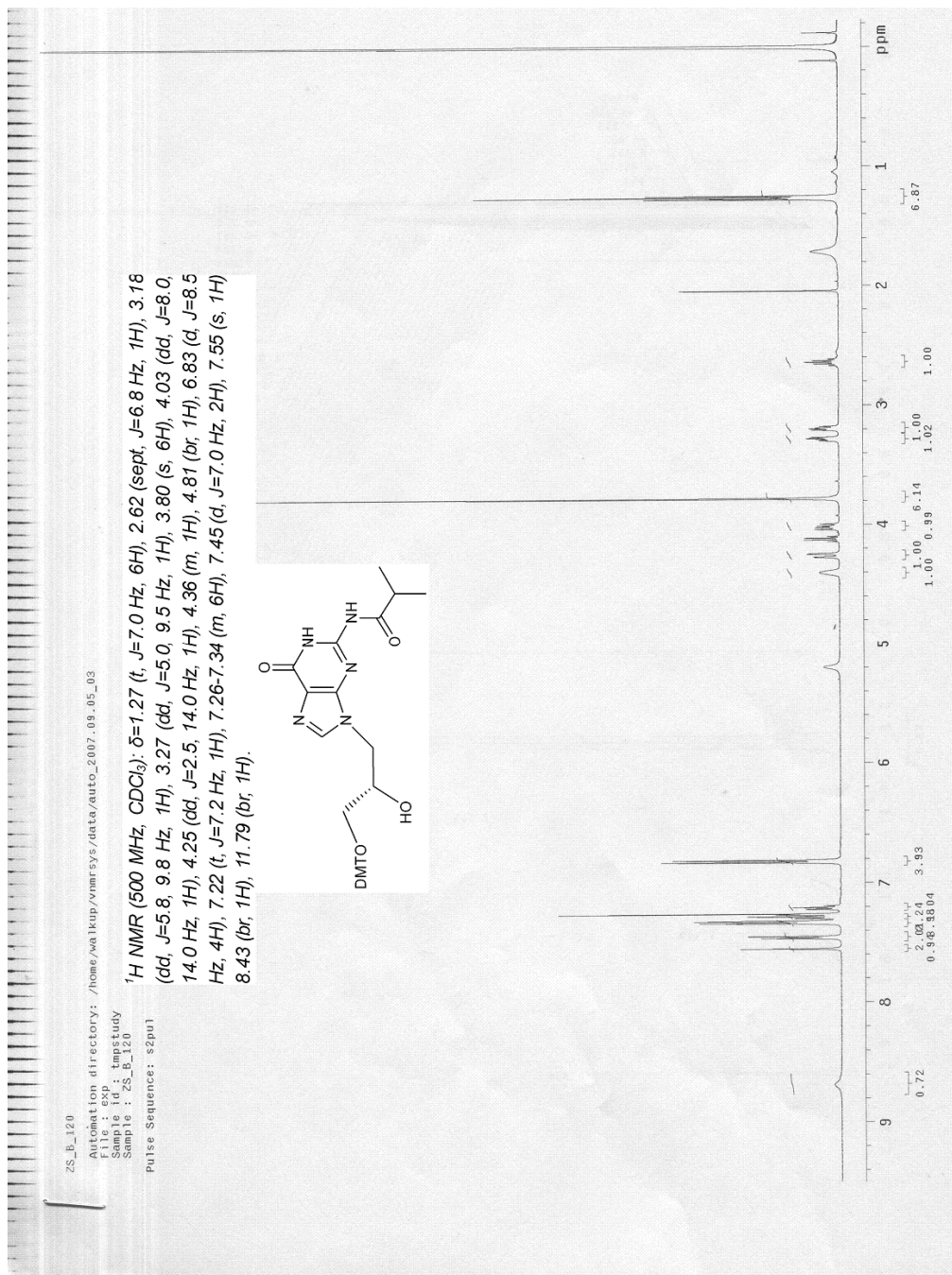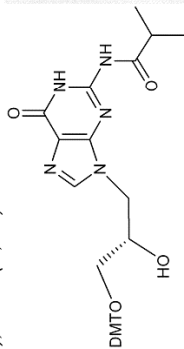
¹H NMR (500 MHz, CDCl₃): δ=1.06 (t, J=7.0 Hz, 2H), 1.10-1.15 (m, 17H), 1.20-1.29 (m, 17H), 2.47 (t, J=6.0 Hz, 1H), 2.52-2.65 (m, 4H), 3.02 (dd, J=4.0, 9.5 Hz, 1H), 3.08 (dd, J=5.2, 9.8 Hz, 1H), 3.13-3.23 (m, 2H), 3.46-3.82 (m, 10H), 3.78 (s, 6H), 3.79 (s, 6H), 4.26-4.38 (m, 5H), 4.53 (m, 1H), 6.80 (m, 8H), 7.18-7.33 (m, 14H), 7.42-7.48 (m, 4H), 7.49 (s, 1H), 7.55 (s, 1H), 8.38 (br, 1H), 8.70 (br, 1H).
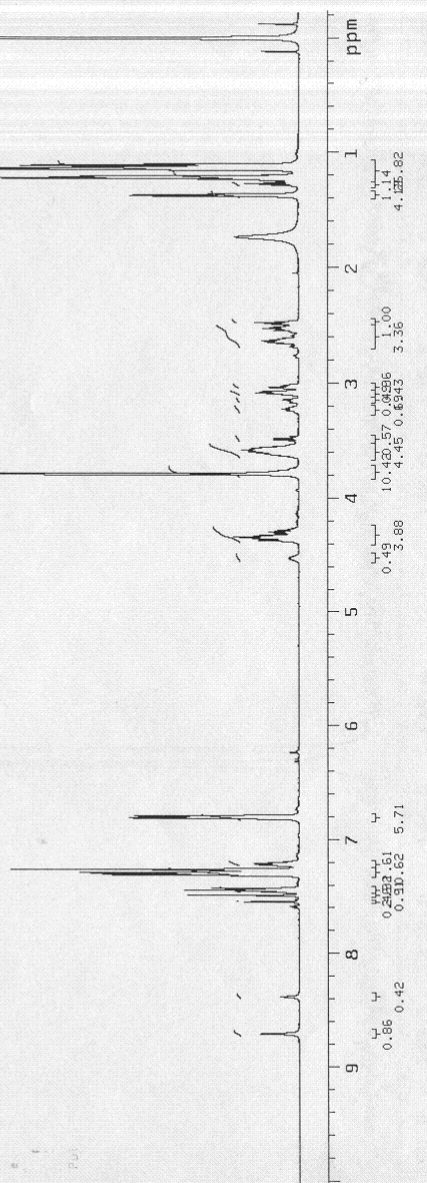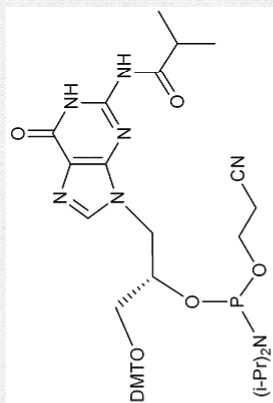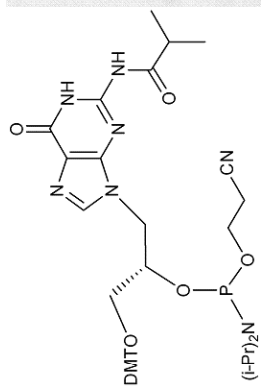
File : home/inova400/UserData/zhangsu/HI_ZS_B_92_09_21_07.fid
Sample id : tmpstudy
Pulse Sequence: s2pul

202

^{31}P NMR (202 MHz, CDCl$_3$): δ=148.20, 148.28.

^1H NMR (500 MHz, CDCl$_3$): δ= 1.73 (s, 3H), 3.07 (d, J=5.6 Hz, 2H), 3.55 (dd, J=7.2, 14.0 Hz, 1H), 3.68 (s, 6H), 3.89-3.96 (m, 2H), 6.73 (m, 4H), 6.94 (d, J=1.2 Hz, 1H), 7.10-7.22 (m, 7H), 7.30 (m, 2H).

STANDARD 1H OBSERVE

Data Collected on:
asnmr6-inova500
Archive directory:
    /home/gemini/vnmrsys6/data
Sample directory:

File: dmts+glycidol zs_a_72
Pulse Sequence: s2pul
Solvent: CDCl3

Pulse 45.0 degrees
Acq. time 1.821 sec
Width 4498.6 Hz
128 repetitions
OBSERVE H1, 300.0770475 MHz
DATA PROCESSING
FT size 16384
Total time 3 min

^1H NMR (300 MHz, CDCl$_3$): δ=1.12-1.18 (m, 24H), 1.84 (m, 6H), 2.41 (t, J=6.3 Hz, 2H), 2.59 (t, J=6.3 Hz, 2H), 3.16-3.32 (m, 4H), 3.49-3.88 (m, 10H), 3.78 (s, 6H), 3.79 (s, 6H), 4.03-4.23 (m, 4H), 6.83 (m, 8H), 7.05 (m, 2H), 7.18-7.36 (m, 14H), 7.44-7.46 (m, 4H).

