UNIVERSITY OF CALIFORNIA,
IRVINE

Synthesis of α-L-Threose Monomers and Polymers for Therapeutic and Diagnostic Applications

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Pharmaceutical Sciences

by

Jen-Yu Liao

Dissertation Committee:
Professor John C. Chaput, Chair
Professor A. Richard Chamberlin
Professor Robert Spitale

2018

DEDICATION

This work is dedicated to my beloved parents Pensheng and Feili for their inspirational, limitless and selfless love and support they made me to be a passionate person.

TABLE OF CONTENTS

Page

Appendix

LIST OF FIGURES

Page

LIST OF TABLES

Page

LIST OF SCHEMES

Page

ACKNOWLEDGEMENTS

Firstly, I would like to express my deepest appreciation to my mentor professor John C Chaput for his advice, encouragement, life experience sharing and continuous support of my Ph.D. study in this collaborative environment. My sincere gratitude is extended to my committee mentors professor Richard Chamberlin, and professor Robert Spitale at University of California Irvine for their insightful comments for my research development, and guidance of my future plan.

I want to thank Dr. Sujay Sau and Dr. Hui Mei for their instructions in chemistry and educations in knowledge. My greatest thanks also goes to my lab mates Cailen McCloskey, Eric Yik and especially to Saikat Bala with many years of companionship in fun, frustrating, and celebration moments in my academic life. I want to thank Arlene, the lab manager, for her time and help in many aspect of my life especially my writing skill. I want to thank Dr. Yajun Wang, my badminton coach, for her intense training every Friday night. I owe significant debts to Dr. Kartik Tembhurnikar and Dr. Su (Richard) Zhang for his remote guidance and interpersonal connection sharing so that I can be prepared for all the upcoming challenges accordingly. I would like to thank all the other current and former Chaput group members for their help and time, it was really delightful to work in this big scientific family.

I owe great debts to Felix and Ben at UCI Mass facility, Dr. Dennison at UCI NMR facility and Dr. Ziller at X-ray Crystallography facility. Without their passionate help and knowledge sharing, I couldn't finish my research projects smoothly in this five years.

I'm really grateful for my master advisor professor Chung-Ming Sun who encouraged me to follow my enthusiasm in chemistry and pursue my graduate study at ASU and UCI. I also want to thank to my senior and professor Meng-Chiao Ho in Academia Sinica in Taiwan for his generosity and patience in taking me as a research assistant despite I didn't have any background in biology field.

I would like to thank to my family especially my dad, Pensheng, and my mom, Feili for their financial and spiritual support to my life. The companionship from them help me to overcome many hardships and frustration moments during my research career. I also want to thank to my grandmom, Fengbi and my grandpa Mingchih for their cultivation in my early ages. I still remembered many fun and interesting memories with them. I hope to have the chance to visit back soon to see them. Finally, I would like to thank to my girlfriend, Yanyang Tang. Whatever I achieved these few years is thanks to you.

Curriculum Vitae

Jen-Yu Liao

Fall 2018

Ph.D. candidate in Pharmaceutical Scieneces
University of California, Irvine
Irvine, CA, 92617

(480)300-9766
jenyul2@uci.edu
www.linkedin.com/in/uci-jenyuliao

Education

University of California-Irvine, Irvine, CA — 09/2015-12/2018
Arizona State University, Tempe, Arizona — 09/2013-08/2015
Ph.D. in Pharmaceutical Sciences

National Chiao-Tung University, Hsinchu, Taiwan — 07/2009-09/2011
Master in Applied Chemistry

National Tsing-Hua University, Taiwan — 09/2004-06/2007
Bachelor in Chemistry

Research Experience

Doctoral Research, University of California-Irvine, Irvine — 09/2015-12/2018

1. Developed the on-scale synthesis of L-threose nucleoside phosphoramidites. 2. Designed, synthesized the novel pyrophosphate reagent for natural/modified nucleoside triphosphates synthesis in gram scale by silica column chromatography. 3. Optimized the oligonucleotide (TNA, GNA, L-RNA, FANA) synthesis by solid phase and Involved the TNA oligonucleotides synthesis by chemical and enzymatic ligations. 4. Established a faithful DNA-TNA replication cycle and evolved a TNA aptamer toward HIV-RT target by SELEX in-vitro. 5. Demonstrated the intermolecular G-quarduplex structure of TNA oligonucleotides. 6. Synthesized, evaluated the stability of TNA polymer in human liver microsome, and snake venom phosphodiesterase.

Research Assistant, Academia Sinica, Taiwan — 08/2012-06/2013

Synthesized the nucleoside-amino acid conjugated molecules to study the mechanism of DNA/RNA methylation by protein arginine methyl transferase (PRMTs).

Master Research, National Chiao-Tung University, Taiwan — 07/2009-09/2011

Synthesized benzimidazole,-thiazole,-oxazole, and indole derivatives via copper catalysis and selenazole-fused isocumarine heterocycles via ruthenium C-H activation.

Patent

1. Chaput, J. C.; Liao, J.-Y.; Bala, S. Chemical reagents for natural/modified nucleoside triphosphate synthesis. Tech ID: 29428/UC Case 2018-245-0.

Awards

- 2018, Poster Award, XXIII Nucleosides, Nucleotides and Nucleic Acids (IRT), San Diego.
- 2018, Travel Grant, Department of Pharmaceutical Sciences, University of California-Irvine, Irvine, CA.
- 2013, Graduate Scholarship, School of Molecular Sciences, Arizona State University, Tempe, AZ.
- 2011, Dissertation Award in Dissertation from the National Chiao-Tung University, Hsinchu, Taiwan.

Publications (†equal contributions)

1. Liao, J.-Y.†; Bala, S.†; Chaput, J. C. An organopyrophosphate reagent for HPLC-free synthesis of natural and modified nucleoside triphosphate. In preparation.
2. McCloskey, C; Liao, J.-Y.; Bala, S.; Chaput, J. C. Ligase-Mediated TNA synthesis. In preparation.
3. Bala, S.†; Liao, J.-Y.†; Zhang, L.; Tran, C.; Chim, N. Chaput, J. C. Synthesis of 2'-Deoxy-α-L-Threofuranosyl Nucleoside Triphosphates. *J. Org. Chem.* 2018, 83, 8840. (†These authors contributed equally to the project)
4. Mei, H.†; Liao, J.-Y.†; Jimenez, R.; Wang, Y. J.; Bala, S.; McCloskey, C.; Switzer, C.; Chaput, J. C. Synthesis and Evolution of a Threose Nucleic Acid Aptamer Bearing 7-Deaza-7-Substituted Guanosine Residues. *J. Am. Chem. Soc*. 2018, 140, 5706. (†These authors contributed equally to the project)
5. Feldman, A. W.; Fischer, E. C.; Ledbetter, M. P.; Liao, J.-Y.; Chaput, J. C.; Romesberg, F. E. A Tool for the Import of Natural and Unnatural Nucleoside Triphosphates into Bacteria. *J. Am. Chem. Soc*. 2018, 140, 1447.
6. Dhole, S.; Liao, J.-Y.; Kumar, S.; Salunke, D. P.; Sun, C. M. Regioselective Synthesis of Angular Isocoumarinselenazoles: A Benzoselenazole-directed, Site-specific, Ruthenium-catalyzed C(sp2)-H Activation. *Adv. Synth. Catal*. 2018, 360, 1.
7. Bala, S.†; Liao, J.-Y.†; Mei, H.; Chaput, J. C. Synthesis of α-L-Threofuranosyl Nucleoside 3'-Monophosphates, 3'-Phosphoro(2-Methyl)imidazolides, and 3'-Triphosphates. *J. Org. Chem*, 2017, 82, 5910. (†These authors contributed equally to the project)
8. Liao, J.-Y.; Anosova, I.; Bala, S.; Van Horn, W. D.; Chaput, J. C. A parallel stranded G-quadruplex composed of threose nucleic acid (TNA). *Biopolymer*, 2017, 107:e22999.
9. Anosova, I.; Kowal, E. A.; Sisco, N. J.; Sau, S. P.; Liao, J.-Y.; Bala, S.; Rozners, F.; Egli, M.; Chaput, J. C.; Van Horn, W. D. Structural Insights into Conformation Differences between DNA/TNA and RNA/TNA Chimeric Duplexes. *ChemBioChem*, 2016, 17, 1.
10. Culbertson, M. C.; Temburnikar, K. W.; Sau, S. P.; Liao, J.-Y.; Bala, S.; Chaput, J. C. Evaluating TNA stability under simulated physiological conditions. *Bioorg. Med. Chem*., 2016, 26, 2418.
11. Sau, S. P.; Fahmi, N. E.; Liao, J.-Y.; Bala, S.; Chaput, J. C. A Scalable Synthesis of α-L-Threose Nucleic Acid Monomers. *J. Org. Chem*, 2016, 81, 2302.
12. Liao, J.-Y.; Selvaraju, M.; Sun, C. M. Multistep divergent synthesis of benzimidazole linked benzoxazole-benzothiazole via copper catalyzed domino annulation. *Organic and Biomolecular Chemistry*, 2013, 11, 2473.

Presentations & Workshops

1. XXIII International Roundtable of Nucleosides, Nucleotides and Nucleic Acids (IRT), University of California, San Diego, 2018
   Poster presentation "Synthesis and Evolution of a Threose Nucleic Acid Aptamer Bearing 7-Deaza-7-Substituted Guanosine Residues"
2. American Chemical Society National Meeting, San Francisco, 2017
   Poster presentation "A Parallel Stranded G-quadruplex Composed of Threose Nucleic Acid (TNA)"
3. Vertex Day, University of California, Irvine, 2016
   Oral presentation "Evaluating the Structural Properties of Threose Nucleic Acid"
4. Graduate Visitation Day, University of California, Irvine, 2016
   Poster presentation "Synthesis of α-L-Threofuranosyl Nucleoside 3'-Monophosphates, 3'-Phosphoro (2-Methyl) imidazolides, and 3'-Triphosphates"
5. Graduate Visitation Day, Arizona State University, 2014
   Poster presentation "Synthesis of Threose Nucleic Acid Monomer and Polymer to Evaluate the Function

References

John C. Chaput
Department of Pharmaceutical Sciences
University of California, Irvine
Tel: (949) 824-8149
Email: jchaput@uci.edu

Richard Chamberlin
Department of Chemistry
University of California, Irvine
Tel: (949) 824-7089
Email: richard.chamberlin@uci.edu

Robert Spitale
Department of Pharmaceutical Sciences
University of California, Irvine
Tel: (949) 824-7655
E-mail: rspitale@uci.edu

ABSTRACT OF THE DISSERTATION

Synthesis of α-L-Threose Monomers and Polymers for Therapeutic and Diagnostic Applications

By

Jen-Yu Liao

Doctor of Philosophy in Pharmaceutical Sciences

University of California, Irvine, 2018

Professor John Chaput, Chair

Current progress in polymerases engineering have enabled the delivery of heredity genetic information between the DNA and unnatural nucleic acid that are composed of threose nucleic acid (TNA). TNA (α-L-threofuranosyl-(3′,2′) nucleic acid) is an artificial genetic polymer with the α-L-threofuranosyl sugar that are connected by 2' to 3' phosphodiester linkage. Structure difference from DNA and RNA, TNA was further explored as a source of nuclease resistant affinity reagents in therapeutic application. Several examples of protein targeting experiments in-vitro using TNA sequences-folded structures have been investigated in our lab. Since TNA molecules are naturally unavailable, synthetic approach has been developed previously for initial research explorations. Further synthetic optimization have been investigated in this dissertation to provide quantitative amount of TNA substrates for the downstream research applications in molecular biology.

Chapter 1 describes a highly optimized synthetic approach for large scale preparation of α-L-threose nucleoside phosphoramidites by ten reaction steps from L-ascorbic acid. This new strategy overcomes several shortcomings from original method presented by

Eschenmoser and his colleagues and allow us to quantitatively synthesize several key intermediate precursors for the following α-L-threose nucleoside phosphoramidite and nucleoside triphosphate synthesis. The α-L-threose nucleoside phosphoramidites are then used as the monomers for the synthesis of TNA oligonucleotides by solid phase chemistry.

Chapter 2 discusses the stepwise approach development for synthesizing α-L-threose nucleoside 3'-monophosphate and 3'-triphosphate. 3'-Secondary hydroxyl group on α-L-threose, having steric encumbered issue from the neighboring nucleobases, presents the low efficiency in the phosphorylation step by using classic P(V) chemistry. Introducing P(III) phosphitylation, substitution, and subsequent oxidation reaction together achieve high phosphorylation yield on 3'-hydroxyl group of α-L-threose nucleoside and facilitate α-L-threose nucleoside 3'-monophosphate and 3'-triphosphate synthesis and purification by HPLC chromatography. The α-L-threose nucleoside 3'-triphosphates are tested as the substrates for enzymatic TNA oligonucleotide synthesis by engineered polymerases.

Chapter 3 discusses the systematic analysis of genetic information delivery in between the DNA and TNA in the TNA replication cycle (DNA->TNA->cDNA) using the engineered polymerases Kod-RI transcriptase and Bst reverse transcriptase. We observed nucleotide mis-incorporation when we introduced the mutagenic manganese (II) for improving the transcription efficiency for four-nucleotide (A, G, C, T) TNA library synthesis. Introducing the new TNA analogues (7-deazatGTPs) allows us to establish the faithful TNA replication cycle by avoiding the G-G miss match issue and create the structural diversity in TNA sequence library for the downstream application towards the biological targets by TNA SELEX *in-vitro*.

Chapter 4 examines the synthesis of 2'-deoxy-α-L-threose nucleoside using two deoxygenation approaches. Introducing the triethylborane as the radical initiator, that avoids the side products formation from the stringent radical initiation reaction condition required for AIBN, allows the deoxygenation reaction to process at room temperature smoothly and efficiently. The following phosphorylation reaction for 2'-deoxy-α-L-threose nucleoside triphosphate synthesis was achieved by stepwise approach described in Chapter 2. The 2'-deoxy-α-L-threose nucleoside triphosphates are designed and synthesized as the chain terminator incorporated by TNA Kod-RI polymerase for crystallography purpose.

Chapter 1

A Scalable Synthesis of α-L-Threose Nucleic Acid Monomers

Publication note

This paper was originally published in the *Journal of Organic Chemistry.*

Sujay P. Sau, Nour Eddine Fahmi, Jen-Yu. Liao, Saikat Bala, and John C. Chaput. A Scalable Synthesis of α-L-Threose Nucleic Acid Monomers. *J. Org. Chem*, 2016, 81, 2302-2307.

1.1 Contribution Statement

Fahmi, N. E. and Sau, S. P. designed the synthetic strategy. All authors contributed to the synthesis, reaction optimization, and data collection. Sau, S. P. and Chaput, J. C. wrote the manuscript with the comments from the rest of the authors.

1.2 Abstract

Recent advances in polymerase engineering have made it possible to copy genetic information back and forth between DNA and synthetic polymers composed of TNA (α-L-threofuranosyl-(3′,2′) nucleic acid). This property, coupled with enhanced nuclease stability related to natural DNA and RNA, encourage us to further investigate into the structural and functional properties of TNA as an artificial genetic polymer for synthetic biology. Here, we report a highly optimized chemical synthesis protocol for synthesizing multigram quantities of TNA nucleosides that can be readily converted to nucleoside 2′-phosphoramidites or 3′-triphosphates for solid-phase and polymerase-mediated synthesis, respectively. The synthetic protocol involves 10 chemical transformations with three crystallization steps and a single chromatographic purification, which results in an overall yield of 16–23%

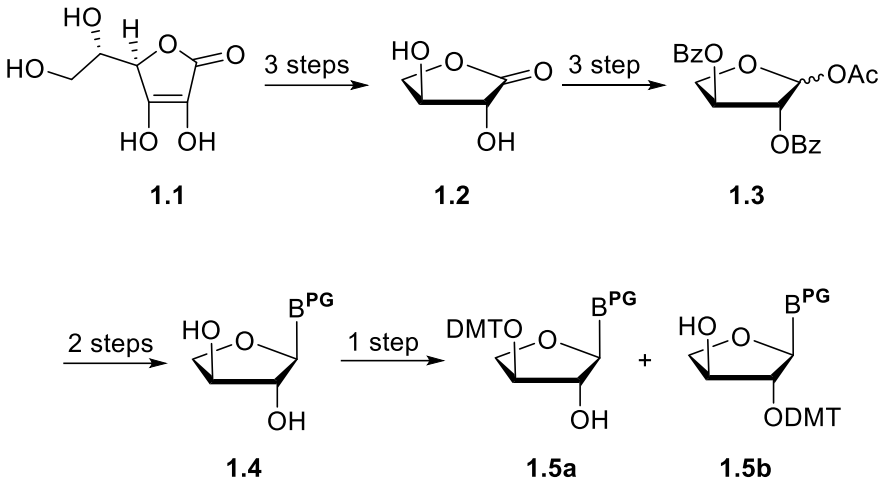
depending on the identity of the nucleoside (A, C, G, T).

1.3 Introduction

TNA (α-L-threofuranosyl-(3′,2′)-nucleic acid) is an artificial genetic polymer in which the natural ribose sugar found in RNA has been replaced by the α-L-threose sugar. In contrast to natural DNA and RNA, which have a six-atom backbone repeat unit connected by 3′, 5′-phosphodiester linkages, TNA has five atoms (or covalent bonds) in the phosphodiester backbone with linkages occurring at the 2′ and 3′ vicinal positions of the threose sugar. Despite this major difference in structure, TNA is capable of forming stable and efficient antiparallel Watson–Crick duplexes with complementary strands of DNA and RNA. [1–3] The ability to transfer genetic information between TNA and RNA, coupled with the chemical simplicity of threose relative to ribose sugar, has fueled interest in TNA as a possible progenitor of RNA in the pre-RNA world hypothesis.[4–6] TNA is also being explored as a catalysts for synthetic biology, source of nuclease resistant reagents, and molecular affinity medicine for biological targets, as the constitutional structure of TNA is stable under biological conditions. Ongoing works in both of these areas have inspired researchers to develop the engineered polymerases that can "transcribe" DNA into TNA and other polymerases that can "reverse transcribe" TNA back into DNA.[7–12] By including a selective amplification step in the replication cycle, methods are now being developed to isolate functional TNA molecules by in vitro selection.[13]

Motivated by a desire to study atomic-level structure by solution NMR and X-ray crystallography and the structural and functional properties of TNA by in vitro selection, we have systematically evaluated the chemical synthesis of TNA monomers with the goal of designing a synthetic approach that is amenable to large-scale (multigram) synthesis. A new

chemical synthesis strategy was necessary because the original approach (**Scheme 1-1**) first described by Eschenmoser and colleagues suffered from a number of shortcomings that



Scheme 1-1 Original Strategy Developed to Synthesize α-L-Threose Nucleosides

included numerous silica gel purification steps, low overall yield (2–6%), and poor regioselectivity during nucleoside tritylation. While the two DMT regioisomers (3′ and 2′) **1.5a** and **1.5b** can be separated chromatographically, only trace amounts of the 2′-isomer is obtained in the case of adenosine and cytosine, which necessitates an additional series of protection–deprotection steps to synthesize the adenine and cytosine threose nucleoside 3′-triphosphates. Here, we describe an optimal chemical synthesis strategy for generating TNA nucleoside precursors that can be readily converted to 3′-triphosphates for polymerase-mediated primer-extension reactions or 2'-phosphoramidites for solid-phase synthesis. Our approach involves a total of 10 chemical transformations with three crystallization and a single chromatographic purification steps and results in an overall yield of 16–23% depending on the identity of the nucleoside (A, C, G, T). We demonstrate that this new strategy can be used to produce multigram quantities of TNA monomers for the downstream exploration in structural and functional properties of TNA polymers.