1H-NMR (300 MHz, D2O): δ=3.63 (dd, J=11.6, 7.8, 1H), 3.70 (dd, J=11.4, 5.4, 1H), 3.98 (ddd, J=12.6, 5.4, 2.1, 1H), 4.04 (d, J=2.1, 1H).

Sample: ZS_D_34
File: exp

Pulse Sequence: s2pul

Solvent: d2o
Ambient temperature
Operator: walkup
VNMRS-400 "400MR"

Relax. delay 1.000 sec
Pulse 45.0 degrees
Acq. time 2.049 sec
Width 6410.3 Hz
16 repetitions
OBSERVE H1, 399.6483305 MHz
DATA PROCESSING
Line broadening 0.2 Hz
FT size 65536
Total time 0 min, 55 sec

1H-NMR (300 MHz, CDCl3): δ=4.38 (m, 1H), 5.02 (m, 1H), 5.81 (m, 2H),7.45-8.12 (m, 10H).

207

1H-NMR (400 MHz, CDCl3): δ=3.97 (dd, J=8.1,3.0, 1H), 4.26 (dd, J=10.2, 4.0), 4.60 (m, 2H), 5.47 (t, J=4.0, 1H), 5.53 (m, 2H), 5.57 (s, 1H), 5.75 (m, 1H), 5.85 (d, J=4.0, 1H), 7.42-7.47 (m, 8H), 7.53-7.59 (m, 4H), 8.01-8.10 (m, 8H).

208

δ=2.04 (s, 3H), 2.16 (s, 3H), 4.05 (dd, J=10.4, 3.6, 1H), 4.22 (dd, J=10.6, 3.8, 1H), 4.62 (dd, J=10.8, 6.4, 1H), 4.67 (dd, J=10.4, 6.0, 1H), 5.56 (m, 1H), 5.65 (s, 1H), 5.70 (t, J=4.6, 1H), 5.76 (m, 1H), 6.38 (s, 1H), 6.64 (d, J=4.4, 1H), 7.26-7.50 (m, 8H), 7.58-7.64 (m, 4H), 8.05-8.10 (m, 8H).

Sample: ZS_0_40
File: exp

Pulse Sequence: s2pul

Solvent: cdcl3
Ambient temperature
Operator: wakup
VNMRS-400 "400MR"

Relax. delay 1.000 sec
Pulse 45.0 degrees
Acq. time 2.049 sec
Width 6410.3 Hz
16 repetitions
OBSERVE H1, 399.6479686 MHz
DATA PROCESSING
Line broadening 0.2 Hz
FT size 65536
Total time 0 min, 55 sec

2-isomer

1H-NMR (300 MHz, CDCl3): δ=4.15 (dd, J=11.6, 3.0, 1H), 4.70 (dd, J= 11.8, 6.0, 1H), 5.84 (m, 2H), 6.86 (d, J=4.0, 1H), 7.38-7.55 (m, 9H), 7.95-7.20 (m, 6H).

210

1H-NMR (300 MHz, CDCl3): δ=4.32 (dd, J=10.5, 3.0, 1H), 4.74 (dd, J=10.7, 5.9, 1H), 5.64 (dd, J=5.1, 3.0, 1H), 5.82 (s, 1H), 6.64 (s, 1H), 7.40-7.51 (m, 6H), 7.56-7.65 (m, 3H), 8.06-8.14 (m, 6H).

1H-NMR (300 MHz, CDCl3): δ=4.58 (dd, J=11.1, 2.0, 1H), 4.68 (dd, J=11.1, 3.3, 1H), 5.77 (m, 1H), 6.30 (dd, J=2.0, 0.5, 1H), 6.47 (d, J=1.8, 1H), 7.43-7.67 (m, 9H), 7.86-7.89 (m, 2H), 8.03-8.12 (m, 4H), 8.33 (s, 1H), 8.78 (s, 1H), 9.02 (br, 1H).

1H-NMR (300 MHz, DMSO-d₆): δ=4.17 (m, 3H), 4.44 (m, 1H), 5.60 (br, 1H), 5.94 (d, J=4.5, 1H), 6.06 (d, J=2.0, 1H), 7.56 (m, 2H), 7.64 (m, 1H), 8.05 (m, 2H), 8.55 (s, 1H), 8.76 (s, 1H), 11.17 (br, 1H).

1H-NMR (300 MHz, CDCl$_3$): δ=3.76 (s, 3H), 3.77 (s, 3H),
3.80 (m, 2H), 4.26 (m, 1H), 4.33 (m, 1H), 5.90 (d, J=2.1, 1H),
6.80 (m, 4H), 7.28-7.61 (m, 12H), 8.02 (d, J=7.2, 2H),
8.30 (s, 1H), 8.78 (s, 1H), 9.02 (br, 1H).

1H-NMR (300 MHz, CDCl$_3$): δ=1.10-1.26 (m, 24H), 2.35-2.44 (m, 2H), 2.53-2.60 (m, 2H), 3.44 (m, 2H), 3.55-3.89 (m, 32H), 4.26 (m, 1H), 4.36 (m, 1H), 4.50 (m, 1H), 4.98 (m, 1H), 6.16 (s, 1H), 6.23 (s, 1H), 6.77-6.81 (m, 8H), 7.08-7.28 (m, 18H), 7.52-7.62 (m, 6H), 8.02-8.05 (m, 4H), 8.34 (s, 1H), 8.39 (s, 1H), 8.80 (s, 1H), 8.81 (s, 1H), 9.00 (br, 2H).

1H-NMR (400 MHz, CDCl3): δ=4.57 (d, J=11.2, 1H), 4.61 (dd, J=11.2, 3.6, 1H), 5.61 (d, J=3.2, 1H), 5.84 (s, 1H), 6.30 (s, 1H), 7.40-7.66 (m, 10H), 7.87 (m, 2H), 7.92 (d, J=7.6, 1H), 8.03 (d, J=7.2, 1H), 8.09 (m, 2H), 8.75 (br, s, 1H).

Sample: ZS_D_44
File: exp

Pulse Sequence: s2pul
Solvent: cdcl3
Ambient temperature
Operator: walkup
VNMRS-400 "400NR"

Relax. delay 1.000 sec
Pulse 45.0 degrees
Acq. time 2.049 sec
Width 6410.3 Hz
16 repetitions
OBSERVE H1, 399.6479681 MHz
DATA PROCESSING
Line broadening 0.2 Hz
FT size 65536
Total time 0 min, 55 sec

2.00
0.98
0.94
0.92
9.47
3.76
0.97
1.90
0.84

1H-NMR (400 MHz, DMSO-d$_6$): δ=4.03 (s, 1H), 4.08 (s, 1H), 4.19 (m, 2H), 5.66 (s, 1H), 7.30 (d, J=7.2, 1H), 7.52 (m, 2H), 7.63 (m, 1H), 8.00 (m, 2H), 8.11 (d, J=7.6, 1H).

Sample: ZS_D_48
File: /home/walkup/NMR_User_Data/New_SuZhang/ZS_D_48.fid

Pulse Sequence: s2pul

Solvent: dmso
Ambient temperature
Operator: walkup
File: ZS_D_48
VNMRS-400 "400MR"

Relax. delay 1.000 sec
Pulse 45.0 degrees
Acq. time 2.049 sec
Width 6410.3 Hz
32 repetitions
OBSERVE H1, 399.6498600 MHz
DATA PROCESSING
Line broadening 0.2 Hz
FT size 65536
Total time 1 min, 44 sec
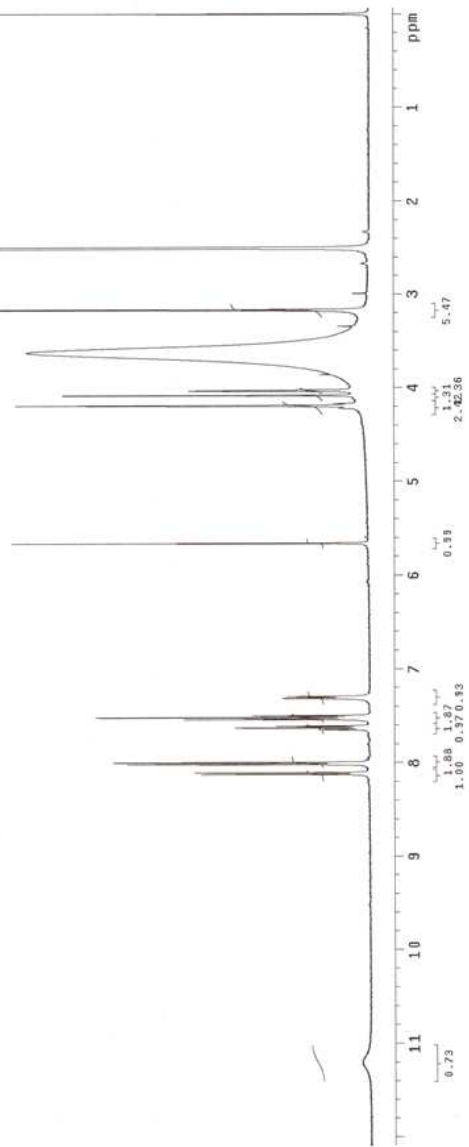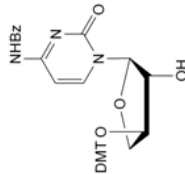
Sample: 3-O-DMT-C TNA
File: exp

Pulse Sequence: s2pul
Solvent: cdcl3
Temp. 24.0 C / 297.1 K
Operator: walkup
VNMRS-500 "wormhole"

Relax. delay 1.000 sec
Pulse 45.0 degrees
Acq. time 2.049 sec
Width 8012.8 Hz
8 repetitions
OBSERVE H1, 499.9081785 MHz
DATA PROCESSING
Resol. enhancement -0.0 Hz
FT size 65536
Total time 0 min, 30 sec

1H-NMR (400 MHz, CDCl$_3$): δ=3.37 (d, J=9.6, 1H), 3.62 (dd, J=10.0, 3.2, 1H), 3.77 (2s, 6H), 3.95 (br, s, 1H), 4.26 (m, 2H), 5.64 (s, 1H), 6.80-6.84 (m, 4H), 7.20-7.36 (m, 9H), 7.53-7.66 (m, 4H), 7.91 (1H), 7.92 (d, J=7.6, 1H), 8.05 (d, J=7.6, 1H), 8.76 (br, s, 1H).

1H-NMR (400 MHz, CDCl$_3$): δ=2.08 (s, 3H), 3.04 (d, J=10.4, 1H), 3.60 (dd, J=10.4, 3.6, 1H), 3.76-3.77 (2s, 6H), 4.24 (d, J= 3.6, 1H), 5.57 (s, 1H), 5.95 (s, 1H), 6.80-6.85 (m, 4H), 7.22-7.65 (m, 13H), 7.92-7.96 (m, 3H), 8.80 (br, s, 1H).

Sample: ZS_D_102_pure
File: exp

Pulse Sequence: s2pul

Solvent: cdcl3
Temp: 24.0 C / 297.1 K
Operator: walkup
VNMRS-400 "400MR"

Relax. delay 1.000 sec
Pulse 45.0 degrees
Acq. time 2.049 sec
Width 6410.3 Hz
8 repetitions
OBSERVE H1, 399.6479678 MHz
DATA PROCESSING
Line broadening 0.2 Hz
FT size 65536
Total time 0 min, 31 sec

219

1H-NMR (400 MHz, DMSO-d$_6$): δ=2.10 (s, 3H), 4.16 (m, 2H), 4.25 (d, J=9.2, 1H), 5.10 (s, 1H), 5.70 (d, J=2.4, 1H), 5.79 (s, 1H), 7.36 (d, J=7.2, 1H), 7.50 (m, 2H), 7.62 (m, 1H), 7.99 (m, 2H), 8.13 (d, J=8.0, 1H), 11.24 (s, 1H).

^{31}P NMR (161.8 MHz, D$_2$O): δ=-10.4 (d, J=19.6), -12.6 (d, J=20.0), -22.9 (t, J=19.6).

File: exp._tCTP_04-19-10

Pulse Sequence: s2pul

Solvent: d2o
Temp. 24.0 C / 297.1 K
Operator: walkup
VNMRS-400 "400MR"

Relax. delay 1.000 sec
Pulse 45.0 degrees
Acq. time 1.600 sec
Width 8928.6 Hz
64 repetitions
OBSERVE P31, 161.7808816 MHz
DECOUPLE H1, 399.6510008 MHz
Power 40 dB
continuously on
WALTZ-16 modulated
DATA PROCESSING
Line broadening 1.0 Hz
FT size 32768
Total time 2 min, 46 sec

221

^{1}H-NMR (400 MHz, CDCl$_3$): δ=1.06-1.22 (m, 24H), 2.33, 2.68 (2m, 4H), 3.13 (d, J=10, 1H), 3.22 (d, J=9.6, 1H), 3.46-3.52 (m, 4H), 3.68, 3.69 (2s, 12H), 3.78-3.83 (m, 4H), 3.98-4.06 (m, 4H), 4.17 (d, J=3.6, 1H), 4.47 (d, J=8.4, 1H), 4.65 (d, J=9.2, 1H), 5.79 (s, 1H), 5.81 (s, 1H), 6.72-6.76 (m, 8H), 7.09-7.19 (m, 14H), 7.45-7.58 (m, 8H), 7.84-7.85 (m, 6H), 7.91 (d, J=7.6, 1H), 7.96 (d, J=7.2, 1H), 8.58 (br, 2H).

³¹P-NMR (161 MHz, CDCl₃): δ=150.60, 152.58.

1H-NMR (400 MHz, DMSO-d6): δ=4.64 (dd, J=10.0, 6.4, 1H), 4.82 (dd, J=9.6, 4.8, 1H), 5.89 (m, 1H), 6.49 (m, 2H), 7.33-8.10 (m, 20H), 8.61 (s, 1H), 10.99 (s, 1H), 11.27 (s, 1H).
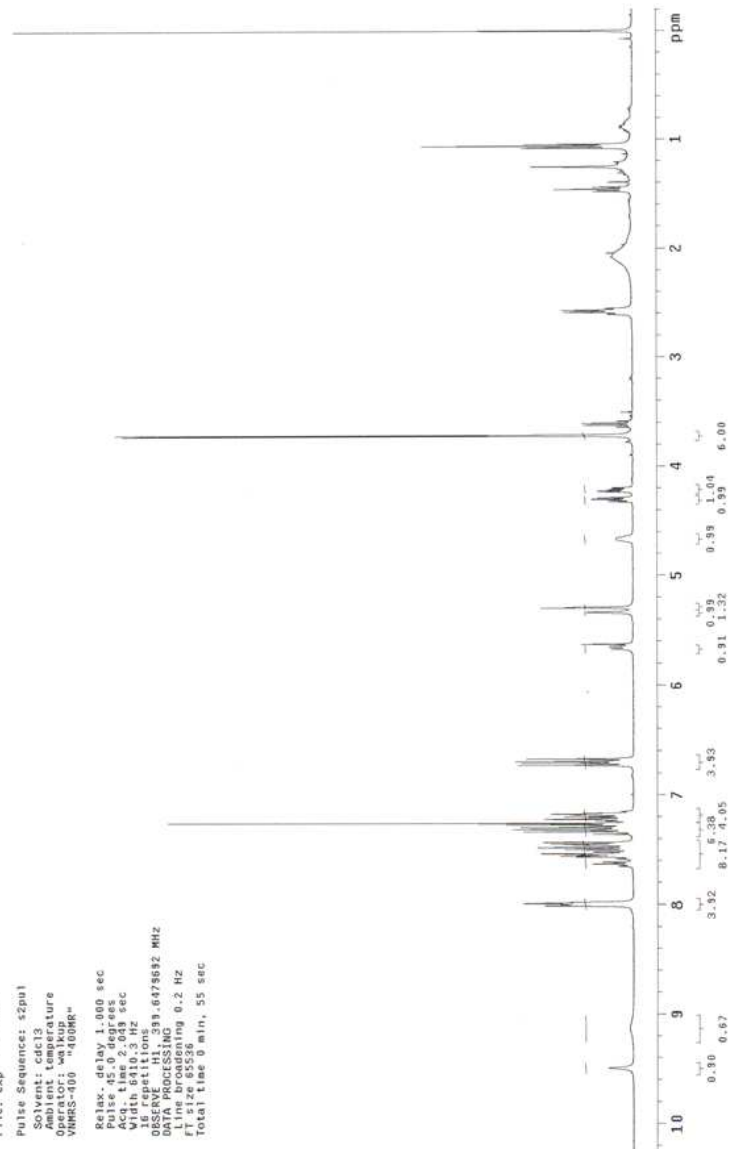
Sample: ZS_D_92
File: exp

Pulse Sequence: s2pul

Solvent: dmso
Ambient temperature
Operator: walkup
VNMRS-400 "400MR5"

Relax. delay 1.000 sec
Pulse 45.0 degrees
Acq. time 2.049 sec
Width 6410.3 Hz
16 repetitions
OBSERVE H1, 399.6498651 MHz
DATA PROCESSING
Line broadening 0.2 Hz
FT size 65536
Total time 0 min, 55 sec

^{1}H-NMR (400 MHz, DMSO-d$_6$): δ=4.07 (dd, J=9.2, 2.4, 1H), 4.14 (dd, J=9.2, 3.6, 1H), 4.18 (m, 1H), 4.41 (m, 1H), 5.61 (br, s, 1H), 5.90 (d, J=4.0, 1H), 5.95 (br, s, 1H), 7.46-7.59 (m, 6H), 7.98-8.05 (m, 4H), 8.05 (s, 1H), 11.00 (br, s, 2H).