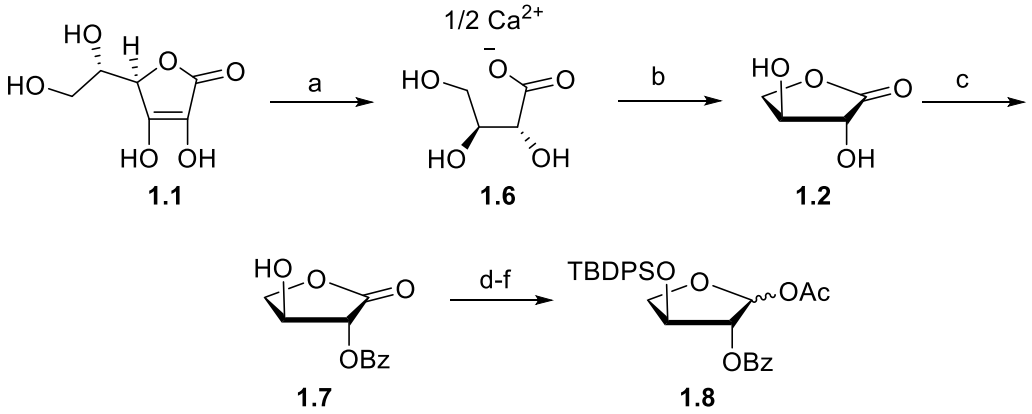
1.4 Results

The synthesis of L-threonolactone (**1.2**) from L-ascorbic acid (vitamin C) has been reported previously (**Scheme 1-1**). Accordingly, L-ascorbic acid undergoes oxidative cleavage in the presence of hydrogen peroxide and $CaCO_3$ to produce the calcium salt of L-threonate (**1.6**). Compound **1.6** is then converted to the desired L-threonolactone (**1.2**) using a Dowex cation exchange resin to promote acidification and intramolecular lactonization. However, due to the limited solubility of **1.6**, both steps require to remove the large volumes of water prior to cyclization and crystallization.

We began by examining the crystallization of **1.6** from the resulting aqueous reaction mixture. Literature methods report to reduce the volume of the aqueous solution by diminished-pressure evaporation prior to crystallization with methanol; however, this is a tedious process, especially when the reaction is performed on a large scale. We found that the addition of 2 volumetric equivalents of methanol directly to the reaction mixture allowed **1.6** to precipitate as a white crystalline material in 85% yield. This subtle change to the protocol resulted in a yield that favorably comparable againsted previous reports (65−79%) and allowed large-scale reactions to be performed with substantially higher throughput.

Next, we investigated the conversion of **1.6** to **1.2** (**Scheme 1-2**) with several feasible approaches. While previous literature methods apply the use of a Dowex cation acid. Of these, dilute sulfuric acid produced the desired compound in crystalline form with yields that are comparable to what has been achieved previously using Dowex. However, because of the hygroscopic nature of L-threonolactone (**1.2**), crystallization failed whenever the compound was not properly dried. As an alternative way to avoid using dilute H_2SO_4, we tested oxalic acid as a calcium exchange reagent in situ. We found that refluxing a heterogeneous mixture

of **1.6** in acetonitrile with 1 equiv of oxalic acid and a catalytic amount of para-toluenesulfonic acid produced the desired L-threonolactone (**1.2**) and calcium oxalate as an insoluble side product precipitate that could be easily separated by filtration. Using this approach, pure **1.2** is obtained as a white solid in 93% yield after coevaporation with ethyl acetate. Additionally, to eliminate the need for an expensive cation exchange resin, oxalic acid provided a streamlined method for obtaining large quantities of L-threonolactone (**1.2**) without the need for rigorous drying of aqueous solution to an anhydrous state.



Scheme 1-2. Synthesis of Universal Glycosyl Donor for TNA Nucleosides

Reagents and conditions: (a) (i) $CaCO_3$, 30% aq. H_2O_2, H_2O, 18 h, 0 °C to rt; (ii) active charcoal, 70 °C, 2 h, 85%; (b) oxalic acid, para-toluenesulfonic acid (cat), CH_3CN, 2 h, reflux, 93%; (c) benzoyl chloride, 1:10 pyridine–CH_2Cl_2, 0.5 h, 0 °C, 64%; (d) tert-butyldiphenylchlorosilane, imidazole, DMAP (cat), CH_2Cl_2, 18 h, 0 °C to rt; (e) 1 M DIBAL-H in toluene, dry 1,2-dimethoxyethane, 0.5 h, −78 °C; (f) Ac_2O–DMAP (5 equiv, 1.5 equiv) in CH_2Cl_2, 2 h, −78 °C to rt, 95% from **1.7**.

A major limitation of the original synthetic route developed by Eschenmoser and colleagues was the low regioselectivity observed when α-L-threofuranosyl nucleosides **1.4** are reacted with DMT-Cl (**Scheme 1-1**) in tritylation step. In most cases, tritylation produced a mixture of 2′ and 3′ regioisomers that could be separated by careful silica gel chromatography to generate pure **1.5a** and **1.5b**. In principle, this synthetic strategy would

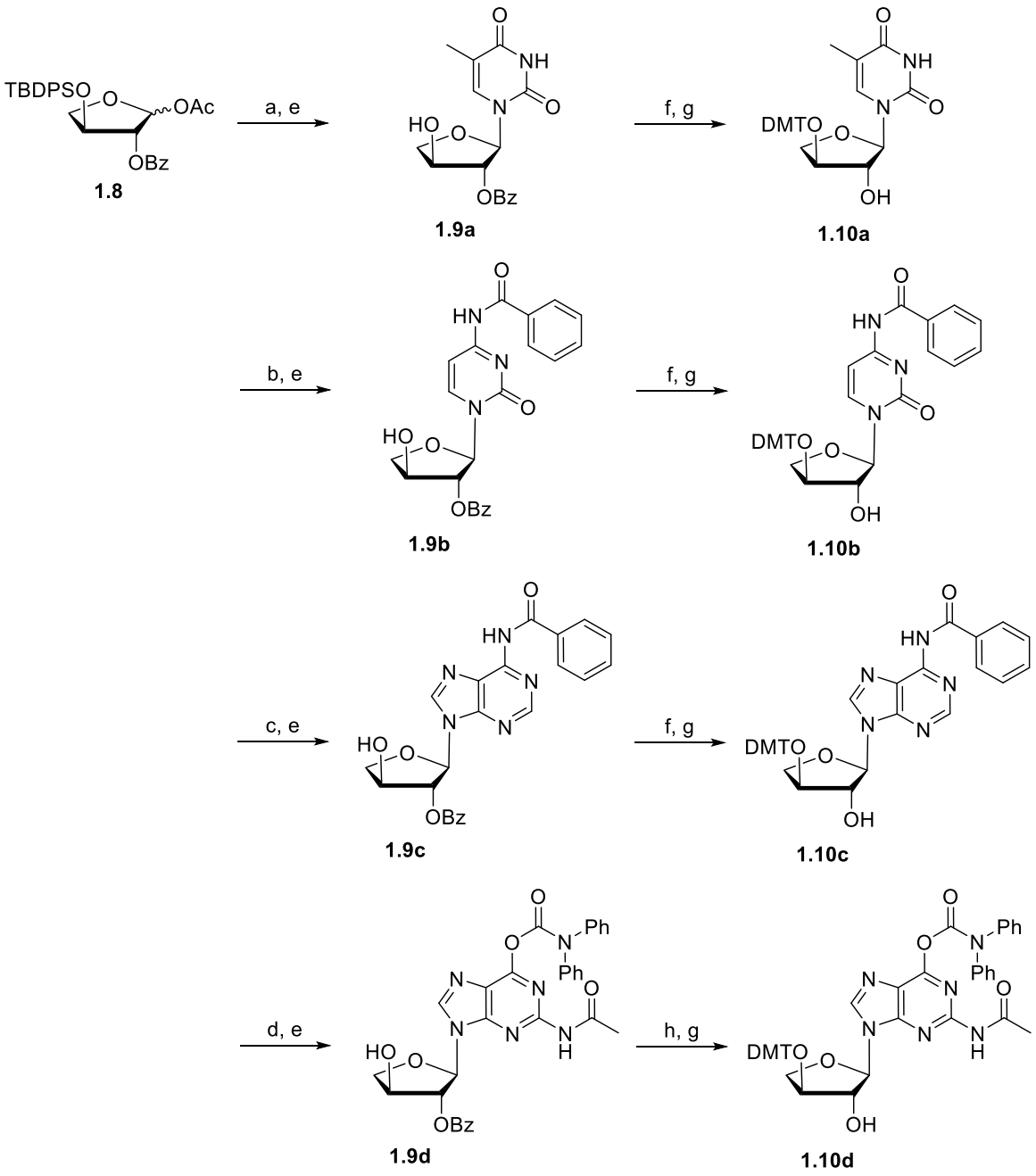
have been ideal if **1.5a** and **1.5b** were obtained in equal amounts; however, in practice, the tritylation of adenosine and cytosine α-L-threofuranosyl nucleosides yielded only trace amounts of the 2′-DMT isomer (**1.5b**). Thus, in order to obtain the precursor compounds required to synthesize TNA 3′-nucleoside triphosphates, a cumbersome strategy of protection and deprotection was developed to convert **1.5a** into **1.5b**.[14]

In an effort to generate antiviral compounds based on the structural framework of TNA nucleosides, Herdewijn and colleagues developed a regioselective strategy that could selectively convert **1.2** to the 2-*O*-benzoyl L-threonolactone derivative (**1.7**) by dropwise adding benzoyl chloride to a solution of **1.2** with imidazole in acetonitrile after 8 h of stirring from −5 °C to room temperature (**Scheme 1-2**). Although this method proved unsatisfactory in our hands, selective benzoylation of **1.2** was achieved with the addition of 1 equiv of benzoyl chloride to **1.2** in 1:10 pyridine−dichloromethane solution after stirring for 30 min at 0 °C. Following the reaction, we further discovered that **1.7** could be precipitated as a pure white powder in 64% yield after stirring overnight at 4 °C from a mixture of dichloromethane and hexanes.

In analogy to previous work on threofuranosyl nucleosides,[18] compound **1.7** was converted to the universal glycosyl donor **1.8** following three successive chemical transformations. We introduced the tert-butyldiphenylchlorosilyl protection at the 3'-hydroxy position of **1.7**, followed by DIBAL-H reduction of the lactone to the lactol, and finally acetylation at the anomeric position to give **1.8** in 96% yield from **1.7**. Filtration of **1.8** through a short-pad of silica gel proved highly facile and efficient due to the strong polarity differences between **1.8** and the nonwashable reagents. From **1.8**, we constructed a complete set of TNA nucleosides for thymidine (T, **1.9c**), cytidine (C, **1.9b**), adenosine (A,

1.9a), and guanosine (G, **1.9d**) by applying a Vorbrüggen glycosylation (**Scheme 1-3**) that involved heating the nucleobase (in protected form) with glycosyl donor in the presence of trimethylsilyl trifluoromethane-sulfonate (TMSOTf). After workup, the 3′-tert-butyldiphenylsilyl protecting group was selectively removed following the treatment with tetrabutylammonium fluoride for 1 h at 0 °C. Glycosylation and desilylation proceeded smoothly for thymine, N^4-benzoyl cytosine, and N^2-acetyl-O^6-diphenylcarbamoyl guanine to give **1.9a**, **1.9b**, and **1.9d**, respectively, in 64–70% yield after crystallization. However, glycosylation with N^6-benzoyl adenine resulted in a 3:2 mixture of N^9- and N^7-regioisomers. Surprisingly, similar isomeric mixtures were obtained when $SnCl_4$ was substituted for TMSOTf, which gave exclusively the N^9-regioisomer for glycosylation with 1,2,3-tribenzoyl threose.[15] Although the N^9- and N^7-regioisomers are separable by silica gel chromatography, we found that the crude isomeric mixture could be converted to the thermodynamically favored N^9-isomer by heating the reaction in anhydrous toluene in the presence of 1 equiv of TMSOTf for 1 h at 80 °C. Further optimization showed that glycosylation in toluene at 95 °C provided the desired N^9-regioisomer as the major synthesized product. Crystallization after 3′-desilylation gave pure N^9-adenosine nucleoside (**1.9c**) in 43.5% yield.

Compounds **1.9a–d** are key intermediates in the divergent synthesis of TNA nucleoside-3′-triphosphates and 2′-phosphoramidites. For L-threofuranosyl nucleoside 3′-triphosphates, compounds **1.9a–d** can be phosphorylated using the standard Ludwig and Eckstein method, followed by treating the concentrated NH_4OH to remove the sugar and nucleobase protecting groups.[14] For L-threofuranosyl nucleoside 2′-phosphoramidites, compounds **1.9a–d** are tritylated with DMT-Cl and treated with 1 N NaOH at 0 °C for 20 min deprotection reaction to remove the 2′-benzoyl group to obtain compounds **1.10a–d**



Scheme 1-3. Synthesis of TNA Phosphoramidite Precursors

Reagents and conditions: (a) thymine, *N,O*-bis(trimethylsilyl)acetamide (BSA), trimethylsilyltriflate (TMSOTf), CH_3CN, 2 h, 60°C; (b) N^4-benzoylcytosine, BSA, TMSOTf, CH_3CN, 2 h, 60 °C; (c) N^6-benzoyladenine, BSA, TMSOTf, toluene, 2.5 h, 95 °C; (d) (i) N^2-acetyl-O^6-diphenylcarbamoyl-guanine, BSA, 1,2-dichloroethane, 0.5 h 70 °C (ii) TMSOTf, CH_3CN, toluene, 1.5 h, 70 °C; (e) 1 M TBAF in THF, 1 h, 0 °C, 64% for **1.9a**, 70% for **1.9b**, 43.5% for **1.9c**, 69% for **1.9d**; (f) DMT-Cl, DMAP, pyridine, 18 h, 70 °C; (g) 1 M NaOH, THF–MeOH, 20 min, 0 °C, 71% for **1.10a**, 65% for **1.10b**, 63% for **1.10c**; (h) DMT-Cl, AgOTf, 2,4,6-trimethylpyridine, dry CH_2Cl_2 18 h, 70 °C, 52% for **1.10d**.

(**Scheme 1-3**).[17] While the tritylation of **1.9a–c** proceeded efficiently under standard conditions, tritylation of **1.9d** required rigorous azeotropic removal of water and the use of AgOTf as the catalyst. Removal of the *O*^{6}-diphenylcarbamoyl (DPC) group with 20% TFA in CH_2Cl_2 allowed tritylation to occur under standard conditions, suggesting that the bulky DPC group creates the steric barrier to limit the addition of the DMT group.

In our laboratory, we routinely synthesize TNA nucleosides **1.9a–d** using the above-mentioned methodology. In a typical synthesis run, 125 g of **1.1** is converted into 80 g of **1.7** with an overall yield of 49%. In contrast to previously methods, modifications to the purification protocol of **1.6** and **1.2** makes it possible to generate large quantities of highly pure **1.7** in just 5 days. The conversion of **1.7** to **1.8** is generally performed with 16.5 g of **1.7** to obtain 35 g of **1.8** with an overall yield of 96% in 24 h. The *N*-glycosylation of **1.8**, followed by desilylation, affords nucleosides **1.9a–d** in 64–70% yield, with the noted exception of **1.9c**, which is obtained in slightly lower yield (43.5%). We have found that the synthesis of **1.9a–d** from **1.7** can be performed on scales as large as 50 g, which is a substantial improvement in the synthesis of TNA nucleoside molecules that will help accelerate the synthesis and characterization of TNA polymers.

1.5 Conclusion

In summary, we provide a scalable and optimized synthesis protocol for constructing L-threofuranosyl nucleosides that are immediate precursors for TNA triphosphates and phosphoramidite monomers synthesis. Several key challenges have been resolved and numbers of purification steps were minimized to increase the yield of key intermediates and product synthesis. The conversion of vitamin C to L-threonolactone was optimized to minimize the cost and for generating a key intermediate in the synthesis pathway. The

universal glycosyl donor **1.8** was prepared from **1.7** in three chemical steps, followed by filtration through a short pad of silica gel. The complete set of TNA nucleosides (**1.9a–d**) were synthesized in two reactions from **1.8** and purified by crystallization. Final tritylation, followed by debenzoylation, provided the desired compounds (**1.10a–d**) after silica gel column chromatography. Together, these changes make it possible to synthesize multigram-scale quantities of TNA monomers that can be used to synthesize TNA polymers.

1.6 Experimental Details

General Procedures

Except as otherwise noted, all nonaqueous reactions were carried out in oven-dried glassware under a balloon pressure of argon or nitrogen. Reagents were commercially available and used as received; anhydrous solvents were purchased as the highest grade. Reactions were monitored by thin-layer chromatography using 0.25 mm Silicycle or EM silica gel 60 F254 plates. Column chromatography was performed using Silicycle 40–60 mesh silica gel. Yields are reported as isolated yields of spectroscopically pure compounds. ^{1}H and ^{13}C NMR spectra were obtained using 400 and 500 MHz NMR spectrometers. Chemical shifts are reported in parts per million (ppm, δ) referenced to the ^{1}H resonance of TMS. ^{13}C spectra are referenced to the residual ^{13}C resonance of the solvent (DMSO-d_6, 39.52 ppm, CD_3OD, 49.00 ppm). Splitting patterns are designated as follows: s, singlet; br, broad; d, doublet; dd, doublet of doublets; t, triplet; q, quartet; m, multiplet.

Calcium-L-threonate (1.6).[15]

To a cold (0–5 °C) solution containing 125 g (0.71 mol) of L-ascorbic acid (**1.1**) in 1 L of H_2O was slowly added 125 g (1.25 mol) of $CaCO_3$ with a spatula over 30 min (evolution of

CO_2 was observed). To the resulting heterogeneous slurry was added 250 mL of 30% aq. H_2O_2 dropwise over a period of ca. 1 h with stirring at 0–5 °C. The reaction mixture was allowed to warm to r.t and stirred overnight. The heterogeneous slurry was filtered, and the filter cake was washed with two-100 mL portions of H_2O. The filtrate was treated with 25 g of activated carbon Darco G-60 and then heated to 50 °C, until peroxide was no longer detected using Merckoquant 1001-1 peroxide test strips. The hot suspension was filtered, and the solid material was washed with two-50 mL portions of H_2O. The washings and the filtrate were combined and crystallized by the addition of 2 volume equivalents of methanol while stirring for 16 h at 4 °C. The solid material was filtered, washed with two-50 mL portions of MeOH, and dried under high vacuum at 40 °C. Calcium L-threonate monohydrate (**1.6**) was obtained as a white solid: yield 99.4 g (85%); ^{1}H NMR (400 MHz, D_2O) δ 3.60 (dd, 1H, *J* = 11.6, 7.6 Hz), 3.66 (dd, 1H, *J* = 11.6, 5.2 Hz), 3.95 (m, 1H) and 4.03 (m, 1H).

L-Threonolactone (1.2).[19]

To a suspension of calcium L-threonate (**1.6**) (99.4 g, 0.30 mol) in dry acetonitrile (500 mL) were added anhydrous oxalic acid (28.8 g, 0.32 mol) and para-toluenesulfonic acid monohydrate (1.0 g). The heterogeneous mixture was stirred at reflux for 3 h. The hot mixture was allowed to cool to room temperature and filtered. The filter cake was washed with 50 mL of acetonitrile, and the combined filtrate was evaporated under reduced pressure to produce a colorless syrup. The residue was dissolved in 100 mL of EtOAc and evaporated to dryness to give L-threonolactone (**1.2**) as a white solid: yield 66.5 g (93%); silica gel TLC (EtOAc): Rf = 0.36; ^{1}H NMR (400 MHz, CD_3OD) δ 3.93 (dd, 1H, *J* = 8.8, 7.2 Hz), 4.20 (d, 1H, *J* = 7.2 Hz), 4.29 (dd, 1H, *J* = 14.0, 6.8 Hz) and 4.41 (dd, 1H, *J* = 9.8, 6.8 Hz), 4.82 (s, 2H).

2-*O*-Benzoyl-L-threonolactone (1.7).[18]

To a cold (0−5 °C) solution containing 66.5 g (0.56 mol) of L-threonolactone (**1.2**) in 1.2 L of CH_2Cl_2 and 135 mL of anhydrous pyridine under argon was added benzoyl chloride (72.0 mL, 0.62 mol) dropwise. The mixture was stirred under argon for 30 min at 4 °C. The crude reaction mixture was then sequentially washed with 1 N HCl (3 × 400 mL) and brine (200 mL). The combined aqueous layer was back extracted with two 100 mL portions of CH_2Cl_2. To the combined organic layers was added 2 volume equivalents of hexane over 1 h, and the solution was left stirring for 16 h at 4 °C. The product was collected by filtration and dried under vacuum. 2-*O*-benzoy-L-threonolactone (**1.7**) was obtained as a white solid: yield 80.1 g (64%); silica gel TLC (1:1 hexanes−EtOAc): Rf = 0.30; ^{1}H NMR (400 MHz, $CDCl_3$) δ 3.81 (brs, 1H), 4.16 (dd, 1H, *J* = 9.2, 7.6 Hz), 4.64 (dd, 1H, *J* = 9.2, 8.0 Hz), 4.73 (q, 1H, *J* = 7.6 Hz), 5.41 (d, 1H, *J* = 6.8 Hz), 7.48 (t, 2H, *J* = 8.0 Hz), 7.64 (t, 1H, *J* = 8.0 Hz) and 8.10 (d, 2H, *J* = 7.6 Hz).

1-*O*-Acetyl-2-*O*-benzoyl-3-*O*-tert-butyldiphenylsilyl-L-threofuranose (1.8).[18]

To a cold (0−5 °C) solution containing 16.5 g (74.3 mmol) of 2-*O*-benzoyl-L-threonolactone (**1.7**), 60 mg of DMAP and 10.2 g (150 mmol) of imidazole in 160 mL of CH_2Cl_2 was added dropwise tert-butyldiphenylchlorosilane (20 mL, 75.0 mmol). The reaction mixture was warmed to r.t. and stirred under argon for 16 h. The solvent was evaporated under reduced pressure, and the residue was dissolved in 300 mL of hexane. The organic phase was sequentially washed with 1 N HCl (100 mL), H_2O (100 mL), and brine (100 mL). The organic layer was dried over $MgSO_4$ and evaporated to give 2-*O*-benzoyl-3-*O*-tert-butyldiphenylsilyl-L-threonolactone as a colorless syrup (35 g). The crude material was used directly without further purification; silica gel TLC (3:1 hexanes−EtOAc) Rf = 0.5.

To a cold (−78 °C, acetone/dry ice) solution containing 35.0 g of crude 2-*O*-benzoyl-3-*O*-tert-butyldiphenylsilyl-L-threonolactone in 150 mL of 1,2-dimethoxyethane was added a 1 M solution of DIBAL-H in toluene (100 mL, 100 mmol) dropwise over a 10 min period. The mixture was stirred under argon for an additional 20 min at −78 °C. TLC showed the reaction to be complete (Rf 0.35 in 3:1 hexanes−EtOAc). A premade solution containing Ac_2O (367 mmol) and DMAP (115 mmol) in CH_2Cl_2 (Ac_2O/DMAP/CH_2Cl_2 = 35 mL/14.0 g/40 mL) was then added dropwise at −78 °C. After 10 min, the reaction mixture was removed from the dry ice bath and allowed to stir for 1.5 h while warming to r.t. The mixture was then diluted with 200 mL of hexanes and poured into a cold stirring 1 N aq. HCl solution (200 mL). The organic layer was separated and washed with H_2O (100 mL), sat. aq. $NaHCO_3$ (100 mL), and brine (100 mL). The organic layer was dried over $MgSO_4$ and evaporated to give 42 g of crude product as a yellow syrup. The product was suspended in 25 mL of 10% EtOAc in hexanes, passed through a short pad of silica (50 g), and washed out with 10% EtOAc in hexanes (200 mL). An anomeric mixture of 1-*O*-acetyl-2-*O*-benzoyl-3-*O*-tert-butyldiphenylsilyl-L-threo-furanose (**1.8**) was obtained as a colorless syrup: yield 36.0 g (96%); silica gel TLC (3:1 hexanes−EtOAc): Rf = 0.5.

1-(2′-*O*-Benzoyl-α-L-threofuranosyl)thymine (1.9a).

To a solution containing 34.0 g (67.4 mmol) of 1-*O*-acetyl-2-*O*-benzoyl-3-*O*-tert-butyldiphenylsilyl-L-threofuranose (**1.8**) and 8.9 g (70.5 mmol) of thymine in 140 mL of anhydrous acetonitrile was added 45 mL (145 mmol) of *N,O*-bis(trimethylsilyl)acetamide, and the mixture was stirred for 30 min at 60 °C. TMSOTf (20.0 mL, 108 mmol) was added dropwise, and stirring was continued for another 2 h at 60 °C, after which time TLC analysis (1:1 hexanes−EtOAc) showed the reaction to be complete. The mixture was cooled to r.t,

diluted with 300 mL of EtOAc, and poured into 200 mL of cold sat. aq. $NaHCO_3$ solution with stirring. The organic layer was separated and washed with H_2O (100 mL) and brine (100 mL), dried over $MgSO_4$, and concentrated under reduced pressure. The crude nucleoside was obtained as a white foam (41 g) and was used directly without further purification; silica gel TLC (1:1 hexanes–EtOAc): Rf = 0.37.

To a cold (0–5 °C) solution containing 41 g of crude 1-(2′-*O*-benzoyl-3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)thymine in THF (250 mL) was added dropwise tetrabutylammonium fluoride (1 M solution in THF, 70.0 mL, 70.0 mmol), and the mixture was stirred for 1 h at 0–5 °C. The solvent was evaporated under reduced pressure, and the residue was dissolved in 300 mL of EtOAc. The organic layer was separated and washed twice with H_2O (100 mL) and brine (100 mL), dried over $MgSO_4$, and concentrated under reduced pressure to give 20 g of the crude nucleoside as a yellow syrup. The syrup was dissolved in 50 mL of CH_2Cl_2 and crystallized by the addition of 50 mL of hexanes. 1-(2′-*O*-Benzoyl-α-L-threofuranosyl)-thymine (**1.9a**) was obtained as a white solid: yield 14.44 g (64%); silica gel TLC (1:2 hexanes–EtOAc): Rf = 0.35; ^{1}H NMR (400 MHz, $CDCl_3$) δ 1.90 (s, 3H), 4.05 (brs, 1H), 4.20 (dd, 1H, *J* = 4.4, 10.0 Hz), 4.29 (d, 1H, *J* = 10.0 Hz), 4.51 (brs, 1H), 5.43 (s, 1H), 5.87 (d, 1H, *J* = 2.0 Hz), 7.36 (s, 1H), 7.47 (t, 2H, *J* = 7.6 Hz), 7.61 (t, 1H, *J* = 7.2 Hz), 8.02 (d, 2H, *J* = 8.2 Hz) and 8.90 (brs, 1H); ^{13}C NMR (125 MHz, CD_3OD): δ 12.5, 74.7, 76.5, 83.9, 91.4, 111.2, 129.7, 130.4, 130.8, 134.8, 138.7, 152.3, 166.4, 166.6; HRMS (ESI-TOF): [M + Na]+ calcd for $C_{16}H_{16}N_2O_6Na$, 355.0906; found, 355.0910.

N^4-Benzoyl-1-(2′-*O*-benzoyl-α-L-threofuranosyl)cytosine (1.9b).

To a solution containing 33.5 g (66.3 mmol) of 1-*O*-acetyl-2-*O*-benzoyl-3-*O*-tert-butyldiphenylsilyl-L-threofuranose (**1.8**) and 15.0 g (70.0 mmol) of N^4-benzoylcytosine in

150 mL of anhydrous acetonitrile was added 40.0 mL (154 mmol) of *N,O*-bis(trimethylsilyl)acetamide, and the mixture was stirred for 30 min at 60 °C. TMSOTf (38.0 mL, 199 mmol) was added dropwise, and stirring was continued at 60 °C for another 2 h, after which time TLC (1:1 hexanes–EtOAc) showed the reaction to be complete. The mixture was cooled to room temperature, diluted with 200 mL of EtOAc, and poured into 200 mL of sat. aq. $NaHCO_3$ solution with stirring. The white suspension was filtered over Celite, and the organic layer was separated and washed with H_2O (150 mL) and brine (150 mL), dried over $MgSO_4$, and concentrated under reduced pressure. The crude nucleoside was obtained as a white foam (36.0 g) and was used directly without further purification; silica gel TLC (EtOAc): Rf = 0.75.

To a cold (0–5 °C) solution containing 36.0 g of crude N^4-benzoyl-1-(2′-*O*-benzoyl-3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)-cytosine in THF (200 mL) was added dropwise tetrabutylammonium fluoride (1 M solution in THF, 70.0 mL, 70.0 mmol), and the mixture was stirred at 0–5 °C for 1 h. The solvent was evaporated under diminished pressure, and the residue was dissolved in 500 mL of EtOAc. The organic layer was washed with 1 N aq. HCl (150 mL) and H_2O (150 mL), and then was stirred with brine (150 mL) at r.t. to precipitate the product. N^4-Benzoyl-1-(2′-*O*-benzoyl-α-L-threofuranosyl)cytosine (**1.9b**) was obtained as a white solid: yield 19.4 g (70%); silica gel TLC (1:2 hexanes–EtOAc): Rf = 0.40; ^{1}H NMR (400 MHz, DMSO-d_6) δ 4.3 (m, 3H), 5.39 (s, 1H), 5.85 (d, 1H, J = 2.8 Hz), 5.99 (s, 1H), 7.41 (d, 1H, J = 7.2 Hz), 7.50–7.64 (m, 5H), 7.72 (t, 1H, J = 7.2 Hz), 8.03 (t, 4H, J = 7.6 Hz), 8.22 (d, 1H, J = 7.6 Hz), and 11.27 (s, 1H); ^{13}C NMR (125 MHz, DMSO-d_6): δ 72.4, 76.6, 81.5, 90.9, 95.7, 128.5, 128.9, 129.5, 132.8, 133.2, 133.9, 145.7, 154.5, 163.4, 164.5, 167.4; HRMS (ESI-TOF): [M + Na]+ calcd for $C_{22}H_{19}N_3O_6Na$, 444.1172; found, 444.1180.

N^6-Benzoyl-9-(2′-*O*-benzoyl-α-L-threofuranosyl)adenine (1.9c).[18]

A mixture of 40.5 g (80.0 mmol) of 1-*O*-acetyl-2-*O*-benzoyl-3-*O*-tert-butyldiphenylsilyl-L-threofuranose (**1.8**) and 19.1 g (80.0 mmol) of N^6-benzoyladenine was coevaporated with 50 mL of anhydrous toluene and then suspended in 160 mL of anhydrous toluene. To the suspension was added 42.0 mL (163.8 mmol) of *N,O*-bis(trimethylsilyl)acetamide, and the mixture was stirred for 30 min at 95 °C. Once all the suspension was dissolved, TMSOTf (22.2 mL, 118.8 mmol) was added dropwise, and stirring was continued at 95 °C for another 2.5 h, after which TLC showed complete consumption of the starting material and more polar major products corresponding to the N^9 glycosides. The mixture was cooled to room temperature, then poured into a stirring mixture of 150 g of ice and 150 mL of sat. aq. $NaHCO_3$ solution and stirred for 30 min. The organic layer was separated and washed with H_2O (50 mL) and brine (50 mL), dried over $MgSO_4$, and concentrated under reduced pressure. The crude nucleoside was obtained as a yellow foam: yield 54.3 g.

To a cold (0−5 °C) solution (ice-bath) containing 54.3 g of N^6-benzoyl-9-(2′-*O*-benzoyl-3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)adenine in THF (160 mL) was added dropwise tetrabutylammonium fluoride (1 M solution in THF, 80.0 mL, 80.0 mmol), and the mixture was stirred at 0−5 °C for 30 min. The solvent was evaporated under diminished pressure, and the residue was dissolved in 200 mL of EtOAc. The organic layer was washed with H_2O (50 mL) and brine (50 mL). The combined aqueous layer was back-extracted with EtOAc (2 × 50 mL). The combined organic layer was dried over $MgSO_4$ and concentrated under reduced pressure. The resulting crude material was dissolved in 150 mL of 1% methanol in CH_2Cl_2 and crystallized by slow addition of 120 mL of hexanes. N^6-

Benzoyl-9-(2′-*O*-benzoyl-α-L-threofuranosyl)adenine (**1.9c**) was obtained as a white solid: yield 15.5 g (43.5%); silica gel TLC (EtOAc): Rf = 0.35; ^{1}H NMR (400 MHz, $CDCl_3$) δ 4.30 (dd, 1H, *J* = 4.0, 10.0 Hz), 4.39 (d, 1H, *J* = 10.0 Hz), 4.67 (d, 1H, *J* = 3.2 Hz), 5.63 (s, 1H), 6.07 (d, 1H, *J* = 1.6 Hz), 7.56 (m, 6H), 8.04 (t, 4H, *J* = 8.1 Hz), 8.27 (s, 1H), and 8.81 (s, 1H).

N^2-Acetyl-O^6-diphenylcarbamoyl-9-(2′-*O*-benzoyl-α-L-threo-furanosyl)guanine 1.9d).

To a solution containing 8.45 g (26.1 mmol) of N^2-acetyl-O^6-diphenylcarbamoyl guanine in 200 mL of a mixture of anhydrous dichloroethane and toluene (1:2 v/v) was added 15.8 mL (64.6 mmol) of *N,O*-bis(trimethylsilyl)acetamide, and the mixture was stirred for 30 min at 70 °C. The solvent was removed under reduced pressure, and the residue was dissolved in 55 mL of anhydrous toluene. 1-*O*-Acetyl-2-*O*-benzoyl-3-*O*-tert-butyldiphenyl-silyl-L-threofuranose (**1.8**) (11 g, 21.8 mmol) in 65 mL of anhydrous toluene was added dropwise using a canula, and the mixture was heated to 70 °C. TMSOTf (8.5 mL, 45.9 mmol) was then added dropwise, and the mixture was stirred at 70 °C for another 1.5 h, after which TLC (1:1 hexanes−EtOAc) showed complete consumption of (**1.8**). The mixture was cooled to room temperature, diluted with 300 mL of EtOAc, and poured into 150 mL of sat. aq. $NaHCO_3$ solution with stirring, resulting in a purple suspension. The suspension was filtered, and the organic layer was separated and washed with H_2O (150 mL) and brine (150 mL), dried over $MgSO_4$, and concentrated under reduced pressure. The crude nucleoside was obtained as a purple foam (16.12 g) and was used directly without further purification; silica gel TLC (1:2 hexanes−EtOAc): Rf = 0.7.