Sample: ZS_D_94
File: exp

Pulse Sequence: s2pul

Solvent: dmso
Temp. 24.0 C / 297.1 K
Operator: walkup
VNMRS-400 "400MR"

Relax. delay 1.000 sec
Pulse 45.0 degrees
Acq. time 2.049 sec
Width 6410.3 Hz
8 repetitions
OBSERVE H1, 399.6498644 MHz
DATA PROCESSING
Line broadening 0.2 Hz
FT size 65536
Total time 0 min, 31 sec

^{1}H-NMR (400 MHz, CDCl$_3$): δ=3.05 (dd, J=9.4, 6.2, 1H), 3.36 (dd, J=9.2, 6.2, 1H), 3.77 (2s, 6H), 4.44 (m, 1H), 4.76 (m, 1H), 5.79 (d, J=4.4, 1H), 6.80-6.85 (m, 4H), 7.18-7.64 (m, 15H), 7.99-8.06 (m, 5H), 9.43 (br, s, 1H), 9.64 (br, s, 1H).

Sample: ZS_D_68_1ST
File: exp

Pulse Sequence: s2pul
Solvent: cdcl3
Ambient temperature
Operator: walkup
VNMRS-400 "400MR"

Relax. delay 1.000 sec
Pulse 45.0 degrees
Acq. time 2.049 sec
Width 6410.3 Hz
16 repetitions
OBSERVE H1, 399.6479696 MHz
DATA PROCESSING
Line broadening 0.2 Hz
FT size 65536
Total time 0 min, 55 sec
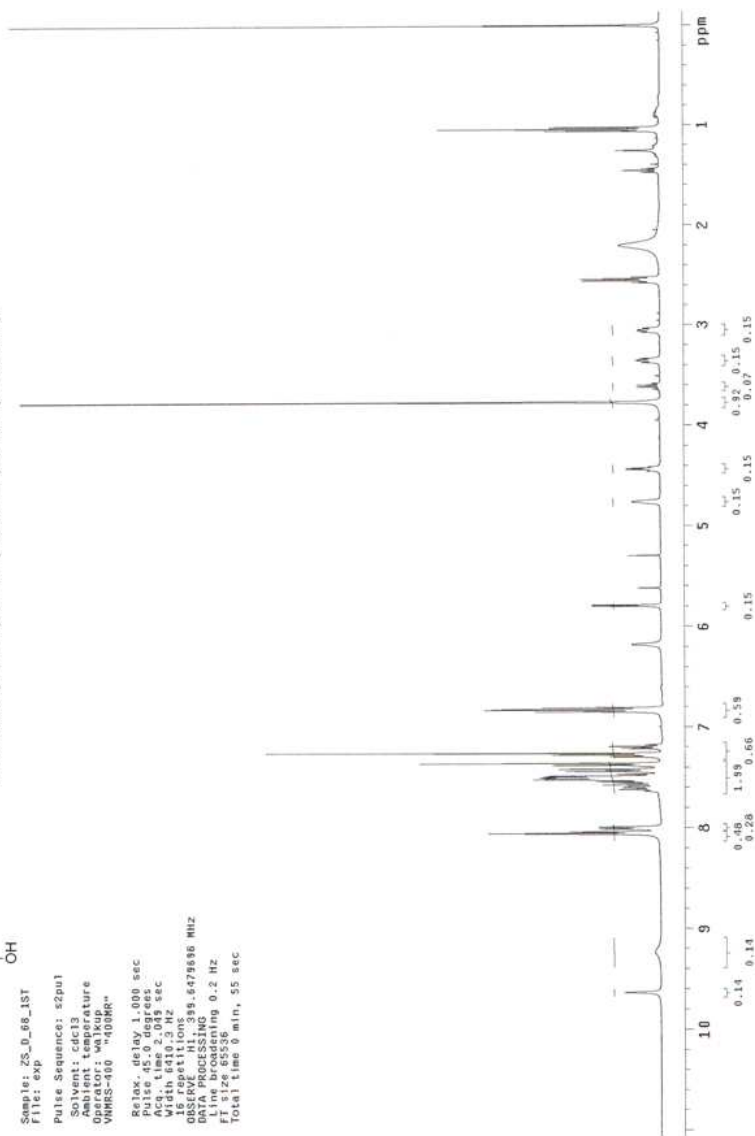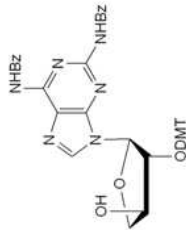
^{1}H-NMR (500 MHz, DMSO-d$_6$): δ=3.62 (2s, 6H), 3.76 (m, 1H), 3.97 (d, J=7.6, 1H), 4.08 (dd, J=7.6, 2.4, 1H), 4.44 (s, 1H), 5.36 (d, J=2.8, 1H), 5.92 (s, 1H), 6.70-6.77 (m, 4H), 7.14-7.65 (m, 16H), 7.98-8.20 (m, 4H), 11.01 (s, 1H), 11.19 (s, 1H).

Sample: ZS_D_68_2ND
File: exp

Pulse Sequence: s2pul
Solvent: cdcl3
Ambient temperature
Operator: walkup
VNMRS-400 "400MR"

Relax. delay 1.000 sec
Pulse 45.0 degrees
Acq. time 2.049 sec
Width 6410.3 Hz
16 repetitions
OBSERVE H1, 399.6479692 MHz
DATA PROCESSING
Line broadening 0.2 Hz
FT size 65536
Total time 0 min, 55 sec

^{1}H-NMR (400 MHz, CDCl$_3$): δ=3.77 (2s, 6H), 4.05 (d, J=9.6, 1H), 4.17 (dd, J=9.6, 3.6, 1H), 4.27 (m, 1H), 4.67 (d, J=5.6, 1H), 5.21 (d, J=1.2, 1H), 5.42 (br, s, 2H), 6.72 (br, 2H), 6.72-6.79 (m, 4H), 7.22-7.41 (m, 9H), 7.58 (s, 1H).

Sample: TNA_D_deprotection
File: exp

Pulse Sequence: s2pul

Solvent: cdcl3
Temp. 24.0 C / 297.1 K
Operator: walkup
VNMRS-400 "400MR"

Relax. delay 1.000 sec
Pulse 45.0 degrees
Acq. time 2.049 sec
Width 6410.3 Hz
16 repetitions
OBSERVE H1, 399.6479682 MHz
DATA PROCESSING
Line broadening 0.2 Hz
FT size 65536
Total time 0 min, 55 sec
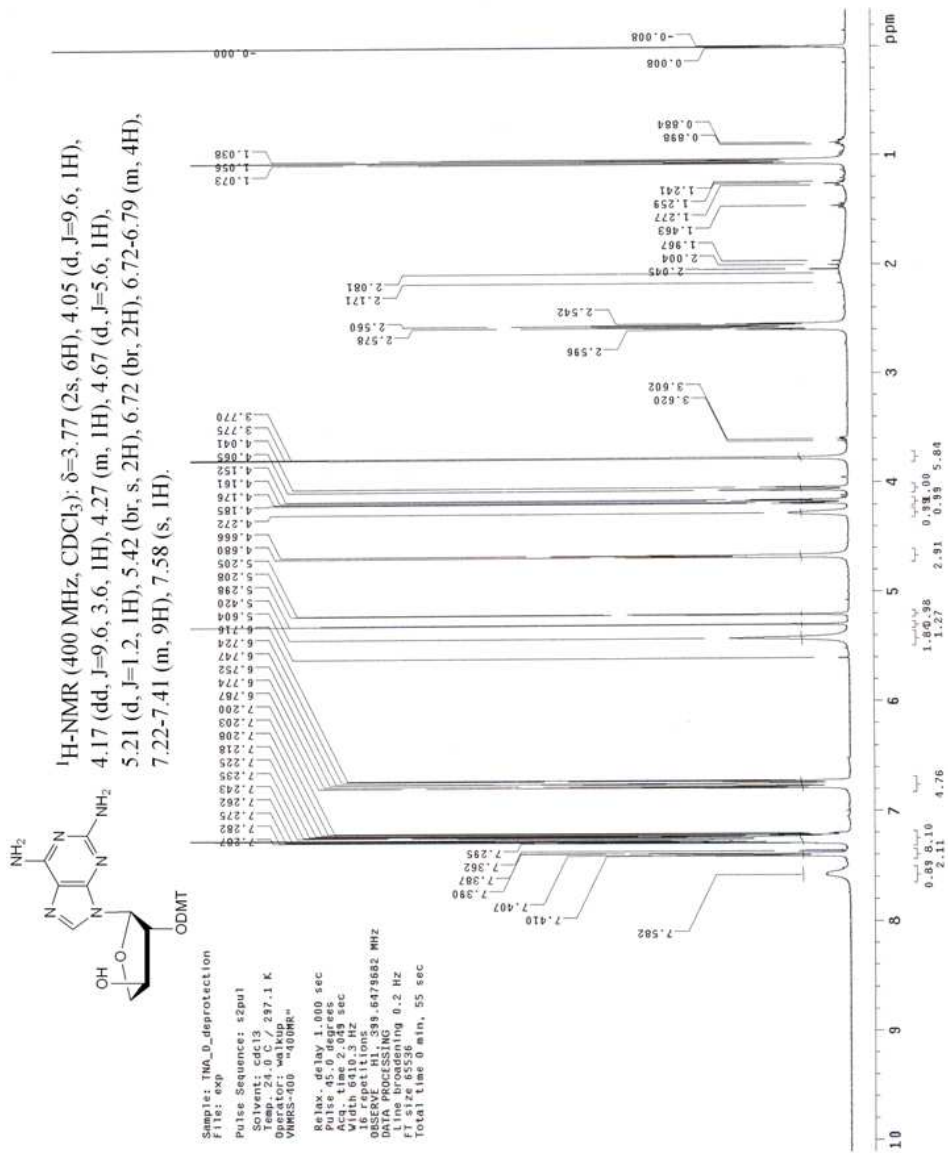
^{31}P NMR (161.8 MHz, D$_2$O): δ=-9.12 (d, J=20.0), -12.1 (d, J=20.0), -22.6 (t, J=18.4)
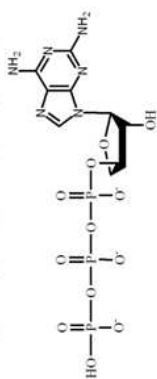
Sample: P31_ref_06_28_10
File: exp