To a cold (0–5 °C) solution containing 16.12 g of crude N^2-acetyl-O^6-diphenylcarbamoyl-9-(2′-*O*-benzoyl-3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl) guanine in THF (50 mL) was added dropwise tetrabutylammonium fluoride (1 M solution in

THF, 19.0 mL, 19.0 mmol), and the mixture was stirred at 0−5 °C. After 15 min, TLC was checked and another 9 mL of TBAF was added, and the mixture was stirred for 45 min. The solvent was evaporated under diminished pressure, and the residue was dissolved in 250 mL of EtOAc. The organic layer washed with 1 N aq. HCl (50 mL), H_2O (50 mL), and brine (50 mL), was dried over $MgSO_4$, and concentrated under reduced pressure. The resulting crude product was dissolved in CH_2Cl_2 (30 mL) and crystallized by addition of an equal amount of hexanes. *N*^{2}-Acetyl-*O*^{6}-diphenylcarbamoyl-9-(2′-*O*-benzoyl-α-L-threofuranosyl)-guanine (**1.9d**) was obtained as a white solid: yield 8.9 g (69%); silica gel TLC (1:2 hexanes−EtOAc): Rf = 0.25; ^{1}H NMR (400 MHz, DMSO-*d*^{6}) δ 2.20 (s, 3H), 4.32 (dq, 2H, *J* = 3.6, 7.6 Hz), 4.60 (m, 1H), 5.78 (t, 1H, *J* = 2.0 Hz), 5.96 (d, 1H, *J* = 4.0 Hz), 6.31 (d, 1H, *J* = 1.6 Hz), 7.33 (t, 2H, *J* = 6.0 Hz), 7.43−7.60 (m, 10H), 7.72 (t, 1H, *J* = 6.0 Hz), 8.03 (dd, 1H, *J* = 0.8, 6.4 Hz), 8.60 (s, 1H), and 10.81 (s, 1H); ^{13}C NMR (125 MHz, DMSO-*d*$_6$) δ 24.6, 73.2, 74.8, 82.3, 87.9, 120.1, 128.7, 128.9, 129.5, 129.6, 134.0, 141.6, 144.3, 150.1, 152.3, 154.2, 155.3, 165.0, 169.2; HRMS (ESI-TOF): [M + Na]+ calcd for $C_{31}H_{26}N_6O_7Na$, 617.1761; found, 617.1750.

1-(3′-O-Dimethoxytrityl-α-L-threofuranosyl)thymine (1.10a).[15]

A mixture containing 3.2 g (9.6 mmol) of 1-(2′-*O*-benzoyl-α-L-threofuranosyl)thymine (**1.9a**), 4.88 g (14.4 mmol, 1.5 equiv) of DMT-Cl, and 100 mg of DMAP (0.82 mmol) was coevaporated twice with 20.0 mL of anhydrous pyridine. The mixture was dissolved in 40 mL of anhydrous pyridine and stirred under argon at 80 °C for 16 h. The solvent was evaporated under diminished pressure, and the residue was partitioned between EtOAc (50 mL) and H_2O (50 mL). The organic layer was washed with brine (50 mL), dried ($MgSO_4$), and evaporated. The residue was dissolved in a mixture of 38 mL of THF and 30.5 mL of MeOH, and then cooled to 0 °C. To the cold solution was added

15.0 mL of ice-cold 1 N aq. NaOH while stirring. After 30 min, the reaction was quenched by addition of 100 mL of water. The volatile solvent was removed under diminished pressure, and the residual aqueous solution was extracted with EtOAc (3 × 50 mL). The combined organic layer was dried ($MgSO_4$) and evaporated. The residue was purified on a silica gel column (50 g silica bed), eluting with 50–100% EtOAc in hexanes containing 2% Et_3N in steps of 5% increase in EtOAc for every 200 mL. 1-(3′-*O*-Dimethoxytrityl-α-L-threofuranosyl)thymine (**1.10a**) was obtained as a white foam: yield 3.6 g (71%); ^1H NMR (400 MHz, $CDCl_3$): δ 1.82 (s, 3H), 3.30 (d, 1H, *J* = 9.6 Hz), 3.67 (s, 3H), 3.71 (s, 3H), 3.91 (d, 1H, *J* = 13.6 Hz), 4.19 (s, 1H), 4.23 (s, 1H), 5.72 (s, 1H), 6.79 (m, 4H), 7.12–7.38 (m, 9H), 7.44 (s, 1H) and 10.92 (s, 1H).

N^4-Benzoyl-1-(3′-*O*-dimethoxytrityl-α-L-threofuranosyl)-cytosine (1.10b).[15]

A mixture of 3.18 g (7.7 mmol) of N^4-benzoyl-1-(2′-*O*-benzoyl-α-L-threofuranosyl)cytosine (**1.9b**), 3.84 g (11.3 mmol, 1.5 equiv) of DMT-Cl, and 120 mg of DMAP (0.99 mmol, 0.1 equiv) was coevaporated twice with 15.0 mL of anhydrous pyridine. The residue was suspended in 50 mL of anhydrous pyridine and stirred under argon for 16 h at 80 °C. The solvent was evaporated under diminished pressure, and the residue was partitioned between EtOAc (50 mL) and H_2O (50 mL). The organic layer was washed with brine (50 mL), dried ($MgSO_4$), and evaporated. The residue was dissolved in a mixture of 38 mL of THF and 30.5 mL of MeOH, and then cooled to 0 °C. To the cold solution was added 7.7 mL of ice-cold 1 N aq. NaOH while stirring. After 10 min, another 7.7 mL of ice-cold 1 N aq. NaOH was added. Ten minutes after the second addition, the reaction was quenched by addition of 40 mL of 10% NH_4Cl in water. The volatile solvent was removed under diminished pressure, and the residue was diluted by addition of 50 mL of EtOAc. The organic

layer was separated, washed with H_2O (50 mL) and brine (50 mL), dried ($MgSO_4$), and evaporated. The residue was purified on a silica gel column (50 g silica bed), eluting with 70−100% EtOAc in hexane containing 1% Et_3N in steps of 5% increase in EtOAc for every 200 mL. *N*^{4}-Benzoyl-1-(3′-*O*-dimethoxytrityl-α-L-threo-furanosyl)cytosine (**1.10b**) was obtained as a white foam: yield 3.12 g (65%); ^{1}H NMR (400 MHz, $CDCl_3$): δ 3.32 (d, 1H, *J* = 10.0 Hz), 3.71 (d, 1H, *J* = 4.0 Hz), 3.74 (s, 6H), 4.25 (s, 1H), 4.33 (s, 1H), 5.70 (s, 1H), 6.79 (m, 4H), 6.75−7.35 (m, 9H), 7.47−7.67 (m, 4H), 7.94 (d, 2H, *J* = 8.0 Hz), 8.05 (d, 1H, *J* = 8.0 Hz), and 9.10 (brs, 1H).

_N_^{6}-Benzoyl-9-(3′-*O*-dimethoxytrityl-α-L-threofuranosyl)adenine (1.10c).[15]

A mixture of 3.39 g (7.6 mmol) of *N*^{6}-benzoyl-9-(2′-*O*-benzoyl-α-L-threofuranosyl)adenine (**1.9c**), 3.86 g (11.4 mmol, 1.5 equiv) of DMT-Cl, and 93 mg (0.76 mmol, 0.1 equiv) of DMAP was coevaporated twice with 7.0 mL of anhydrous pyridine. The residue was dissolved in 35 mL of anhydrous pyridine, and the mixture was stirred at 75 °C under argon for 16 h. The solvent was evaporated under diminished pressure, and the residue was partitioned between EtOAc (50 mL) and H_2O (50 mL). The organic layer was washed with brine (50 mL), dried ($MgSO_4$), and evaporated. The residue was dissolved in a mixture of 38 mL of THF and 30.5 mL of MeOH, and then cooled to 0 °C. To the cold stirred solution was added 7.6 mL of ice-cold 1 N aq. NaOH solution. After 10 min, another 7.6 mL of ice-cold 1 N aq. NaOH solution was added. Ten minutes after the second addition, the reaction was quenched with 40 mL of 10% aq. NH_4Cl. The volatile solvent was removed under diminished pressure, and the aqueous residue was diluted with 50 mL of EtOAc. The organic layer was separated, washed with H_2O (50 mL) and brine (50 mL), dried ($MgSO_4$), and evaporated. The residue was purified on a silica gel column (70 g silica bed), eluting with

30–80% EtOAc in hexanes containing 1% Et_3N. Stepwise elution was performed with 10% increases in EtOAc for each 200 mL volume. *N*^{6}-Benzoyl-9-(3′-*O*-dimethoxytrityl-α-L-threofuranosyl)adenine (**1.10c**) was obtained as a white foam: yield 3.6 g (63%); ^{1}H NMR (400 MHz, DMSO-*d*^{6}) 3.45 (dd, 1H, *J* = 5.2, 9.6 Hz), 3.59 (dd, 1H, *J* = 3.2, 9.6 Hz), 3.73 (s, 6H), 4.17 (m, 1H), 4.50 (m, 1H), 5.91 (d, 1H, *J* = 2.8 Hz), 5.93 (d, 1H, *J* = 4.8 Hz), 6.86 (m, 4H), 7.13–7.33 (m, 9H), 7.56 (t, 2H, *J* = 7.6 Hz), 7.65 (t, 1H, *J* = 7.2 Hz), 8.06 (d, 2H, *J* = 7.2 Hz), 8.56 (s, 1H), 8.75 (s, 1H), 11.22 (brs, 1H).

_N_^{2}-Acetyl-_O_^{6}-diphenylcarbamoyl-9-(3′-_O_-dimethoxytrityl-α-L-threofuranosyl) guanine (1.10d).[15]

5.2 g (8.74 mmol) of *N*^{2}-acetyl-*O*^{6}-diphenylcarbamoyl-9-(2′-*O*-benzoyl-α-L-threofuranosyl)guanine (**1.9d**) was coevaporated with 20 mL of anhydrous pyridine under reduced pressure, and the solid residue was dried under vacuum for 18 h. To the residue was added 5.92 g (17.48 mmol, 2 equiv) of DMT-Cl, anhydrous dichloromethane (60 mL), 2.3 mL of 2,4,6-trimethylpyridine (17.48 mmol, 2 equiv), and 675 mg (2.62 mmol, 0.3 equiv) of silver triflate. The mixture was stirred under nitrogen for 18 h. TLC (3:1 hexanes–EtOAc) showed the reaction to be complete. The solvent was removed under diminished pressure, and the residue was dissolved in 150 mL of dichloromethane. The organic layer was washed with 0.2 (N) HCl (300 mL), brine (60 mL), dried ($MgSO_4$), and evaporated. The residue was purified on a silica gel column (4 ×15 cm), eluting with 100 mL of 99:1 DCM–Et_3N to 300 mL of 97:2:1 DCM–MeOH–Et_3N. The product was obtained as a pale yellow powder: yield 5.8 g (74.3%). The compound (3.0 g, 3.34 mmol) was dissolved in 45 mL of THF–MeOH–water (5:4:1) and cooled to 0–5 °C in a water-ice bath. Ice-cold 1 N aq. NaOH (2.1 mL) was added dropwise to the mixture, and the mixture was stirred for 10–15 min. TLC (hexane–EtOAc 2:1) showed

that the reaction was not complete. Another 2.1 mL of cold 1 N aq. NaOH was added dropwise, and stirring continued for another 15–20 min. The reaction was quenched with 15 mL of 10% aq. NH_4Cl, and the crude product was extracted with EtOAc (2 ×50 mL). The combined organic layer was dried over $MgSO_4$ and evaporated under reduced pressure. The residue was purified on a silica gel column (4 × 15 cm) using a gradient elution with 100 mL of (85:14:1) hexane–EtOAc–Et_3N, followed by 100–300 mL of (75:24:1) hexane–EtOAc–Et_3N and (65:34:1) hexane–EtOAc–Et_3N. *N*^{2}-Acetyl-*O*^{6}-diphenylcarbamoyl-9-(3′-*O*-dimethoxytrityl-α-L-threofuranosyl)guanine (**1.10d**) was obtained as a pale yellow powder: yield 1.9 g (71%); ^{1}H NMR (400 MHz, $CDCl_3$): δ 2.17 (s, 3H), 2.73 (s, 1H), 3.42 (s, 2H), 3.71 (s, 6H), 4.33 (m, 1H), 4.45 (s, 1H), 5.83 (d, 1H, *J* = 2.4 Hz), 6.78 (m, 4H), 7.12–7.45 (m, 19H), 8.19 (s, 1H), 8.81 (s, 1H).

References

1. Schoning, K. U.; Scholz, P.; Guntha, S.; Wu, X.; Krishnamurthy, R.; Eschenmoser, A. *Science* **2000**, *290*, 1347.
2. Yang, Y.-W.; Zhang, S.; McCullum, E. O.; Chaput, J. C. *J. Mol. Evol.* **2007**, *65*, 289.
3. Ebert, M.-O.; Mang, C.; Krishnamurthy, R.; Eschenmoser, A.; Jaun, B. *J. Am. Chem. Soc*. **2008**, *130*, 15105.
4. Orgel, L. E. *Science* **2000**, *290*, 1306.
5. Joyce, G. F. *Nature* **2002**, *418*, 214.
6. Chaput, J. C.; Yu, H.; Zhang, S. *Chem. Biol*. **2012**, *19*, 1360.
7. Chaput, J. C.; Szostak, J. W. *J. Am. Chem. Soc.* **2003**, *125*, 9274.
8. Chaput, J. C.; Ichida, J. K.; Szostak, J. W. *J. Am. Chem. Soc*. **2003**, *125*, 856.
9. Kempeneers, V.; Vastmans, K.; Rozenski, J.; Herdewijn, P. *Nucleic Acids Res*. **2003**, *31*, 6221.
10. Horhota, A.; Zou, K.; Ichida, J. K.; Yu, B.; McLaughlin, L. W.; Szostak, J. W.; Chaput, J. C. *J. Am. Chem. Soc*. **2005**, *127*, 7427.
11. Yu, H.; Zhang, S.; Dunn, M.; Chaput, J. C. *J. Am. Chem. Soc*. **2013**, *135*, 3583.
12. Dunn, M. R.; Larsen, A. C.; Fahmi, N.; Zahurancik, W. J.; Meyers, M.; Suo, Z.; Chaput, J. C. *J. Am. Chem. Soc.* **2015**, *137*, 4014.
13. Yu, H.; Zhang, S.; Chaput, J. C. *Nat. Chem*. **2012**, *4*, 183.
14. Zou, K.; Horhota, A.; Yu, B.; Szostak, J. W.; McLaughlin, L. W. *Org. Lett*. **2005**, *7*, 1485.
15. Schoning, K.-U.; Scholz, P.; Wu, X.; Guntha, S.; Delgado, G.; Krishnamurthy, R.; Eschenmoser, A. *Helv. Chim. Acta* **2002**, *85*, 4111.
16. Zhang, S.; Yu, H.; Chaput, J. C. *Curr. Protoc. Nucleic Acid Chem*. **2013**, *52*, 4.54.1.
17. Zhang, S.; Chaput, J. C. *Curr. Protoc. Nucleic Acid Chem.* **2012**, *4*, 4.51.1.
18. Dumbre, S. G.; Jang, M.-Y.; Herdewijn, P. *J. Org. Chem.* **2013**, *78*, 7137.

19. Wei, C. C.; De Bernardo, S.; Tengi, J. P.; Borgese, J.; Weigele, M. *J. Org. Chem.* **1985**, *50*, 3462.
20. Wu, T.; Froeyen, M.; Kempeneers, V.; Pannecouque, C.; Wang, J.; Busson, R.; De Clercq, E.; Herdewijn, P. *J. Am. Chem. Soc.* **2005**, *127*, 5056.
21. Milecki, J.; Foldesi, A.; Fischer, A.; Adamiak, R. W.; Chattopadhyaya, J. *J. Labelled Compd. Radiopharm.* **2001**, *44*, 763.
22. Jang, M.-Y.; Froeyen, M.; Herdewijn, P. *Tetrahedron Lett.* **2013**, *54*, 6084.
23. Zou, R.; Robins, M. J. *Can. J. Chem*. **1987**, *65*, 1436.

Chapter 2

Synthesis of α-L-Threofuranosyl Nucleoside 3′-Monophosphates, 3′-Phosphoro(2-Methyl)imidazolides, and 3′-Triphosphates

Publication note

This paper was originally published in the *Journal of Organic Chemistry.*

Saikat Bala,† Jen-Yu Liao,† Hui Mei, and John C. Chaput Synthesis of α-L-Threofuranosyl Nucleoside 3′-Monophosphates, 3′-Phosphoro(2-Methyl)imidazolides, and 3′-Triphosphates. *J. Org. Chem*, 2017, 82, 5910−5916. († Equal contribution).

2.1 Contribution Statement

Chaput, J. C. and Bala, S. designed the synthetic strategy. All authors contributed to the synthesis, optimization, and characterization of the synthetic compounds. All authors contributed to the manuscript writing with the comments from Mei, H..

2.2 Abstract

α-L-Threofuranosyl nucleic acid (TNA) is an artificial genetic polymer composed of vicinal 2′,3′-phosphodiester bonds linking adjacent to threofuranosyl nucleosides. TNA is one of a small number of genetic polymers that capable of cross-pairing with DNA and RNA and are resistant to nuclease digestion. Although an efficient approach has been reported for synthesizing TNA nucleosides, very few advances have been made in the synthesis of phosphorylated TNA compounds. Here, we describe an highly efficient method for synthesizing α-L-threofuranosyl nucleoside 3′-monophosphates (tNMPs), 3′-phosphoro(2-methyl)imidazolides (2-MeImptNs), and 3′-triphosphates (tNTPs) bearing the four genetic

bases of adenine (A), cytosine (C), thymine (T), and guanine (G). We suggest that this strategy, which provides access to grams of tNMPs, hundreds of milligrams of 2-MeImptNs, and tens of milligrams of tNTPs, will help advance the use of TNA monomers in the downstream exobiology and biotechnology applications.

2.3 Introduction

Nucleoside triphosphates are indispensable reagents for a broad range of synthetic biology, molecular biology and biotechnology applications.[1–3] Particularly, modified nucleoside triphosphates that carry new chemical moieties at the sugar, nucleobase, and phosphate positions are critical to many therapeutic and diagnostic applications,[4–7] as well as fundamental studies intended for evaluating the specificity of natural enzymes or identifying emergent properties of life either here on Earth or elsewhere in the universe.[8] Unfortunately, despite extensive efforts to develop chemical and/or enzymatic strategies for producing nucleoside triphosphates, no universal strategy has been developed that can be applied to all nucleosides.[9] This is especially true for chemically modified analogues where a large number of synthetic methods have been developed to accommodate the idiosyncrasies of different nucleoside chemistries.[10–14] As expected, each strategy comes with its own strengths and weaknesses that must be evaluated on a case-by-case basis.[15] In general, enzymatic strategies suffer from poor substrate specificity, scale, yield, and cost, whereas chemical approaches struggle with functional group compatibility, regiochemistry, yield, and purification.[15]

Synthesizing nucleoside triphosphates of α-L-threofuranosyl nucleic acid (TNA) requires phosphorylating the 3′-hydroxyl position on the furanose ring.[16] However, the 3′-hydroxyl position of α-L-threofuranosyl nucleosides is substantially less reactive than the 5′-

hydroxyl position of DNA and RNA nucleosides because secondary alcohols are less nucleophilic than primary alcohols and the 3′-position is more sterically encumbered than the normal 5′-hydroxyl position of ribose sugar (**Figure 2-1**).