Pulse Sequence: s2pul

Solvent: d2o
Temp. 19.0 C / 292.1 K
Operator: walkup
VNMRS-400 "400MR"

Relax. delay 1.000 sec
Pulse 45.0 degrees
Acq. time 1.600 sec
Width 9124.1 HZ
128 repetitions
OBSERVE P31, 161.7807952 MHz
DECOUPLE H1, 399.6510008 MHz
Power 40 dB
continuously on
WALTZ-16 modulated
DATA PROCESSING
Line broadening 1.0 Hz
FT size 32768
Total time 5 min, 34 sec

-22.721
-22.621
-22.493
-12.191
-12.067
-11.716
-9.244
-9.124
-9.000

ppm

229

^{1}H-NMR (400 MHz, CDCl$_3$): δ=1.08-1.30 (m, 24H), 2.39 (m, 2H), 2.60 (m, 2H), 3.36-3.39 (m, 1H), 3.54-3.90 (m, 8H), 3.72 (2s, 12H), 4.27 (m, 2H), 4.35 (m, 1H), 4.71 (d, J=8.0, 2H), 5.16 (d, J=9.6, 1H), 6.12 (s, 1H), 6.22 (s, 2H),6.78 (m, 8H), 7.12-7.35 (m, 17H), 7.46-7.50 (m, 13H), 7.94-8.04 (m, 8H), 8.26, 8.31 (2s, 2H), 8.90, 9.11, 9.23 (3br, 4H).

Sample: ZS_f_35
File: exp

Pulse Sequence: s2pul

Solvent: cdc13
Temp. 24.0 C / 297.1 K
Operator: walkup
VNMRS-400 "NMR400"

Relax. delay 1.000 sec
Pulse 45.0 degrees
Acq. time 2.049 sec
Width 6410.3 Hz
64 repetitions
OBSERVE H1, 399.6464127 MHz
DATA PROCESSING
Line broadening 0.1 Hz
FT size 65536
Total time 3 min, 21 sec

230
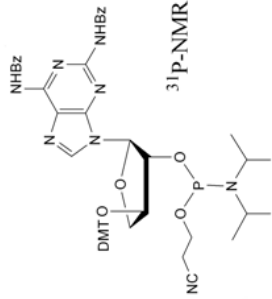
^{31}P-NMR (161 MHz, CDCl$_3$): δ=150.5, 151.9.

1H-NMR (400 MHz, CDCl₃): δ=2.48 (s, 3H), 4.57 (dd, J=11.0, 1.8, 1H), 4.66 (dd, J=11.0, 4.6, 1H), 5.74 (dd, J=2.8, 1.6, 1H), 6.21 (s, 1H), 6.32 (d, J=1.6, 1H), 7.34-8.09 (m, 20H), 8.23 (s, 1H).
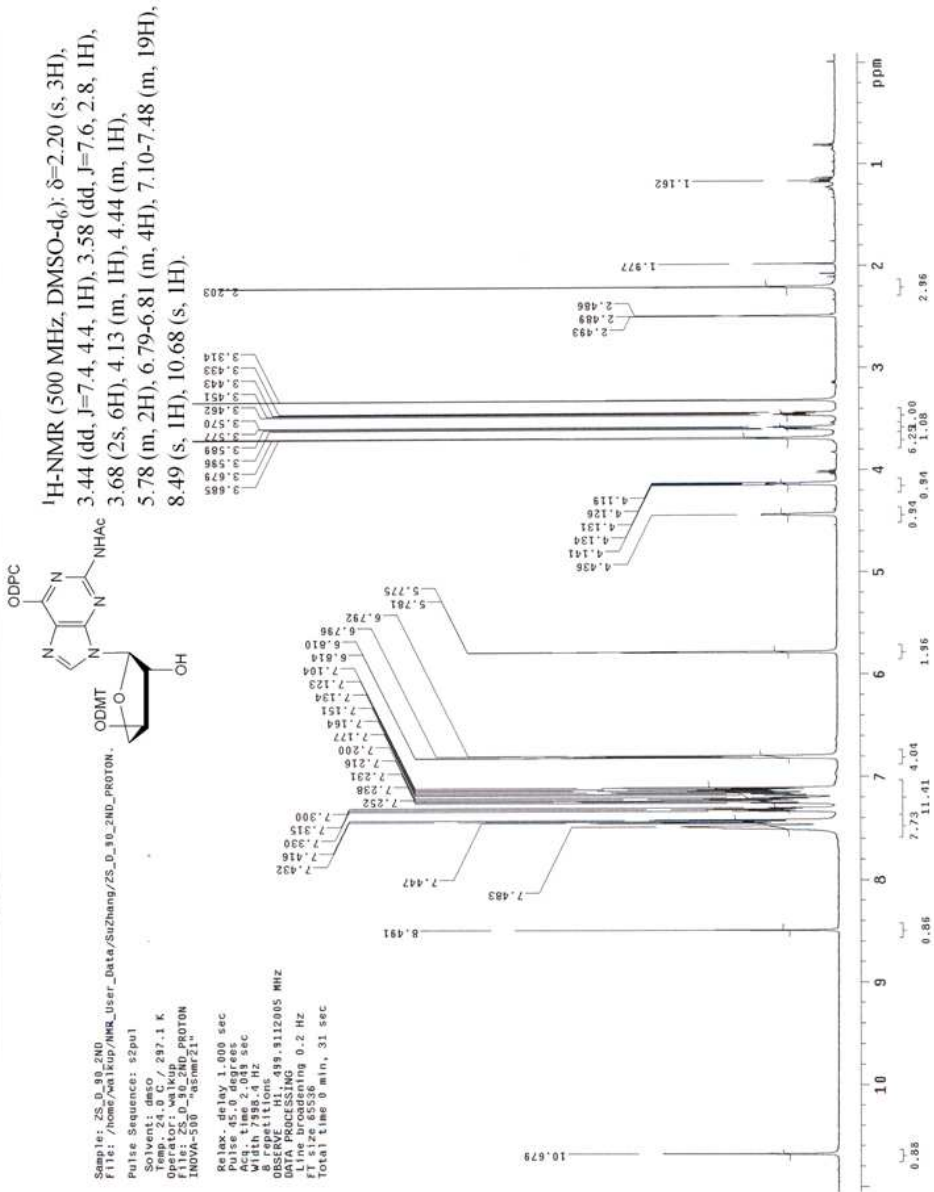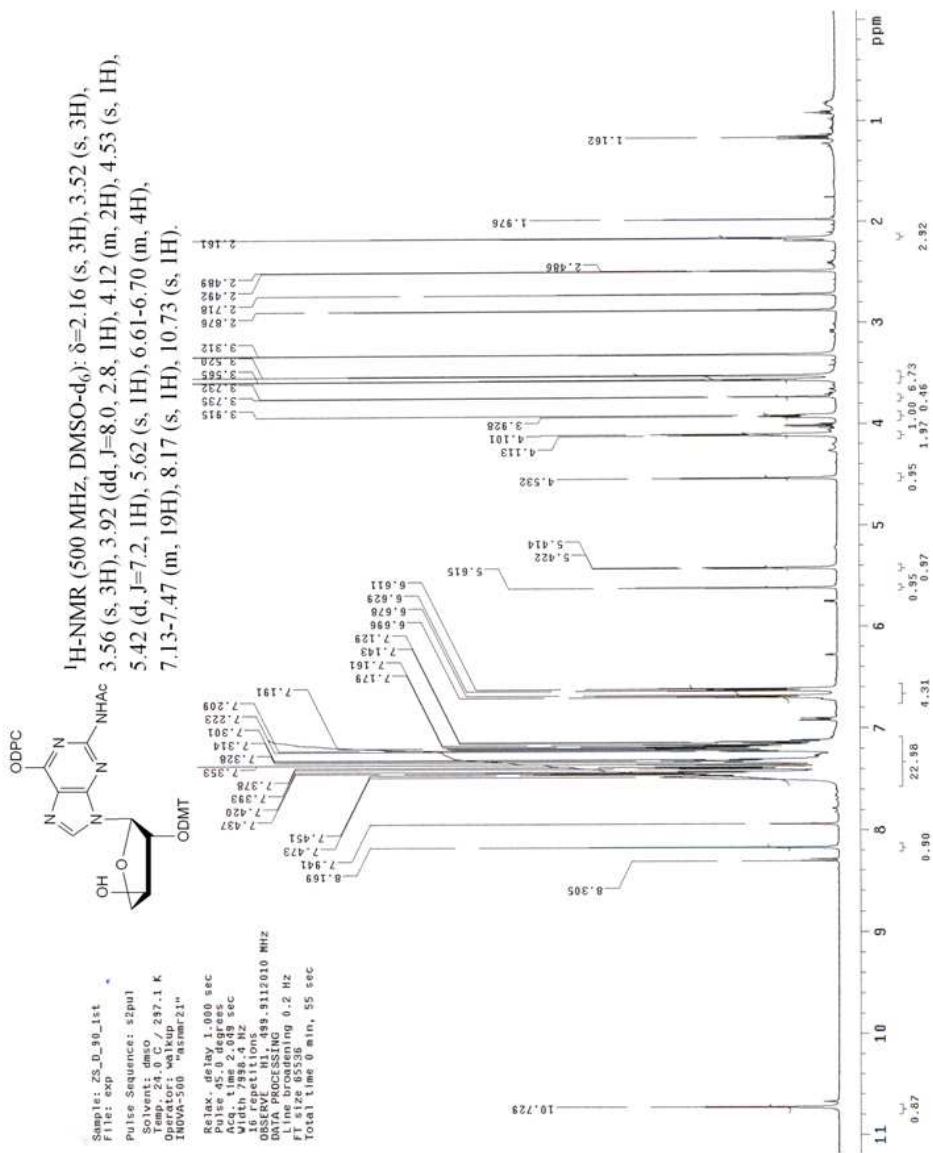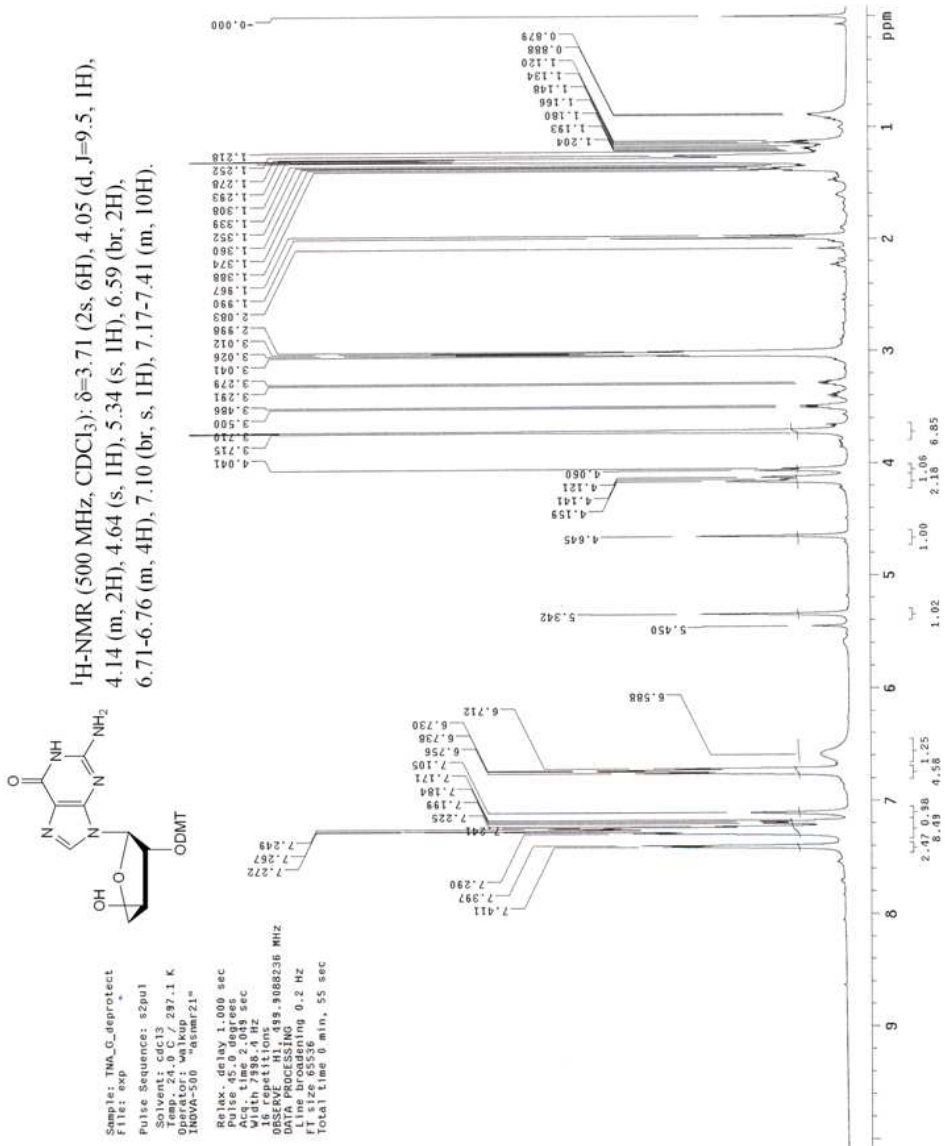
^{1}H-NMR (400 MHz, DMSO-d$_6$): δ=4.08-4.16 (m, 3H), 4.42 (m, 1H), 5.53 (d, J=3.2, 1H), 5.88 (d, J=4.4, 1H), 5.92 (d, J=1.6, 1H), 7.29-7.48 (m, 10H), 8.52 (s, 1H), 10.73 (br, s, 1H).

Sample: ZS_D_56
File: exp

Pulse Sequence: s2pul
  Solvent: dmso
  Ambient temperature
  Operator: walkup
  VNMRS-400 "400MR"

  Relax. delay 1.000 sec
  Pulse 45.0 degrees
  Acq. time 2.049 sec
  Width 6410.3 Hz
  32 repetitions
  OBSERVE H1, 399.6498652 MHz
  DATA PROCESSING
  Line broadening 0.2 Hz
  FT size 65536
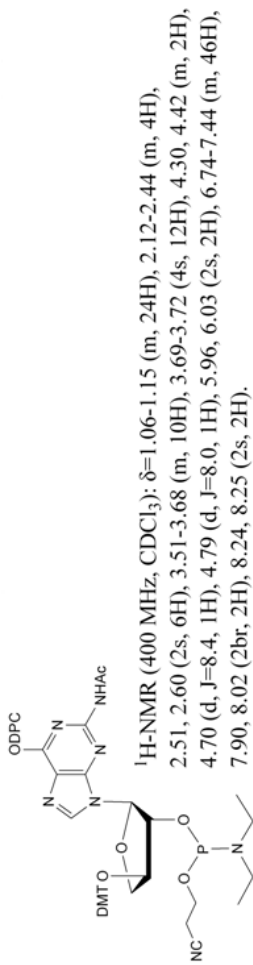 Total time 1 min, 43 sec

^{1}H-NMR (500 MHz, DMSO-d$_6$): δ=2.20 (s, 3H), 3.44 (dd, J=7.4, 4.4, 1H), 3.58 (dd, J=7.6, 2.8, 1H), 3.68 (2s, 6H), 4.13 (m, 1H), 4.44 (m, 1H), 5.78 (m, 2H), 6.79-6.81 (m, 4H), 7.10-7.48 (m, 19H), 8.49 (s, 1H), 10.68 (s, 1H).

¹H-NMR (500 MHz, DMSO-d₆): δ=2.16 (s, 3H), 3.52 (s, 3H), 3.56 (s, 3H), 3.92 (dd, J=8.0, 2.8, 1H), 4.12 (m, 2H), 4.53 (s, 1H), 5.42 (d, J=7.2, 1H), 5.62 (s, 1H), 6.61-6.70 (m, 4H), 7.13-7.47 (m, 19H), 8.17 (s, 1H), 10.73 (s, 1H).

^{1}H-NMR (500 MHz, CDCl$_3$): δ=3.71 (2s, 6H), 4.05 (d, J=9.5, 1H), 4.14 (m, 2H), 4.64 (s, 1H), 5.34 (s, 1H), 6.59 (br, 2H), 6.71-6.76 (m, 4H), 7.10 (br, s, 1H), 7.17-7.41 (m, 10H).

Sample: TNA_G_deprotect
File: exp

Pulse Sequence: s2pul

Solvent: cdcl3
Temp. 24.0 C / 297.1 K
Operator: walkup
INOVA-500 "astnmr21"

Relax. delay 1.000 sec
Pulse 45.0 degrees
Acq. time 2.045 sec
Width 7998.4 Hz
16 repetitions
OBSERVE H1 499.9088236 MHz
DATA PROCESSING
Line broadening 0.2 Hz
FT size 65536
Total time 0 min, 55 sec

^{31}P NMR (161.8 MHz, D$_2$O): δ=-8.57 (d, J=20.0), -12.2 (d, J=20.1), -22.6 (t, J=20.0).