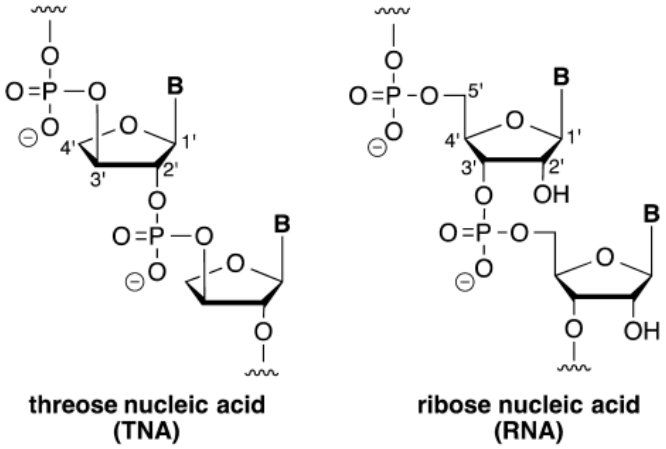
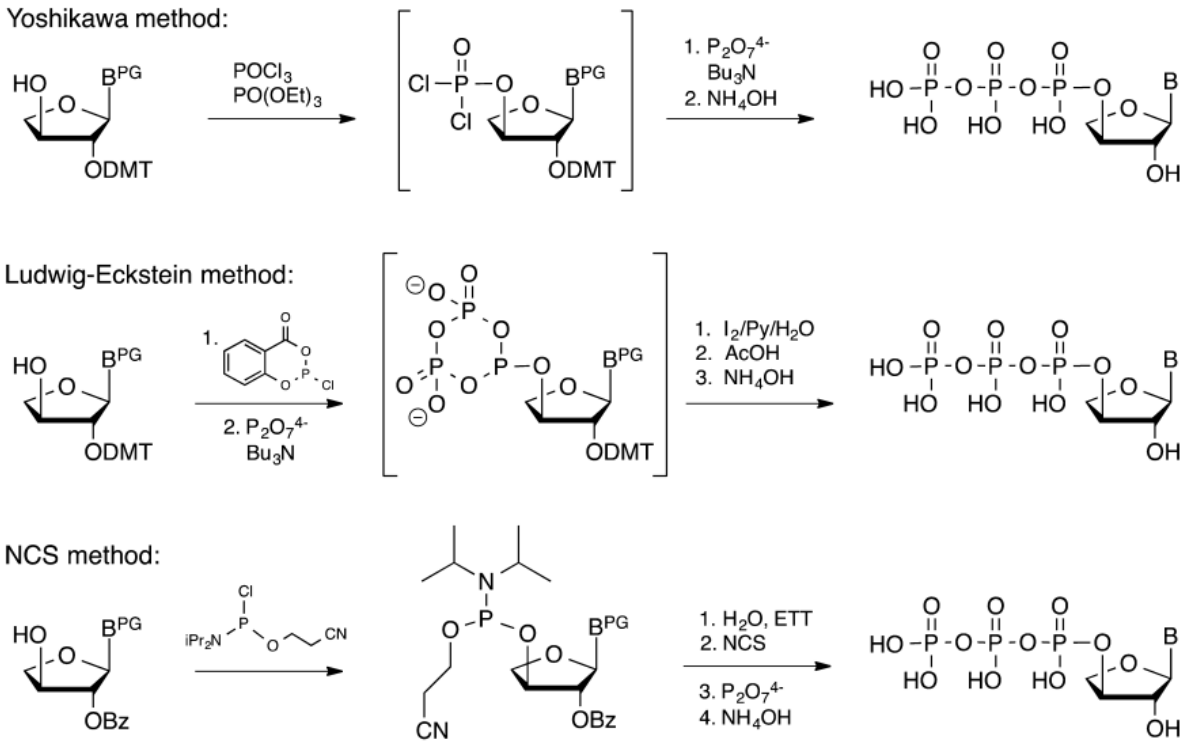


Figure 2-1. Molecular structure of TNA and RNA.

Given the importance of TNA building blocks in many exobiology and synthetic biology projects, a concerted effort has been made to develop synthetic approaches for constructing phosphorylated TNA compounds (**Scheme 2-1**) on the scales required to meet the needs of current downstream applications.[17–20] Early efforts by our laboratory and others applying the Yoshikawa method with $POCl_3$ proved unsuccessful due to unwanted side reactions. Szostak and co-workers overcame this problem using the classic Ludwig–Eckstein method, which was achieved all four TNA triphosphates synthesis from their corresponding 2′-*O*-DMT protected nucleosides.[21]

Although a detailed protocol was subsequently established,[22] this method proved cumbersome due to the poor regioselectivity observed when TNA nucleosides are tritylated on 3’-OH with DMT.[21] As an alternative strategy, we developed a one-pot four-step synthesis of TNA triphosphates from their 2′-*O*-benzoyl derivatives that is considerably less sensitive

to moisture than traditional approaches.[23] Using this method, H-phosphonate intermediates are oxidized in situ with N-chlorosuccinimide (NCS) to produce chlorophosphate intermediates that are treated with pyrophosphate to generate the desired nucleoside triphosphates.[23] Although TNA nucleoside 3′-triphosphates generated in this way are easy to purify by HPLC, the NCS method suffers from low overall yield (~10% isolated yield) due to an undesired side reaction that produces large quantities of the 3′-monophosphate derivatives as the major product.



Scheme 2-1. Previous Strategies for Synthesizing TNA Triphosphates

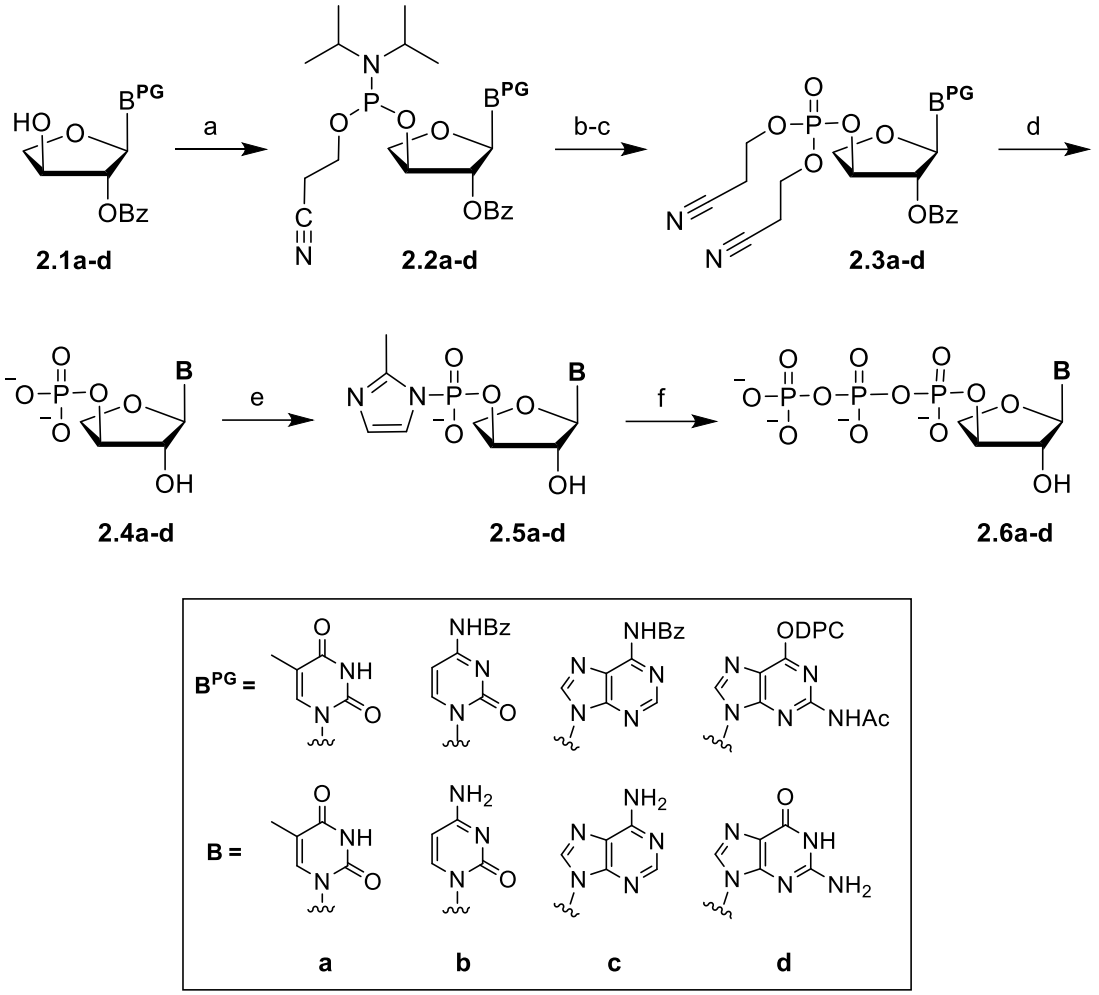
To overcome the shortcomings of previous methods, we developed a robust strategy for generating TNA nucleoside monophosphates, monophosphate derivatives, and nucleoside triphosphates that is straightforward to perform and could be transferrable to other modified nucleosides. Our approach consists of four synthetic steps to generate TNA

monophosphates from a TNA nucleoside intermediate that can be produced in large quantities from L-ascorbic acid.[24] TNA 3′-monophosphates are then converted to 3′-phosphoro(2-methyl)imidazolides, which are reacted with pyrophosphate to produce nucleoside 3′-triphosphates. HPLC analysis of the crude reaction mixture indicates that the phosphorylation steps proceed cleanly with high yield of the desired compounds. We suggest that the current strategy will be useful for constructing TNA nucleotides for a wide range of molecular biology and biotechnology applications, as TNA nucleoside 3′-monophosphates, 3′-monophosphate derivatives, and 3′-triphosphates are obtained as isolated compounds that can be stored for long periods of time without degradation.

2.4 Results and Discussion

We designed a synthetic approach (**Scheme 2-2**) for constructing 3′-phosphorylated TNA compounds from 2′-*O*-benzoyl TNA nucleosides **2.1a–d**. TNA nucleosides **2.1a–d** were obtained from L-ascorbic acid using known methodology (**Scheme 1-2 and 1-3**) that involves eight chemical transformations with three crystallization steps.[24] While several phosphorylation reagents were evaluated as possible routes to α-L-threofuranosyl 3′-monophosphates (**2.4a–d**), a stepwise process that invoked the synthesis of a 3′-phosphoramidite intermediate proved highly efficient for each of the four TNA nucleosides. Accordingly, 2′-*O*-benzoyl nucleosides **2.1a–d** were converted to their corresponding 3′-*O*-phosphoramidite derivatives **2.2a–d** upon treatment with 2-cyanoethoxy-*N*,*N*-diisopropylchlorophosphine in the presence of Hunig's base and DMAP. A slight excess of the chlorophosphoramidite reagent was crucial for obtaining 3′-*O*-phosphoramidites **2.2a–d** in high yield. Intermediates **2.2a–d** were subsequently activated by 1H-tetrazole and converted to their corresponding trialkyl phosphite by replacing the *N*,*N*-diisopropylamino

group with a base-labile cyanoethoxy group. The trialkyl phosphite was then oxidized with H_2O_2 to obtain the phosphate triesters **2.3a–d**, which were purified by silica gel



Scheme 2-2. Synthesis of TNA Triphosphates Using Activated TNA Monophosphates

Reagents and conditions: (a) 2-cyanoethyl *N,N*-diisopropylchlorophophoramidite, DMAP, DIPEA, CH_2Cl_2, rt, 30–45 min, 68–75%; (b) 3-hydroxypropionitrile, tetrazole, acetonitrile, rt, 1 h; (c) 30% H_2O_2, rt, 15 min, 71–75% (2 steps); (d) 30–33% NH_4OH, 38 °C, 18 h, 91–98%; (e) 2-methylimidazole, triphenylphosphine, 2,2′-dipyridyldisulfide, DMSO, triethylamine, rt, 6–8 h; (f) tributylammonium pyrophosphate, tributylamine, DMF, 10–12 h, 81–87% (2 steps).

chromatography. The resulting TNA nucleoside 3′-monophosphates **2.4a–d** were obtained as an ammonium salt following deprotection with 30% aqueous ammonium hydroxide. Although this strategy is more cumbersome than traditional monophosphate reactions,[25] we

found that the higher reactivity of the chlorophosphoramidite reagent by P(III) chemistry, coupled with the ability to purify the nucleoside monophosphate precursor **2.3a-d** by silica gel chromatography, provided an efficient method for obtaining large quantities (grams) of highly pure TNA 3′-monophosphates.

Next, TNA monophosphates **2.4a–d** were converted to their 3′-phosphoro(2-methyl)-imidazolide derivatives **2.5a–d** using classic methodology previously developed by Mukaiyama and co-workers.[26] Accordingly, a Mitsunobu-like reaction is performed by incubating **2.4a–d** with excess 2-methylimidazole and triphenylphosphine in the presence of 2,2′-dipyridyldisulfide under basic reaction condition. In some cases, complete conversion of the starting material required the addition of excess amount of 2-methylimidazole and triphenylphosphine. The desired TNA nucleoside 3′-phosphoro-(2-methyl)imidazolides

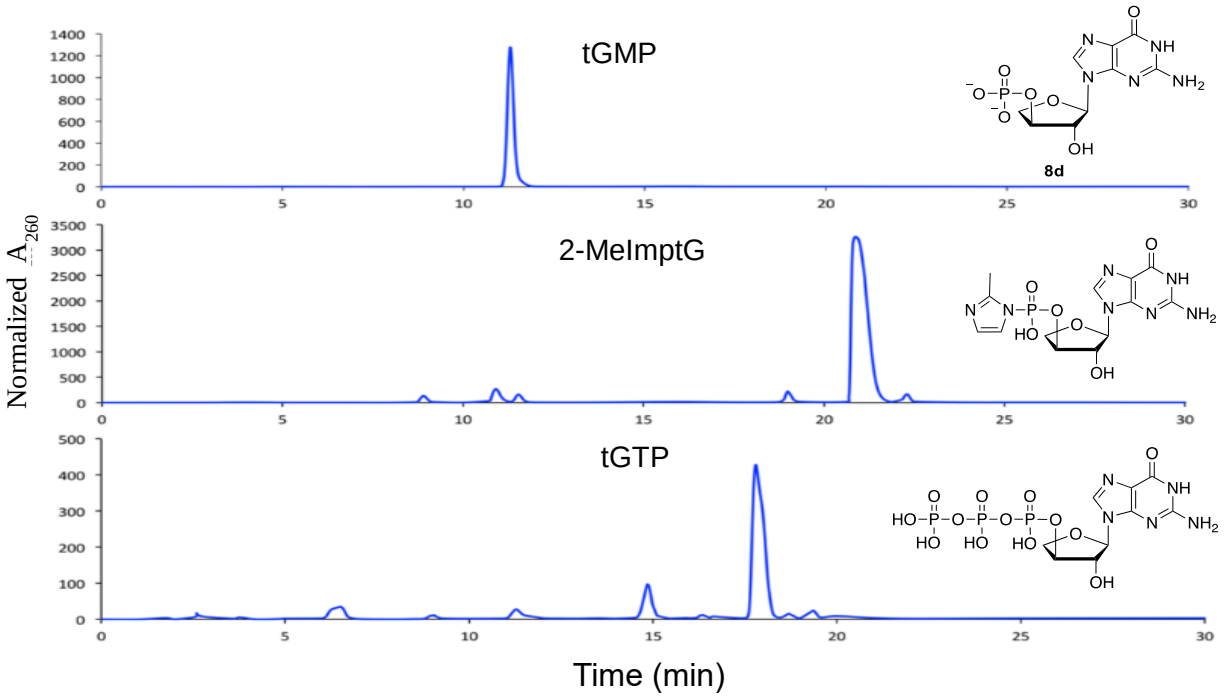


Figure 2-2. HPLC analysis of the crude reaction for tGMP, 2-MeImptG, and tGTP.

2.5a–d were isolated as a sodium salt by precipitating the reaction with excess of sodium perchlorate in acetone solution. The reactions were monitored by analytical HPLC until >90% conversion of the starting material to the desired product was observed (**Figure 2-2**)

Interestingly, we found that the phosphorimidazolide reaction proceeds more efficiently when the monophosphates are precipitated as the ammonium salt rather than the sodium salt.

TNA nucleoside 3′-triphosphates **2.6a–d** were synthesized from the activated nucleoside monophosphates **2.5a–d** by displacing the activated 2-methylimidazole leaving group with pyrophosphate. In these reactions, TNA nucleosides **2.5a–d** were incubated tributylammonium pyrophosphate and tributylamine in a DMF solution for 8–12 h at room temperature. The reaction was monitored by analytical HPLC, and additional pyrophosphate and tributylamine were added as needs to drive the reaction to completion (**Figure 2-2**). In all cases, >90% of the starting nucleoside was consumed within 12 h. The desired TNA nucleoside triphosphates **2.6a–d** were purified by reverse-phase HPLC, concentrated to dryness, resuspended in methanol, and precipitated as the sodium salt from acetone. During the concentration step, we noticed that the pH of the solution dropped due to a buildup of acetic acid from the ammonium acetate running buffer. If left unchecked, this change in pH leads to unwanted acid-catalyzed degradation of the nucleotide triphosphate. Fortunately, this problem is easily overcome by addition the triethylamine for pH adjustment as needed.

Next, we evaluated the thermal stability of TNA triphosphates (**2.6a–d**) bearing the four genetic bases of adenine (A), cytosine (C), thymine (T), and guanine (G). In this assay, the stability of each TNA triphosphate was monitored by analytical HPLC using small

aliquots of TNA solutions that were stored at temperatures of 4, 24, and 37 °C. A time course analysis performed over 8 days (**Table 2-1**) revealed that tNTPs are resistant to thermal degradation when stored in a buffered solution containing 10 mM Tris (pH 8.0). Even after 8 days at 37 °C, >70% of tGTP and >80% of tATP, tCTP, and tTTP remained undegraded. This result is consistent with previous thermal stability studies of tNTPs performed under standard PCR conditions.[23]

temperature (°C)	days	ATP (%)	GTP (%)	CTP (%)	TTP (%)
4	1	99.7	95.4	96.9	94.4
	5	98.4	92.7	93.6	89.4
	8	95.6	88.7	91.3	84.5
24	1	98.8	94.1	95.6	93.6
	5	90.5	86.5	94.9	83.3
	8	82.8	76.8	91.2	81.8
37	1	97.0	93.1	95.4	93.1
	5	89.6	82.4	93.2	82.4
	8	81.7	73.9	86.9	80.8

[a]Analysis performed in 10 mM Tris, pH 8.0.

Table 2-1. Thermal Stability of Purified TNA Triphosphates[a]

To confirm that tNTPs synthesized using our new strategy can function as substrates for an engineered TNA polymerase, we performed a standard primer-extension assay in which a library of DNA templates were copied into TNA product. To date, the polymerase that most efficiently uses tNTP substrates is Kod-RI, a polymerase that was developed through a combination of directed evolution and scaffold sampling.[20,27] Kod-RI functions with a catalytic rate of ~1 nucleotide per minute and is able to copy most DNA templates into TNA.[27] As a demonstration, an IR-labeled DNA primer was annealed to a degenerate DNA library and incubated with Kod-RI in the presence of chemically synthesized of tNTP substrates. Analysis of the primer-extension reaction by denaturing polyacrylamide gel

electrophoresis (**Figure 2-3**) reveals that the primer is extended to full-length product after a 3 h incubation at 55 °C. This result, which is consistent with the known activity of Kod-RI,[27] demonstrates that the new synthetic approach produces TNA triphosphates that are recognized by an engineered TNA polymerase.

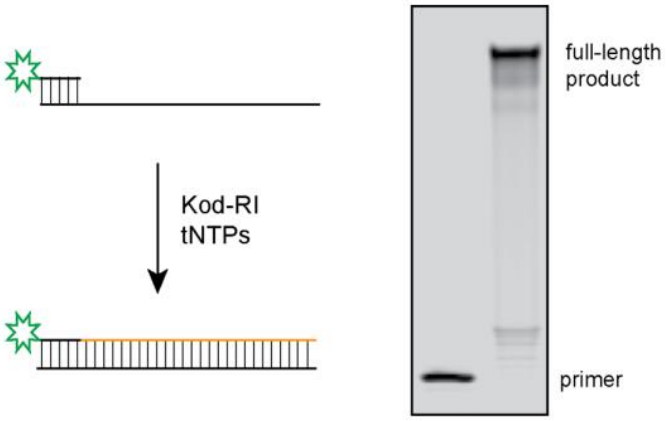


Figure 2-3. Primer-extension analysis

A 5′-IR-labeled DNA primer annealed to a DNA template was extended with 70 sequential TNA residues. Polymerization reaction was performed at 55 °C for 3 h using Kod-RI TNA polymerase and chemically synthesized tNTPs. The reaction was analyzed by denaturing polyacrylamide gel electrophoresis.

2.5 Conclusion

In summary, we report a novel synthetic route for constructing TNA nucleotide 3′-monophosphates, 3′-monophosphate derivatives, and 3′-triphosphates bearing all four nucleobases (A, T, G, C). This strategy overcomes the poor phosphorylation reaction of the 3′-hydroxyl group on the threose sugar by introducing a more convenient one-pot phosphitylating reagent with following oxidation reaction. Our approach involves five chemical transformations with one silica gel purification, two precipitation steps, and one HPLC purification. The overall isolated yield of the purified tNTP compounds from their starting nucleosides is 36–49% depending on the identity of nucleoside (A, T, G, C). Finally,

we show that the resulting TNA substrates are thermally stable and function as substrates for an engineered TNA polymerase. We suggest that the simplicity of this approach, coupled with the ability to monitor the phosphorylation reaction and isolate intermediate compounds, provides a general approach that could be applied to the facile synthesis for the phosphorylation of other modified nucleotides.

2.6 Experimental Details

General Procedures

All non-aqueous reactions were performed using oven-dried glassware under an atmosphere of argon or nitrogen. All chemicals were commercially available and used as received. Anhydrous solvents were purchased as the highest grade from Sigma-Aldrich. Reactions were monitored by thin layer chromatography using UV-activated TLC plates with silica gel 60 F254 and aluminum backing (Sigma-Aldrich, St. Louis, MO). Flash column chromatography was performed using SiliCycle 40−60 mesh silica gel (SiliCycle Inc., Quebec City, Canada). Yields are reported as isolated yields of pure compounds. UV quantification data are analyzed on NanoDrop 2000c using Beer's Law. ^{1}H, ^{13}C, and ^{31}P NMR spectra were analyzed on 400 and 500 MHz NMR spectrometers (Bruker, Billerica, MA). ^{1}H values are reported in parts per million relative to Me_4Si or corresponding deuterium solvents as internal standard. ^{13}C values are reported in parts per million relative to corresponding deuterium solvents as internal standards. ^{31}P NMR values are reported in parts per million relative to an external standard of 85% H_3PO_4. Splitting patterns are designated as follows: s, singlet; d, doublet; dd, doublet of doublets; t, triplet; m, multiplet. HPLC analysis was performed on a reverse-phase C18 150 × 4.6 mm column with 5 μm particle size, and TNA triphosphate was purified on a preparative reverse-phase C18 250 ×21.2 mm column

(Thermo Scientific, USA) using a mobile phase of 100 mM triethylammonium acetate buffer (pH 7.0)/acetonitrile.

1-(2′-O-Benzoyl-3′-O-[bis(2-cyanoethyoxy)phosphoryl]-α-L-threofuranosyl)thymine (2.3a).

To a stirring solution containing 1 g (3.01 mmol) of 1-(2′-O-benzoyl-α-L-threofuranosyl) thymine (**2.1a**)[24] and 360 mg (3.01 mmol) of DMAP in 12 mL anhydrous CH_2Cl_2 was added 1 mL (7.53 mmol) of *N,N*′-diisopropylethylamine followed by the addition of 0.81 mL (3.61 mmol) of 2-cyanoethoxy-*N,N*-diisopropylchlorophosphoramidite. After being stirred for 40 min at room temperature, the reaction mixture was diluted with 40 mL of CH_2Cl_2 and extracted twice with 50 mL of saturated aqueous $NaHCO_3$, washed with brine, dried over $MgSO_4$, and evaporated under reduced pressure. The crude 1-[2′-O-benzoyl-3′-O-(*N,N*-diisopropylamino-2-cyanoethoxyphosphinyl)-α-L-threofuranosyl]-thymine (**2.2a**)[23] was used without further purification.

To a stirring solution containing crude 1-[2′-O-benzoyl-3′-O-(*N,N*-diisopropylamino-2-cyanoethoxyphosphinyl)-α-L-threofuranosyl]-thymine (**2.2a**)[23] in 10 mL of anhydrous acetonitrile was added 0.62 mL (9.03 mmol) of 3-hydroxypropionitrile followed by the addition of 26.8 mL of tetrazole (0.45 M solution in anhydrous acetonitrile, 12.04 mmol). The reaction mixture was stirred for 1.5 h at room temperature with monitoring by TLC (1:1 hexanes−EtOAc). After complete consumption of the starting material, to the reaction was added 10 mL of 30% H_2O_2 and left to stir for 20 min at room temperature. The crude product was evaporated under reduced pressure and dissolved in 50 mL of ethyl acetate. The organic layer was washed with brine and water, dried over $MgSO_4$, and evaporated under reduced pressure. The crude was purified with silica gel column chromatography affording 1-(2′-*O*-

benzoyl-3′-*O*-[bis(2-cyanoethyoxy)phosphoryl]-α-L-threofuranosyl)thymine (**2.3a**) as a white foam: yield 0.87 g (56%); silica gel TLC (DCM/MeOH, 9:1) Rf = 0.35; ^{1}H NMR (500 MHz, CDCl3) δ 9.74 (s, 1H), 8.02 (d, 2H, *J* = 5 Hz), 7.61 (t, 2H, *J* = 7.5 Hz), 7.46 (t, 2H, *J* = 7.5 Hz), 7.30 (s, 1H), 6.11 (d, 1H, *J* = 2 Hz), 5.62 (s, 1H), 5.10 (m, 1H), 4.53 (m, 1H), 4.30−4.39 (m, 5H), 2.82 (t, 4H, *J* = 6 Hz), 1.92 (s, 3H); ^{13}C NMR (125.8 MHz, $CDCl_3$) δ 165.2, 164.1, 150.5, 135.3, 134.1, 130.0, 128.7, 128.3, 116.6, 116.6, 111.0, 89.5, 80.4 (d, $J_{C,P}$ = 6.5 Hz), 79.6 (d, $J_{C,P}$ =5.2 Hz), 73.5 (d, $J_{C,P}$ = 4.7 Hz), 63.1 (d, $J_{C,P}$ = 5.0 Hz), 63.0 (d, $J_{C,P}$ =5.0 Hz), 19.7 (d, $J_{C,P}$ = 7.4 Hz), 19.7 (d, $J_{C,P}$ = 7.2 Hz), 12.7; ^{31}P NMR (162 MHz, $CDCl_3$) δ −2.14; HRMS (ESI-TOF) calcd for $C_{22}H_{23}N_4O_9PNa$ [M + Na]+ 541.1100; found 541.1094.

***N*^{4}-(2′-*O*-Benzoyl-3′-*O*-[bis(2-cyanoethyoxy)phosphoryl]-α-L-threofuranosyl)-cytosine (2.3b).**

To a stirring solution containing 1 g (2.37 mmol) of *N*^{4}-benzoyl-1-(2′-*O*-benzoyl-α-L-threofuranosyl)cytosine (**2.1b**)[24] and 290 mg (2.37 mmol) of DMAP in 12 mL anhydrous CH_2Cl_2 was added 1 mL (5.93 mmol) of *N,N*′-diisopropylethylamine followed by the addition of 1.05 mL (4.74 mmol) of 2-cyanoethoxy-*N,N*-diisopropylchlorophosphoramidite. After being stirred for 40 min at room temperature, the reaction mixture was diluted with 40 mL of CH_2Cl_2 and extracted twice with 50 mL of saturated aqueous $NaHCO_3$, washed with brine, dried over $MgSO_4$, and evaporated under reduced pressure. The crude *N*^{4}-benzoyl-1-[2′-*O*-benzoyl-3′-*O*-(*N,N*-diisopropylamino-2-cyanoethoxy-phosphinyl)-α-L-threofuranosyl]cytosine (**2.2b**)[23] was used without further purification.

To a stirring solution containing crude *N*^{4}-benzoyl-1-[2′-*O*-benzoyl-3′-*O*-(*N,N*-diisopropylamino-2-cyanoethoxyphosphinyl)-α-L-threofuranosyl]cytosine (**2.2b**)[23] in 10 mL of anhydrous acetonitrile was added 0.65 mL (7.1 mmol) of 3-hydroxypropionitrile

followed by the addition of 21.07 mL of tetrazole (0.45 M solution in anhydrous acetonitrile, 9.48 mmol). The reaction mixture was stirred for 1.5 h at room temperature with monitoring by TLC (1:1 hexanes–EtOAc).

After complete consumption of the starting material, to the reaction was added 10 mL of 30% H_2O_2 and left to stir for 20 min at room temperature. The crude product solution was evaporated to dryness and redissolved with 50 mL of ethyl acetate. The organic layer was washed with brine and water, dried over $MgSO_4$, and evaporated under reduced pressure. The crude product was purified with silica gel column chromatography affording *N*^{4}-(2′-*O*-benzoyl-3′-*O*-[bis(2-cyanoethyoxy)phosphoryl]-α-L-threofuranosyl)cytosine (**2.3b**) as a white foam: yield 0.68 g (47.2%); silica gel TLC (DCM/acetone, 10:1) Rf = 0.35; ^{1}H NMR (400 MHz, DMSO-*d*$_6$) δ 11.30 (s, 1H), 8.21 (d, 1H, *J* = 7.2 Hz), 8.00–8.07 (m, 4H), 7.73 (t, 1H, *J* = 7.6 Hz), 7.57–7.65 (m, 3H), 7.52 (t, 2H, *J* = 7.2 Hz), 7.44 (d, 1H, *J* = 6.4 Hz), 6.07 (d, 1H, *J* = 1.2 Hz), 5.73 (d, 1H, *J* = 1.2 Hz), 5.21 (m, 1H), 4.63 (d, 1H, *J* = 10.4 Hz), 4.47 (dd, 1H, *J* = 3.6, 10.4 Hz), 4.23 (m, 4H), 2.94 (m, 4H); ^{13}C NMR (100.6 MHz, DMSO-*d*$_6$) *δ* 167.4, 164.4, 163.7, 154.5, 145.1, 134.1, 133.2, 132.8, 129.7, 128.9, 128.6, 128.5, 118.1, 118.0, 96.3, 90.9, 79.5 (d, $J_{C,P}$ = 6.3 Hz), 78.2 (d, $J_{C,P}$ = 5.0 Hz), 74.1 (d, $J_{C,P}$ = 3.8 Hz), 63.1 (d, $J_{C,P}$ = 4.9 Hz), 63.1 (d, $J_{C,P}$ = 4.9 Hz), 19.1 (d, $J_{C,P}$ = 7.4 Hz), 19.0 (d, $J_{C,P}$ = 7.5 Hz); ^{31}P NMR (162 MHz, DMSO-*d*$_6$) *δ* −2.61; HRMS (ESI-TOF) calcd for $C_{28}H_{26}N_5O_9PNa$ $[M + Na]^+$ 630.1366; found 630.1382.

_N_^{6}-(2′-*O*-Benzoyl-3′-*O*-[bis(2-cyanoethyoxy)phosphoryl]-α-L-threofuranosyl)adenine (2.3c).

To a stirring solution containing 1 g (2.24 mmol) of *N*^{6}-benzoyl-9-(2′-*O*-benzoyl-α-L-threofuranosyl) adenine (**2.1c**)[24] and 273 mg (2.24 mmol) of DMAP in 12 mL anhydrous CH_2Cl_2 was added 0.78 mL (4.48 mmol) of *N,N*′-diisopropylethylamine followed by the

addition of 0.85 mL (3.81 mmol) of 2-cyanoethoxy-*N,N*-diisopropylchlorophosphoramidite. After being stirred for 40 min at room temperature, the reaction mixture was diluted with 40 mL of CH_2Cl_2 and extracted twice with 50 mL of saturated aqueous $NaHCO_3$, washed with brine, dried over $MgSO_4$, and evaporated under reduced pressure. The crude N^6-benzoyl-9-[2′-*O*-benzoyl-3′-*O*-(*N,N*-diisopropylamino-2-cyanoethoxyphosphinyl)-α-L-threofuranosyl] adenine (**2.2c**)[23] was used without further purification.

To a stirring solution containing crude N^6-benzoyl-9-[2′-*O*-benzoyl-3′-*O*-(*N,N*-diisopropylamino-2-cyanoethoxy-phosphinyl)-α-L-threofuranosyl] adenine (**2.2c**)[23] in 10 mL of anhydrous acetonitrile was added 0.6 mL (6.72 mmol) of 3-hydroxypropionitrile followed by the addition of 19.9 mL of tetrazole (0.45 M solution in anhydrous acetonitrile, 8.96 mmol). The reaction mixture was stirred for 1.5 h at room temperature with monitoring by TLC (1:1 hexanes–EtOAc). After complete consumption of the starting material, to the reaction was added 10 mL of 30% H_2O_2 and left to stir for 20 min at room temperature. The crude product solution was evaporated to dryness and redissolved with 50 mL of ethyl acetate. The organic layer was washed with brine and water, dried over $MgSO_4$, and evaporated under reduced pressure. The crude product was purified with silica gel column chromatography affording N^6-(2′-*O*-benzoyl-3′-*O*-[bis(2-cyanoethyoxy)phosphoryl]-α-L-threofuranosyl)adenine (**2.3c**) as a white foam: yield 0.65 g (46%); silica gel TLC (DCM/MeOH, 9:1) Rf = 0.35; ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.80 (s, 1H), 8.62 (s,1H), 8.06 (m, 4H), 7.72 (m, 1H), 7.66 (m, 1H), 7.57 (m, 4H), 6.52 (d, 1H, J = 2.8 Hz), 6.29 (t, 1H, J = 2.8 Hz), 5.45 (m, 1H), 4.6 (dd, 1H, J = 4, 10.4 Hz), 4.52 (dd, 1H, J = 4.8, 10.4 Hz), 4.24 (m, 4H), 2.96 (m, 4H); ^{13}C NMR (125.8 MHz, DMSO-d_6) δ 165.6, 164.7, 152.0, 151.9, 150.5, 143.0, 134.1, 133.3, 132.5, 129.7, 128.9, 128.5, 128.5, 128.4, 125.8, 118.2, 118.1, 87.5, 79.7, 78.6 (d, $J_{C,P}$ =

6.9 Hz), 72.1 (d, $J_{C,P}$ = 4.5 Hz), 63.1 (d, $J_{C,P}$ = 5.0 Hz), 63.0 (d, $J_{C,P}$ = 5.0 Hz), 19.1 (d, $J_{C,P}$ = 7.3 Hz), 19.0 (d, $J_{C,P}$ = 7.2 Hz); ^{31}P NMR (162 MHz, DMSO-d_6) δ −2.25; HRMS (ESI-TOF) calcd for $C_{29}H_{26}N_7O_8PNa$ $[M + Na]^+$ 654.1485; found 654.1478.

N^2-Acetyl-O^6-diphenylcarbamoyl-9-(2′-*O*-benzoyl-3′-*O*-[bis-(2-cyanoethyoxy)phosphoryl]-α-L-threofuranosyl)guanine(2.3d).

To a stirring solution containing 1 g (1.68 mmol) of N^2-acetyl-O^6-diphenylcarbamoyl-9-(2′-*O*-benzoyl-α-L-threofuranosyl) guanine (**2.1d**)[24] and 210 mg (1.68 mmol) of DMAP in 12 mL anhydrous CH_2Cl_2 was added 0.58 mL (3.36 mmol) of *N,N*′-diisopropylethylamine followed by the addition of 0.821 mL (3.67 mmol) of 2-cyanoethoxy-*N,N*-diisopropylchlorophosphoramidite. After being stirred for 40 min at room temperature, the reaction mixture was diluted with 40 mL of CH_2Cl_2 and extracted twice with 50 mL of saturated aqueous $NaHCO_3$, washed with brine, dried over $MgSO_4$, and evaporated under reduced pressure. The crude N^2-acetyl-O^6-diphenylcarbamoyl-9-[2′-*O*-benzoyl-3′-*O*-(*N,N*-diisopropylamino-2-cyanoethoxyphosphinyl)-α-L-threofuranosyl]guanine (**2.2d**)[23] was used without further purification.