-22.645

-12.244
-12.120

-8.574

Sample: RGF6[BE8q08]DS8510
File: exp

Pulse Sequence: s2pul

Solvent: d2o
Temp. 24.0 C / 297.1 K
Operator: walkup
VNMRS-400 "400MR"

Relax. delay 1.000 sec
Pulse 45.0 degrees
Acq. time 1.690 sec
Width 9689.9 Hz
256 repetitions
OBSERVE P31, 161.7808009 MHz
DECOUPLE H1, 399.6510008 MHz
Power 40 dB
continuously on
WALTZ-16 modulated
DATA PROCESSING
Line broadening 1.0 Hz
FT size 32768
Total time 11 min, 8 sec

237

^{1}H-NMR (400 MHz, CDCl$_3$): δ=1.06-1.15 (m, 24H), 2.12-2.44 (m, 4H), 2.51, 2.60 (2s, 6H), 3.51-3.68 (m, 10H), 3.69-3.72 (4s, 12H), 4.30, 4.42 (m, 2H), 4.70 (d, J=8.4, 1H), 4.79 (d, J=8.0, 1H), 5.96, 6.03 (2s, 2H), 6.74-7.44 (m, 46H), 7.90, 8.02 (2br, 2H), 8.24, 8.25 (2s, 2H).

238

^{31}P-NMR (161 MHz, CDCl$_3$): 150.99, 151.34.

150.987

151.344

SampleexBS_F_50
File: exp
Pulse Sequence: s2pul
Pulse Sequence: s2pul
Solvent: CDCl3
Temp: 24.0 C / 297.1 K
Operator: walkup
VNMRS-400 "MR400"

Relax. delay 1.000 sec
Pulse 45.0 degrees
Acq. time 1.600 sec
Width 3787.8 Hz
64 repetitions
OBSERVE P31, 161.7798332 MHz
DECOUPLE H1, 399.6484109 MHz
Power 40 dB
continuously on
WALTZ-16 modulated
DATA PROCESSING
Line broadening 1.0 Hz
FT size 131072
Total time 2 min, 46 sec

1H-NMR (400 MHz, CDCl₃): δ=1.88 (s, 3H), 4.48 (m, 2H), 5.60 (t, 1H), 5.67 (s, 1H), 6.21 (m, 1H), 7.27-8.09 (m, 11H), 8.25 (br, s, 1H).

Sample: 2S_D_d
File: exp

Pulse Sequence: s2pul
Solvent: cdcl3
Ambient temperature
Operator: walkup
VNMRS-400 "400MR"

Relax. delay 1.000 sec
Pulse 45.0 degrees
Acq. time 2.049 sec
Width 6410.3 Hz
16 repetitions
OBSERVE H1, 399.6479683 MHz
DATA PROCESSING
Line broadening 0.2 Hz
FT size 65536
Total time 0 min, 55 sec

3.00

2.07

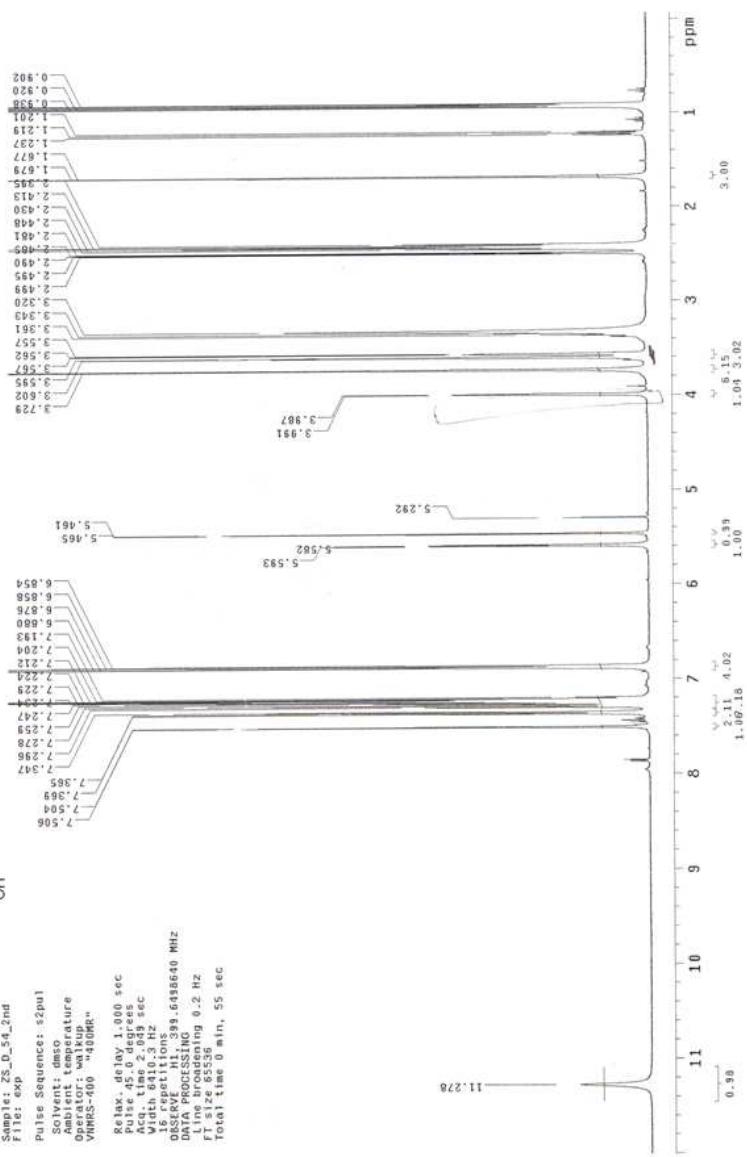0.96
0.97

0.92

1.00
3.84
1.88

1.84
1.85
0.93

¹H-NMR (400 MHz, DMSO-d₆): δ=1.74 (s, 3H), 3.32 (br, s, 2H), 4.00 (m, 4H), 5.64 (m, 1H), 7.56 (m, 1H), 11.27 (br, s, 1H).

Sample: 25_D_46
File: exp

Pulse Sequence: s2pul

Solvent: dmso
Ambient temperature
Operator: walkup
VNMRS-400 "400MR"