To a stirring solution containing crude N^2-acetyl-O^6-diphenylcarbamoyl-9-[2′-*O*-benzoyl-3′-*O*-(*N,N*-diisopropylamino-2-cyanoethoxy-phosphinyl)-α-L-threofuranosyl] guanine (**2.2d**)[23] in 10 mL of anhydrous acetonitrile was added 0.36 mL (5.04 mmol) of 3-hydroxypropionitrile followed by the addition of 14.9 mL of tetrazole (0.45 M solution in anhydrous acetonitrile, 6.72 mmol). The reaction mixture was stirred for 1.5 h at room temperature with monitoring by TLC (1:1 hexanes−EtOAc). After complete consumption of the starting material, to the reaction was added 10 mL of 30% H_2O_2 and left to stir for 20 min at room temperature. The crude product solution was evaporated to dryness and

redissolved with 50 mL of ethyl acetate. The organic layer was washed with brine and water, dried over $MgSO_4$, and evaporated under reduced pressure. The crude product was purified with silica gel column chromatography affording *N*^{2}-acetyl-*O*^{6}-diphenylcarbamoyl-9-(2′-*O*-benzoyl-3′-*O*-[bis(2-cyanoethyoxy)-phosphoryl]-α-L-threofuranosyl)guanine (**2.3d**) as a white foam: yield 0.59 g (45%); silica gel TLC (DCM/acetone, 10:1) Rf = 0.35; ^{1}H NMR (500 MHz, DMSO-*d*^{6}) δ 10.78 (s, 1H), 8.55 (s, 1H), 8.03 (d, 2H, *J* = 7.5 Hz), 7.71 (t, 1H, *J* = 7.5 Hz), 7.44−7.57 (m, 10H), 7.33 (t, 2H, *J* = 6.5 Hz), 6.39 (s, 1H), 6.21 (s, 1H), 5.41 (s, 1H), 4.75 (m, 1H), 4.47 (m, 1H), 4.24 (m, 4H), 2.94 (m, 4H), 2.20 (s, 3H); ^{13}C NMR (125.8 MHz, DMSO-*d*$_6$) δ 168.9, 164.8, 155.3, 154.2, 152.4, 150.1, 144.0, 141.6, 134.1, 129.7, 129.5, 128.9, 128.4, 127.4, 120.3, 118.2, 118.0, 116.7, 87.3, 79.5 (d, $J_{C,P}$ = 7.0 Hz), 78.0 (d, $J_{C,P}$ = 4.3 Hz), 71.6 (d, $J_{C,P}$ = 3.8 Hz), 63.0 (d, $J_{C,P}$ = 5.0 Hz), 63.0 (d, $J_{C,P}$ = 5.0 Hz), 24.7, 19.0 (d, $J_{C,P}$ = 7.9 Hz), 19.0 (d, $J_{C,P}$ = 7.7 Hz); ^{31}P NMR (162 MHz, DMSO-*d*$_6$) δ −2.24; HRMS (ESI-TOF) calcd for $C_{37}H_{33}N_8O_{10}PH$ $[M+H]^+$ 780.2281; found 780.2296.

α-L-Threofuranosyl nucleoside 3′-monophosphates (2.4a−d).

In a sealed tube, 1-(2′-*O*-benzoyl-3′-*O*-[bis(2-cyanoethyoxy)-phosphoryl]-α-L-threofuranosyl) nucleosides (0.3 mmol) (**2.3a−d**) were combined with 10 mL of saturated NH_4OH. The reaction was stirred at 37 °C for 16 h. The mixture was cooled to room temperature, and the solvent was evaporated to dryness. The residue was resuspended in 3 mL of methanol at 40 °C with stirring. The 40 mL of acetone was dropwise added into mixture to precipitate the product as ammonium salt. The precipitate was collected by centrifugation at 4400 rpm at room temperature for 15 min, and the resulting pellet was washed twice with 30 mL of acetone and dried under high vacuum. The product α-L-

threofuranosyl nucleoside 3′-monophosphates (**2.4a–d**) were obtained as the ammonium salt (white solid) in near quantitative yield.

α-L-Threofuranosyl thymine-3′-monophosphate (2.4a).

Product yield: 90.8 mg (98.3%, ε_{267} = 9600); ^{1}H NMR (400 MHz, D_2O) δ 7.61 (d, 1H), 5.82 (s, 1H), 4.62 (s, 1H), 4.46 (m, 2H), 4.34 (m, 1H), 1.90 (s, 3H); ^{13}C NMR (100.6 MHz, D_2O) δ 166.9, 151.8, 138.1, 110.5, 92.0, 79.3 (d, $J_{C,P}$ = 4.4 Hz), 78.2 (d, $J_{C,P}$ = 4.2 Hz), 75.1 (d, $J_{C,P}$ = 4.7 Hz), 11.9; ^{31}P NMR (162 MHz, D_2O) δ 0.21; HRMS (ESI-TOF) calcd for $C_9H_{13}N_2O_8PNa$ $[M + Na]^+$ 331.0307; found 331.0311.

α-L-Threofuranosyl cytosine-3′-monophosphate (2.4b).

Product yield: 84.4 mg (96%, ε_{280} = 13100); ^{1}H NMR (400 MHz, DMSO-d_6) δ 7.65 (d, 1H, J = 7.2 Hz), 5.83 (d, 1H, J = 6.8 Hz), 5.65 (s, 1H), 4.38 (s, 1H), 4.28 (s, 2H), 4.07 (d, 1H, J = 7.6 Hz); ^{13}C NMR (125.8 MHz, DMSO-d_6) δ 163.8, 153.4, 143.1, 93.8, 92.7, 78.8, 78.3, 75.1; ^{31}P NMR (162 MHz, DMSO-d_6) δ −0.65; HRMS (ESI-TOF) calcd for $C_8H_{11}N_3O_7P$ $[M - H]^-$ 292.0326; found 292.0335.

α-L-Threofuranosyl adenine-3′-monophosphate (2.4c).

Product yield: 88.5 mg (93%, ε_{259} = 15200); ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.30 (s, 1H), 8.16 (s, 1H), 5.91 (d, 1H, J = 1.6 Hz), 4.57 (s, 2H), 4.23 (d, 1H, J = 8 Hz), 4.12 (dd, 1H, J = 5.2, 4.0 Hz); ^{13}C NMR (125.8 MHz, DMSO-d_6) δ 156.0, 153.0, 149.5, 140.2, 118.6, 89.3, 79.3 (d, $J_{C,P}$ = 5.2 Hz), 78.2 (d, $J_{C,P}$ = 4.7 Hz), 73.5 (d, $J_{C,P}$ = 4.8 Hz), 75.1; ^{31}P NMR (162 MHz, DMSO-d_6) δ 1.49; HRMS (ESI-TOF) calcd for $C_9H_{11}N_5O_6P$ $[M - H]^-$ 316.0479; found 316.0447.

α-L-Threofuranosyl guanine-3′-monophosphate (2.4d).

Product yield: 90.9 mg (91%, ε_{253} = 13700); ^{1}H NMR (400 MHz, DMSO-d_6) δ 7.88 (s, 1H), 5.68 (s, 1H), 4.45 (d, 2H, J = 26.4 Hz); ^{13}C NMR (125.8 MHz, DMSO-d_6) δ 157.0, 154.1, 151.2, 136.3, 116.0, 88.6, 79.2 (d, $J_{C,P}$ = 4.8 Hz), 78.0 (d, $J_{C,P}$ = 3.6 Hz), 73.1 (d, $J_{C,P}$ = 4.0 Hz); ^{31}P NMR (162 MHz, DMSO-d_6) δ −0.17; HRMS (ESI-TOF) calcd for $C_9H_{11}N_5O_7P$ $[M - H]^-$ 332.0407; found 332.0396.

α-L-Threofuranosyl nucleosides 3′-phosphor-2-methylimidazolides (2.5a–d).

To a solution containing α-L-threofuranosyl thymidine-3′-monophosphate (**2.4a–d**) (0.27 mmol) and 2-methylimidazole (2.7 mmol) in 5 mL of anhydrous DMSO were added triethylamine (2.7 mmol), triphenylphosphine (1.08 mmol), and 2,2′-dipyridyldisulfide (1.08 mmol). The reaction was stirred under a nitrogen atmosphere for 6–8 h at room temperature with monitoring by analytical HPLC. After consumption of the starting material, the product was precipitated by the dropwise addition of the reaction mixture to a stirring solution containing 80 mL of acetone, 60 mL of diethyl ether, 5 mL of triethylamine, and 5 mL of saturated $NaClO_4$ in acetone. The precipitate was collected by centrifugation at 4400 rpm for 15 min at room temperature. The pellet was washed twice with 30 mL of washing solution (acetone/diethyl ether 1:1) and dried under high vacuum to afford the α-L-threofuranosyl nucleoside 3′-phosphor-2-methylimidazolides (**2.5a–d**) as the sodium salt.

α-L-Threofuranosyl thymine-3′-phosphor-2-methylimidazolide (2.5a).

Product yield: 97.9 mg (97.4%, ε_{267} = 9600); ^{31}P NMR (162 MHz, D_2O) δ −7.88; HRMS (ESI-TOF) calcd for $C_{13}H_{17}N_4O_7PNa$ $[M + Na]^+$ 395.0732; found 395.0717.

α-L-Threofuranosyl cytosine-3′-phosphor-2-methylimidazolide (2.5b).

Product yield: 93.1 mg (96.6%, ε_{280} = 13100); ^{31}P NMR (162 MHz, D_2O) δ −8.07; HRMS (ESI-TOF) calcd for $C_{12}H_{16}N_5O_6PNa$ $[M + Na]^+$ 380.0736; found 380.0735.

α-L-Threofuranosyl adenine-3′-phosphor-2-methylimidazolide (2.5c).

Product yield: 100.5 mg (97.7%, ε_{259} = 15200); ^{31}P NMR (162 MHz, D_2O) *δ* −8.37; HRMS (ESI-TOF) calcd for $C_{13}H_{16}N_7O_5PNa$ $[M + Na]^+$ 404.0848; found 404.0847.

α-L-Threofuranosyl guanine-3′-phosphor-2-methylimidazolide (2.5d).

Product yield: 104 mg (97%, ε_{253} = 13700); ^{31}P NMR (162 MHz, D_2O) δ −8.40; HRMS (ESI-TOF) calcd for $C_{13}H_{16}N_7O_6PNa$ $[M + Na]^+$ 420.0797, found 420.0807.

α-L-Threofuranosyl nucleosides 3′-Triphosphates (2.6a–d).[21–23]

To a solution containing 0.1 mmol of α-L-threofuranosyl nucleosides 3′-phosphor-2-methylimidazolides (**2.5a–d**) and 2 mL of in anhydrous DMF were added tributylamine (0.2 mmol) and tributylammonium pyrophosphate (0.2 mmol). The reaction mixture was then stirred under nitrogen atmosphere for 8–12 h at room temperature with monitoring by analytical HPLC. After the reaction was finished, the reaction mixture was added dropwise to a stirring solution containing 30 mL of acetone and 5 mL of saturated $NaClO_4$ in acetone. The precipitate was collected by centrifugation at 4400 rpm for 15 min at room temperature and dried under vacuum for 1 h. The crude precipitate was dissolved in 3 mL of 0.1 M triethylammonium acetate buffer and purified by a semipreparative HPLC. Fractions containing triphosphates were collected and concentrated, pH adjusted by triethylamine to 7.0, and lyophilized to afford the product as a triethylammonium salt. The solid product was resuspended in 3 mL of methanol and was added dropwise to a solution containing 40 mL of acetone and 2 mL of saturated $NaClO_4$ in acetone. The solution was centrifuged at 4400 rpm

for 15 min at room temperature. The supernatant was discarded, and the pellet was washed with 30 mL of acetone and dried under vacuum for 1 h. The resulting white solid was dissolved in RNase-free water containing 10 mM of Tris pH 8.0 to afford the α-L-threofuranosyl nucleotide 3′-triphosphate (**2.6a–d**) solution.

α-L-Threofuranosyl thymine-3′-triphosphate (2.6a).

Product yield after HPLC purification: 42.6 mg (91.1%, ε_{267} = 9600).

α-L-Threofuranosyl cytosine-3′-triphosphate (2.6b).

Product yield after HPLC purification: 40 mg (88.3%, ε_{280} = 13100).

α-L-Threofuranosyl adenine-3′-triphosphate (2.6c).

Product yield after HPLC purification: 43 mg (90.1%, ε_{259} = 15200).

α-L-Threofuranosyl guanine-3′-triphosphate (2.6d).

Product yield after HPLC purification: 44.2 mg (89.7%, ε253 = 13700).

Thermal Stability Analysis of TNA Triphosphates.

A small volume (30 μL) of 4 mM of **2.6a–d** in 10 mM of Tris buffer (pH 8) was maintained at different temperatures (4, 24, and 37 °C) for a period of 8 days. At specific time periods (1, 5, and 8 days), 5 μL of sample was transferred to 150 μL of HPLC running buffer [0.1 M triethylammoniumacetate (TEAA)], and the samples (50 μL) were analyzed by analytical reverse-phase HPLC with a gradient of 0 to 7% acetonitrile in 0.1 M TEAA buffer over 40 min. Analytical HPLC traces were analyzed, and relative peak areas are reported in Table 1. Polymerase-Mediated Primer Extension. The primer extension experiment was done in a single PCR tube with 100 μL of reaction volume containing a DNA primer–template

complex (50 pmol). The complex was labeled with an IR800 dye at the 5′-end of the DNA primer. The primer–template complex was annealed in 1× thermoPol buffer (20 mM Tris·HCl, 0 mM $(NH_4)_2SO_4$), 10 mM KCl, 2 mM $MgSO_4$, 0.1% Triton X-100, pH 8) by heating for 5 min at 95 °C and cooling for 10 min at 4 °C. An engineered polymerase, Kod-RI (10 μL), was pretreated with $MnCl_2$ (1 mM) and added to the reaction mixture. Newly formed **2.6a–d** (100 μM) were then added to the reaction mixture, and the solution was incubated for 3 h at 55 °C. The reaction was analyzed by denaturing polyacrylamide gel electrophoresis.

References

1. Cameron, D. E.; Bashor, C. J.; Collins, J. J. *Nat. Rev. Microbiol*. **2014**, *12*, 381–390.
2. Chen, Y. Y.; Galloway, K. E.; Smolke, C. D. *Genome Biol.* **2012**, 13, 240.
3. Chaput, J. C.; Yu, H.; Zhang, S. *Chem. Biol.* **2012**, *19*, 1360–1371.
4. Pinheiro, V. B.; Taylor, A. I.; Cozens, C.; Abramov, M.; Renders, M.; Zhang, S.; Chaput, J. C.; Wengel, J.; Peak-Chew, S. Y.; McLaughlin, S. H.; Herdewijn, P.; Holliger, P. *Science* **2012**, 336, 341–344.
5. Sefah, K.; Yang, Z.; Bradley, K. M.; Hoshika, S.; Jimenez, E.; Zhang, L.; Zhu, G.; Shanker, S.; Yu, F.; Turek, D.; Tan, W.; Benner, S. A. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 1449–1454.
6. Kimoto, M.; Yamashige, R.; Matsunaga, K.-I.; Yokoyama, S.; Hirao, I. *Nat. Biotechnol.* **2013**, *31*, 453–457.
7. Malyshev, D. A.; Dhami, K.; Lavergne, T.; Chen, T.; Dai, N.; Foster, J. M.; Correa, I. R.; Romesberg, F. E. *Nature* **2014**, *509*, 385–388.
8. Adamala, K.; Szostak, J. W. *Science* **2013**, *342*, 1098–1100.
9. Burgess, K.; Cook, D. *Chem. Rev.* **2000**, *100*, 2047–2060.
10. Dellafiore, M. A.; Montserrat, J. M.; Iribarren, A. M. *Front. Chem.* **2016**, *4*, 18.
11. Hocek, M. *J. Org. Chem.* **2014**, *79*, 9914–9921.
12. Hottin, A.; Marx, A. *Acc. Chem. Res.* **2016**, *49*, 418–427.
13. Hollenstein, M. *Molecules* **2012**, *17*, 13569–13591.
14. Cremosnik, G. S.; Hofer, A.; Jessen, H. *J. Angew. Chem., Int. Ed.* **2014**, *53*, 286–289.
15. Kore, A. R.; Srinivasan, B. *Curr. Org. Synth.* **2014**, *10*, 903–934.
16. Schöning, K. U.; Scholz, P.; Guntha, S.; Wu, X.; Krishnamurthy, R.; Eschenmoser, A. *Science* **2000**, *290*, 1347–1351.
17. Blain, J. C.; Ricardo, A.; Szostak, J. W. *J. Am. Chem. Soc.* **2014**, *136*, 2033–2039.
18. Chaput, J. C.; Szostak, J. W. *J. Am. Chem. Soc.* **2003**, *125*, 9274–9275.
19. Yu, H.; Zhang, S.; Chaput, J. C. *Nat. Chem*. **2012**, *4*, 183–187.
20. Larsen, A. C.; Dunn, M. R.; Hatch, A.; Sau, S. P.; Youngbull, C.; Chaput, J. C. *Nat. Commun.* **2016**, *7*, 11235.
21. Zou, K.; Horhota, A.; Yu, B.; Szostak, J. W.; McLaughlin, L. W. *Org. Lett.* **2005**, *7*, 1485–1487.
22. Zhang, S.; Yu, H.; Chaput, J. C. *Current Protocols in Nucleic Acid Chemistry*; *John Wiley & Sons: Hoboken, NJ,* **2013**, *52*, p 4.54.1.
23. Sau, S. P.; Chaput, J. C. *Bioorg. Med. Chem. Lett.* **2016**, *26*, 3271–3273.

24. Sau, S. P.; Fahmi, N. E.; Liao, J.-Y.; Bala, S.; Chaput, J. C. *J. Org. Chem.* **2016**, *81*, 2302–2307.
25. Yoshikawa, M.; Kato, T.; Takenishi, T. *Tetrahedron Lett.* **1967**, *8*, 5065–5068.
26. Mukaiyama, T.; Hashimoto, M. *J. Am. Chem. Soc.* **1972**, 94, 8528–8532.
27. Dunn, M. R.; Otto, C.; Fenton, K. E.; Chaput, J. C. *ACS Chem. Biol.* **2016**, *11*, 1210–1219

Chapter 3

Synthesis and Evolution of a Threose Nucleic Acid Aptamer Bearing 7-Deaza-7-Substituted Guanosine Residues

Publication note

This paper was originally published in the *Journal of American Chemical Society.*

Hui Mei,† Jen-Yu Liao,† Randi M. Jimenez, Yajun Wang, Saikat Bala, Cailen McCloskey, Christopher Switzer, and John C. Chaput. Synthesis and Evolution of a Threose Nucleic Acid Aptamer Bearing 7-Deaza-7-Substituted Guanosine Residues. *J. Am. Chem. Soc*. 2018, 140, 5706. († Equal contribution).