Relax. delay 1.000 sec
Pulse 45.0 degrees
Acq. time 2.049 sec
Width 6410.3 Hz
16 repetitions
OBSERVE H1, 399.6498653 MHz
DATA PROCESSING
Line broadening 0.2 Hz
FT size 65536
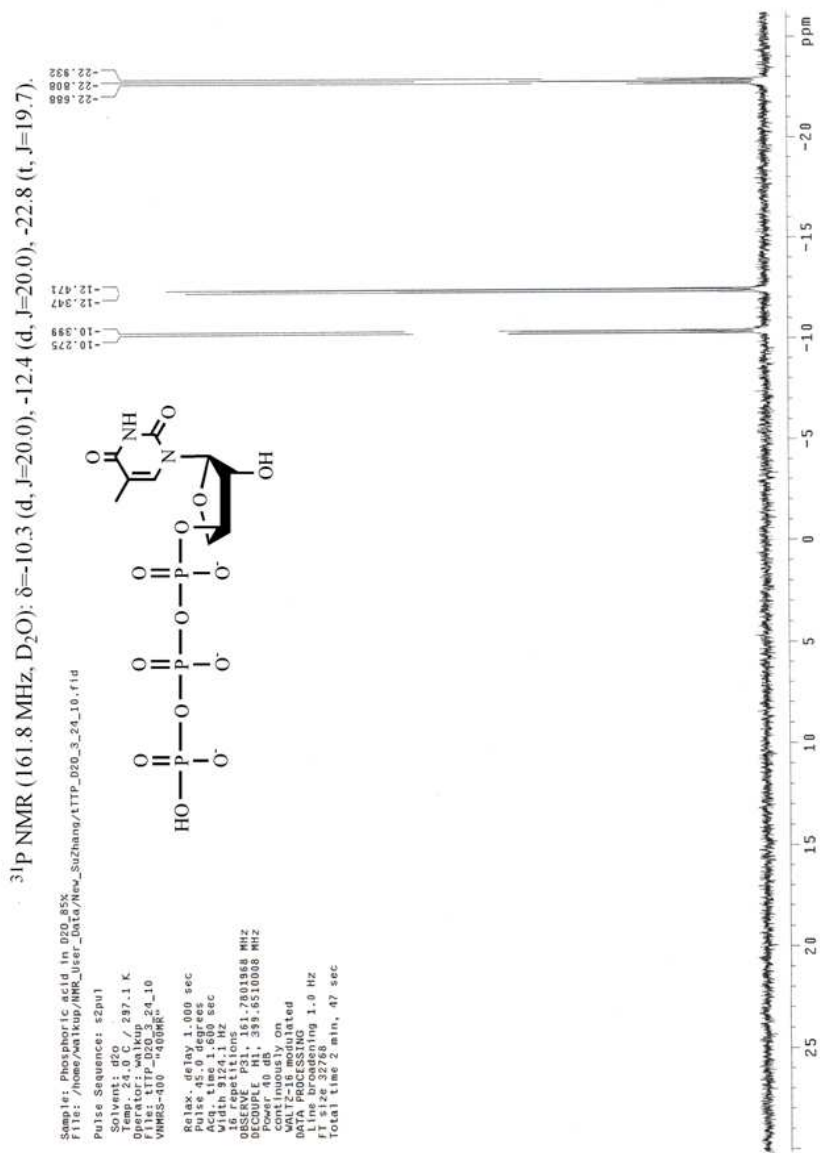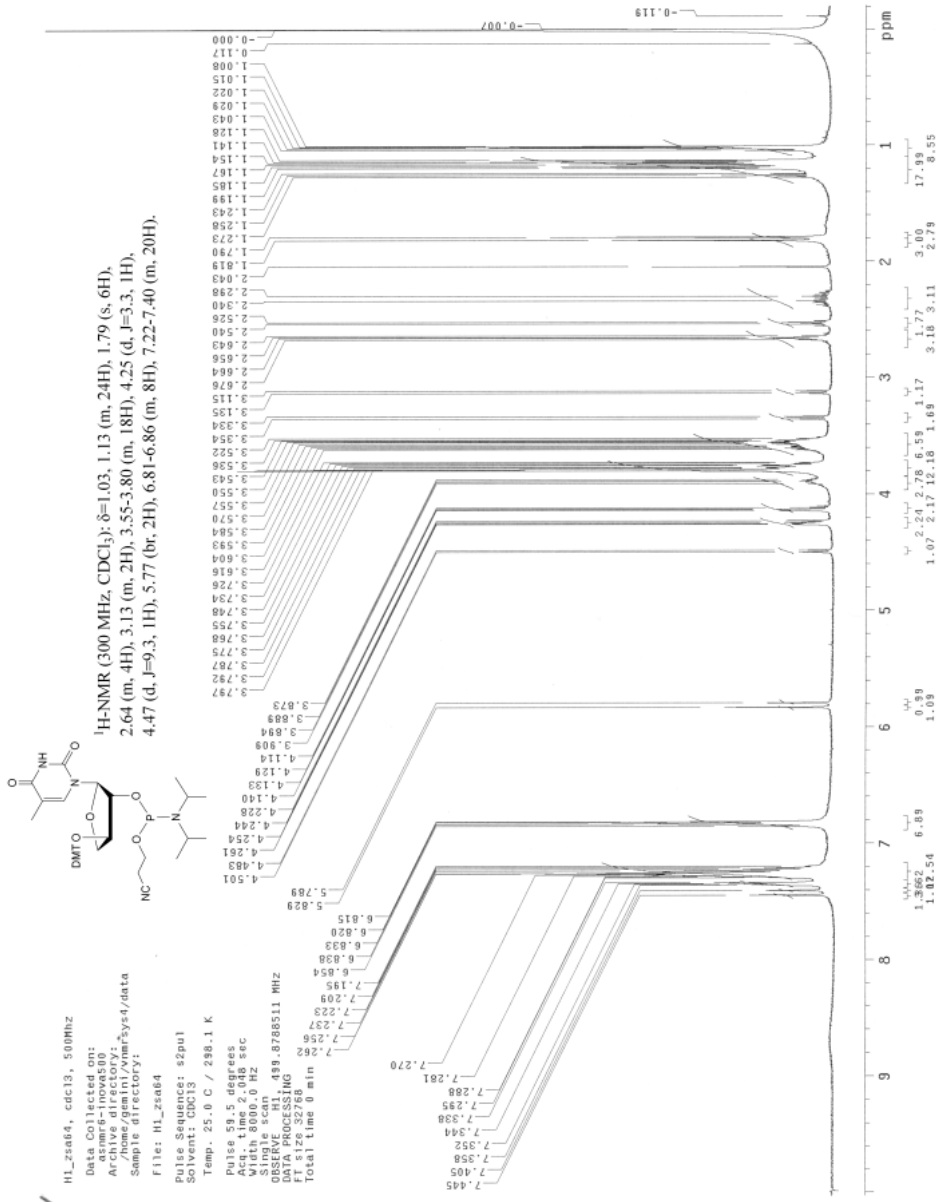Total time 0 min, 55 sec
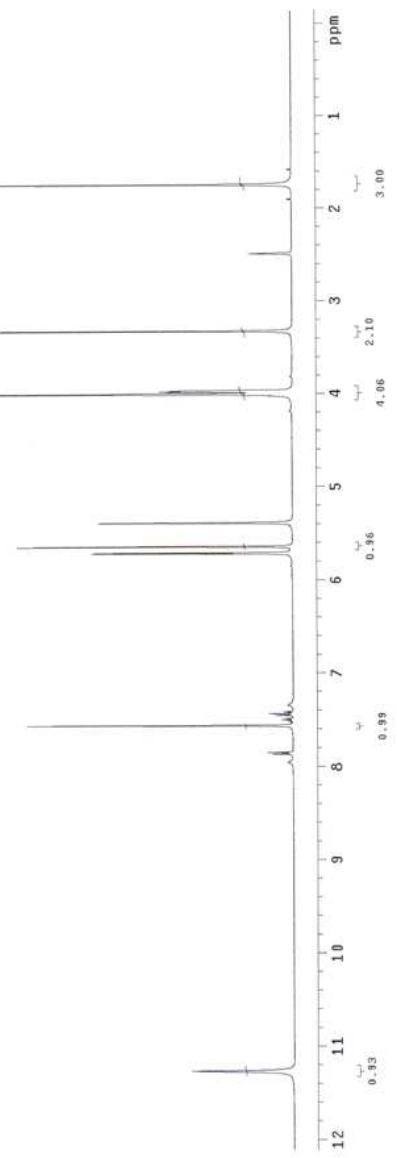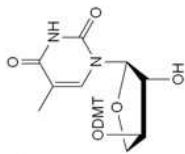
¹H-NMR (400 MHz, DMSO-d₆): δ=1.68 (s, 3H), 3.56-3.57 (m, 3H), 3.60 (s, 3H), 3.73 (s, 3H), 3.99 (d, J=1.6, 1H), 5.46 (d, J=1.6, 1H), 5.58 (d, J=4.4, 1H), 6.85-6.88 (m, 4H), 7.20-7.50 (m, 10H), 11.28 (s, 1H).

OH group seen

Sample: ZS_D_54_2nd
File: exp

Pulse Sequence: s2pul

Solvent: dmso
Ambient temperature
Operator: walkup
VNMRS-400 "400MR"

Relax. delay 1.000 sec
Pulse 45.0 degrees
Acq. time 2.049 sec
Width 6410.3 Hz
16 repetitions
OBSERVE H1, 399.6498640 MHz
DATA PROCESSING
Line broadening 0.2 Hz
FT size 65536
Total time 0 min, 55 sec

3'-O-DMT

^{1}H-NMR (400 MHz, DMSO-d$_6$): δ=1.66 (s, 3H), 3.32 (s, 3H), 3.34 (s, 3H), 3.71 (d, J=2.8, 1H), 3.82 (m, 2H), 3.92 (d, J=1.6, 1H), 5.10 (d, J=2.8, 1H), 5.91 (d, J=1.6, 1H), 6.80-7.20 (m, 4H), 7.21-7.39 (m, 10H), 11.33 (s, 1H).

^{31}P NMR (161.8 MHz, D$_2$O): δ=-10.3 (d, J=20.0), -12.4 (d, J=20.0), -22.8 (t, J=19.7).

^{1}H-NMR (300 MHz, CDCl$_3$); δ=1.03, 1.13 (m, 24H), 1.79 (s, 6H), 2.64 (m, 4H), 3.13 (m, 2H), 3.55-3.80 (m, 18H), 4.25 (d, J=3.3, 1H), 4.47 (d, J=9.3, 1H), 5.77 (br, 2H), 6.81-6.86 (m, 8H), 7.22-7.40 (m, 20H).

H1_zsa64, cdc13, 500Mhz

Data Collected on:
asnmrf6-inova500
Archive directory:
/home/gemini/vnmrsys4/data
Sample directory:

File: H1_zsa64

Pulse Sequence: s2pul
Solvent: CDCl3

Temp. 25.0 C / 298.1 K

Pulse 59.5 degrees
Acq. time 2.048 sec
Width 8000.0 Hz
Single scan
OBSERVE H1, 499.8788511 MHz
DATA PROCESSING
FT size 32768
Total time 0 min

^{31}P-NMR (202 MHz, CDCl$_3$): 153.4, 154.4.

153.409
154.418

APPENDIX B

CO-AUTHOR APPROVALS

VERIFICATION OF CO-AUTHOR APPROVALS

I verify that the following co-authors have approved of my use of our

publications in my dissertation.

John C. Chaput (Arizona State University)

Elizabeth O. McCullum (Arizona State University)

Ying-Wei Yang (Arizona State University)

Hanyang Yu (Arizona State University)

Note: The author's address is listed as when the research was performed.