3.1 Contribution Statement

Liao, J.-Y. and Hui, M. designed the synthetic strategy. Liao, J.-Y., Hui, M., and Bala, S. contributed to the optimization and compound characterization. Jimenez, R. M., and McCloskey, C. performed the aptamer selection and compound characterization. Wang, Y. determined the fidelity of TNA replication. Christopher Switzer contributed to the computational analysis. All authors contributed to the experimental section writing and supporting information compiling. All authors contributed to the manuscript writing.

3.2 Abstract

In vitro selection experiments carried out on artificial genetic polymers require robust and faithful methods for copying genetic information back and forth between DNA and xeno-nucleic acids (XNA). Previously, we have shown that Kod-RI, an engineered polymerase developed to transcribe DNA templates into threose nucleic acid (TNA), can function with high fidelity in the absence of manganese ions. However, the transcriptional efficiency of this enzyme diminishes greatly when individual templates are replaced with

libraries of DNA sequences, indicating that manganese ions are still required for in vitro selection. Unfortunately, the presence of manganese ions in the transcription mixture leads to the misincorporation of tGTP nucleotides opposite dG residues in the templating strand, which are detected as G-to-C transversions when the TNA is reverse transcribed back into DNA. Here we report the synthesis and fidelity measurement of TNA replication using 7-deaza-7-modified guanosine base analogues in the DNA template and incoming TNA nucleoside triphosphate. The results reveal that tGTP misincorporation occurs via a Hoogsteen base pair in which the incoming tGTP residue adopts a syn conformation with respect to the sugar by polymerase Kod-RI incorporation. Substitution of tGTP for 7-deaza-7-phenyl tGTP enabled the synthesis of TNA polymers with >99% overall fidelity. A TNA library containing the 7-deaza-7-phenyl guanine analogue was used to evolve a biologically stable TNA aptamer that binds to HIV reverse transcriptase with low nanomolar affinity.

3.3 Introduction

Aptamers are nucleic acid molecules that fold into stable three-dimensional structures. Those molecules mimic antibodies with ligand binding sites that are complementary in shape and charge to the surface of a desired target.[1,2] Although aptamers have made a significant impact on a wide range of scientific fields from origins of life to healthcare,[3] aptamers of natural genetic polymers (DNA and RNA) have limited potential in practical applications that demand high biological stability.[4] As an instance, an unmodified DNA aptamer developed to bind human α-thrombin exhibited an half-life of <2 min when assayed in a primate animal model.[5] Efforts to overcome this problem have focused on the development of modifications on nucleic acid structure chemically to against nucleic acid degrading enzymes.[6] In particular, amino, fluoro, or methoxy groups has been exploited as a

common approach to substitute the 2′-hydroxyl moiety of ribonucleotides for producing aptamers with enhanced nuclease stability.[7] However, 2′-modified aptamers are still prone to degrade in the presence of natural digesting nucleases, albeit at reduced levels, indicating that biological stability is a spectrum of activities.[8,9]

A fundamentally different approach to nuclease stability involves generating xeno-nucleic acid (XNA) polymers with backbone structures that are distinct from those found in nature.[10] This strategy requires polymerases that can convert genetic information back and forth between DNA and XNA for the evolution of functional XNA molecules by in vitro selection.[11,12] To date, functional aptamers and catalysts have been evolved to five different types of XNA polymers: 2′-fluoro-arabino nucleic acid (FANA), arabino nucleic acid (ANA), 1,5-anhydrohexitol nucleic acid (HNA), cyclohexenyl nucleic acid (CeNA), and α-L-threose nucleic acid (TNA).[13–16] Of these, TNA stands out as one of the most structurally distinct nucleic acid polymers developed for molecular evolution, as it has a backbone repeat unit that is one atom shorter than that of DNA or RNA (**Figure 3-1**).[17] By comparison, the remaining XNA systems each maintain the same six-atom backbone repeat unit found in natural DNA and RNA.[18] Remarkably, despite a shorter backbone repeat unit, TNA is still capable of forming the efficient antiparallel Watson−Crick base pairing with complementary strands of DNA and RNA, as well as with other strands of TNA.[17,19] This feature, coupled with a backbone structure that is refractory to nuclease digestion, makes TNA a promising candidate for diagnostic and therapeutic applications that require high biological stability.[20]

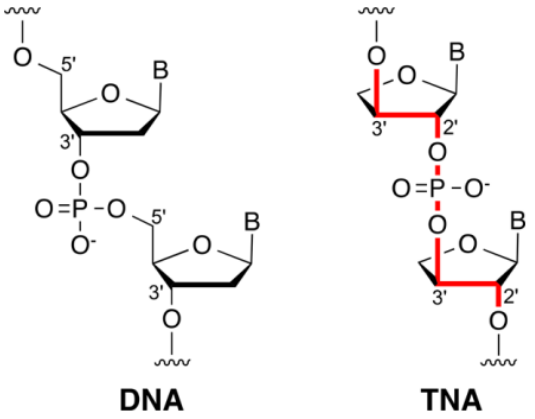


Figure 3-1. Constitutional structures for the linearized backbone of DNA (left) and α-L-threofuranosyl-(3′,2′) nucleic acid, TNA (right).

Previously, we have developed an engineered polymerase, termed Kod-RI, that is capable of transcribing DNA templates into TNA using standard buffer conditions that do not require manganese ions.[21] Unfortunately, subsequent studies revealed that Kod-RI functions with reduced transcriptional efficiency when individual templates are replaced with degenerate random DNA sequences libraries. In an effort to identify conditions that would allow the evolution of TNA libraries in vitro, we analyzed the impact of mutagenic manganese ions on the fidelity of TNA replication.[22] Here we show that the presence of manganese ions in the transcription buffer leads to an increase in tGTP misincorporation events observed as G-to-C transversions when the TNA strand is reverse transcribed back into DNA. By systematically analyzing the effects of different base analogues, we discovered that G-G mispairing in the enzyme active site occurs via a Hoogsteen base pair in which the incoming tGTP residue adopts a syn conformation with respect to the sugar. Substitution of tGTP for 7-deaza-7-phenyl tGTP enable to achieve >99% overall fidelity in the replication of TNA polymers. Finally, a TNA library containing the 7-deaza-7-phenyl guanine analogue was used to evolve a biologically stable TNA aptamer that binds to HIV reverse transcriptase with

a solution binding affinity constant (KD) of 1–2 nM and remains functional in the presence of a strong hydrolytic nuclease that rapidly degraded a control DNA aptamer in seconds to minutes. Our findings provide a mechanism for replicating TNA polymers in vitro with the full complement of genetic letters and suggest a strategy for producing chemically modified TNA molecules with enhanced ligand binding and catalytic activity.

3.4 Results

3.4.1 Fidelity Measurements with Natural and Modified Templates.

We measured the fidelity of TNA replication using an assay that involves transcribing a DNA template into TNA with Kod-RI TNA polymerase, separating the extended TNA product from the starting DNA template, reverse transcribing the TNA back into cDNA with Bst DNA polymerase, amplifying the cDNA by PCR reaction, cloning the DNA into E. coli, and sequencing the replicated DNA sequence (**Figure 3-2**).[23] This assay measures the aggregate fidelity of a complete replication cycle (DNA → TNA → DNA), which is operationally different than the more restricted view of fidelity as the accuracy of a single-nucleotide incorporation event. The fidelity determined by this assay demonstrates the accuracy with which full length TNA is synthesized and reverse transcribed, and therefore reflects the combined effects of nucleotide misincorporation, insertions and deletions (indel), and any mutations that occur during PCR amplification and cloning.

Several controls were implemented to ensure that the sequencing results reflect the true fidelity of TNA replication. First, to eliminate any possibility of contamination by the starting DNA template, the DNA primer-template complex used for TNA transcription was partially unpaired and contained additional nucleotides in the primer strand that added a unique primer binding site to the replicated product that was not present in the starting

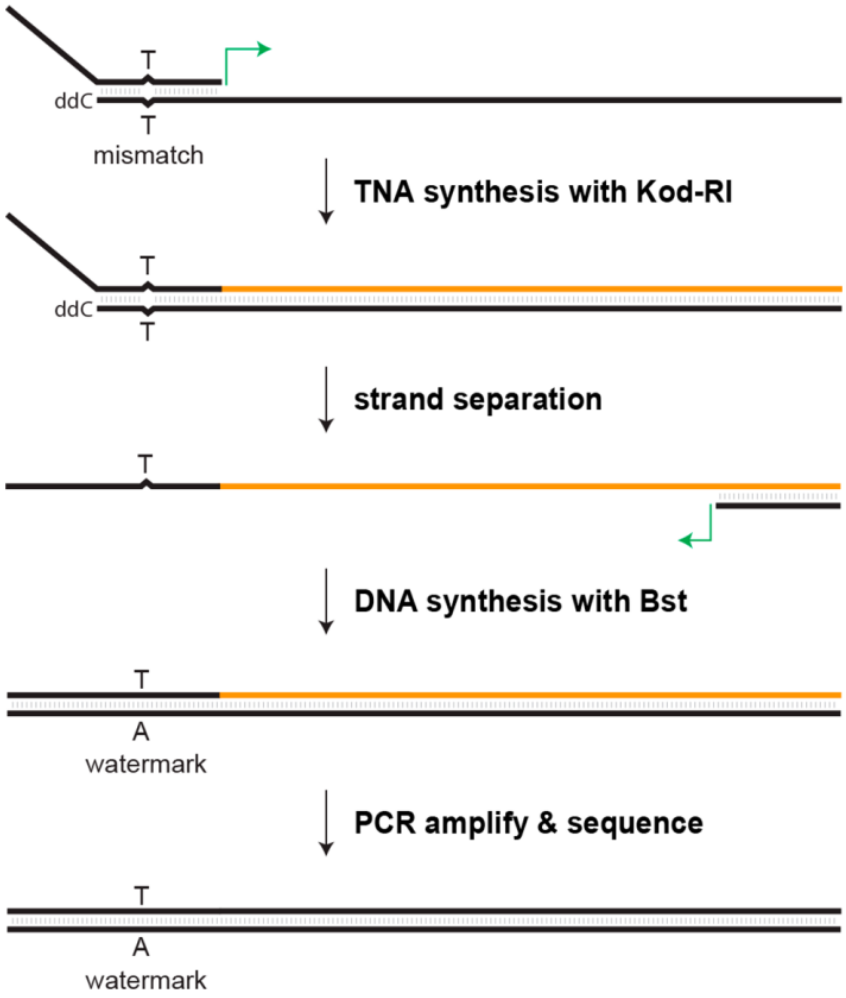


Figure 3-2. Replication strategy used to measure TNA fidelity.

A primer-template complex containing a T-T mismatch in the primer binding site is extended with TNA (orange). DNA is shown in black. The TNA strand is separated from the DNA template and reverse transcribed back into DNA. The cDNA is PCR amplified using a primer binding site downstream of the T-A watermark, cloned, and sequenced. DNA sequences carrying the T-A watermark confirm that the DNA sequence is the product of a TNA replication cycle. ddC refers to dideoxycytidine.

template. Second, the template strand was constructed with a 3′-dideoxycytidine residue to prevent elongation across the unpaired region on the primer strand. Third, all PCR amplification steps were performed with a no reverse transcription control to demonstrate that thepurified TNA product was free of the starting DNA template. Fourth, and most

importantly, the DNA primer used for TNA transcription was engineered to contain a single T-T mismatch that resulted in an A→T transversion when the TNA strand was reverse transcribed back into DNA. The mismatch was positioned inside the primer binding site to ensure that the primers used for PCR did not overlap with the mutation. Together, these controls allowed us to measure the fidelity of TNA replication with high confidence by analyzing the sequencing results obtained for different template and substrate combinations.

All of the fidelity measurements described in this study were performed by extending a DNA primer-template complex for 3 h at 55 °C with chemically synthesized TNA triphosphates (**Table 3-1**).[24,25] The transcription buffer was supplemented with 1 mM $MnCl_2$ to enable efficient TNA synthesis on DNA templates.[26] The primer-template complex and

Oligo Name	Sequence (5'-3' or 3'-2')
3NT8/10	CCATCTTCACTCCATTCAACTATCCTTAAACTCCCATTTACATCTCCAATACAC CTTAATCTCACTGGTGGTATCCCCTTGGGGA/3ddC/
4NT9G	GGATCGTCAGTGCATTGAGATTAAGACTCGCCATGTTACGATCTGCCAAGTA CAGCCTTGAATCGTCACTGGTGGTATCCCCT†AGGGGA/3ddC/
PBS8_extra	/56-FAM/CTTTTAAGAACCGGACGAACGTCCCCTTGGGGATACCACC
PBS7	GGATCGTCAGTGCATTGAGA
PEG-PBS7	AACAAACAAACAAACAAACAAACAAACAAACAAACAAACA/iSp18/ GGATCGT CAGTGCATTGAGA
Extra	CTTTTAAGAACCGGACGAAC
PEG-PBS8	AACAAACAAACAAACAAACAAACAAACAAACAAACAAACA/iSp18/GTCCCCTT GGGGATACCACC
PBS9	/5Hexynyl/CTTTTAAGAACCGGACGAAC
PBS10	CCATCTTCACTCCATTCAACTATCC
L16 library	GGATCGTCAGTGCATTGAGAATTCCTN$_{40}$(N=A:T:C:G=35:35:15:15)GTTCGT CCGGTTCTTAAAAG
3-2 template	GGATCGTCAGTGCATTGAGAATTCCTCGGCCAAGAATGATAACTGTGTTGGA ACTGCTTTTGATAGGTTCGTCCGGTTCTTAAAAGGGCAAAAAAAAAAA
R1T aptamer	CGCCTGATTAGCGATACTCAGGCGTTGGGTGGGTGGGTGGG

DNA sequences are written using IDT notations. †Underlined Ts form a T:T mismatch.

Table 3-1. DNA Sequences used in this study.

TNA polymerase were poised at concentrations of 1 μM each, and the reactions were initiated with the addition of 100 μM tNTP substrate. The TNA products were purified by denaturing PAGE and reverse transcribed back into cDNA using Bst DNA polymerase with standard manganese-free buffer conditions.[27] The cDNA was PCR amplified, cloned into E. coli, and sequenced. Statistical data were obtained by sequencing at least 1000 nucleotide positions for each template and nucleotide combination.

We began by measuring the fidelity of TNA synthesis for a DNA template that contained only three genetic letters (A, C, and T). Fidelity measurements indicate that this template copies with an overall fidelity of 99.7% (**Table 3-2**), which is high for nucleic acid polymers with modified sugar–phosphate backbones.[13] However, when the analogous experiment was performed on a DNA template that contained all four genetic letters (A, T, C, and G), the fidelity dropped to a modest 93% (**Table 3-2**). Close inspection of the sequencing results revealed a large abundance of G-to-C transversions, which were likely due to dG:tGTP mispairing during TNA synthesis with Kod-RI. Similar mutations have been observed with

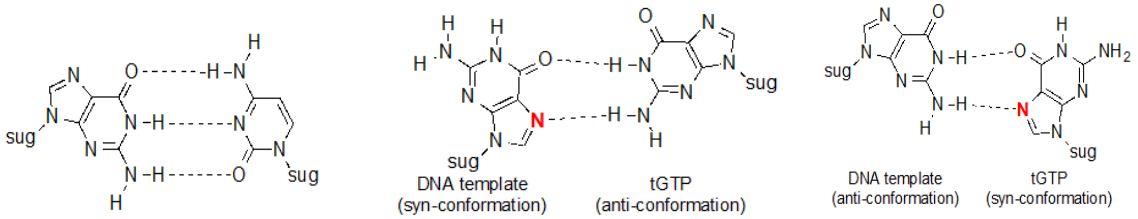
	DNA template design		
	3 nt	4 nt	7dG
transitions	–	2.5	3.3
transversions[b]	–	67	54.9
insertions	3	0.6	1.7
deletions	–	–	–
error rate	0.3×10^{-2}	7.0×10^{-2}	6.0×10^{-2}
fidelity	99.7%	93.0%	94.0%

Table 3-2. Fidelity of TNA Synthesis Using tNTPs with Natural Bases[a]

[a]Fidelity measurements of TNA synthesis were performed using DNA templates that contained three nucleotides (dA, dC, and dT), four nucleotides (dA, dC, dT, and dG), and four nucleotides with dG substituted for 7-deaza-2′-deoxyguanosine (7dG). Values were normalized to 1000 sequenced bases. [b]Most of the transversions (>90%) were G-to-C mutations in the DNA template.

Therminator DNA polymerase, which is a homologous polymerase that was previously used for TNA synthesis.[28] In contrast to natural G:C Watson–Crick base pairing (**Figure 3-3a**), Hoogsteen base pairing between guanine bases occurs when the guanine base on one side of the helix adopts a syn conformation relative to the sugar, while the other guanine base adopts the anti conformation (**Figure 3.3b**).

In this configuration, the G:G base pair is stabilized by two hydrogen bonds that form between the Watson–Crick face of one guanine residue and the major groove face of the other guanine base.[29] Crystallographic analysis of a DNA duplex containing an internal G:G base pair reveals that a G:G Hoogsteen base pair is structurally similar to a natural G:C Watson–Crick base pair.[30] Structural similarity of the G:C and G:G base pairs is thought to be one reason why polymerases struggle to read-through G-rich motifs in DNA templates.[30]



a. Canonical Watson-Crick base pair b. G-G Hoogsteen base pair (i) b. G-G Hoogsteen base pair (ii)

Figure 3-3. Molecular structures of (a) Watson–Crick and (b) Hoogsteen base pairs with guanine. The Hoogsteen base pair can occur via two different base-pairing modes depending on the glycosidic bond conformation (syn vs anti) of the guanine base in the templating and incoming nucleoside triphosphate. Chemical modifications made to the N-7 position (red) can be used to reveal which Hoogsteen base pairing mode predominates in the active site of a polymerase.

To test the hypothesis that G:G Hoogsteen base pairing was responsible for the increase in G-to-C transversions observed in the four-letter DNA template, we repeated the

fidelity experiment using a DNA template in which the natural guanine base (dG) was replaced with the unnatural base analogue 7-deazaguanine (7dG). The unnatural 7dG analogue contains a methine group in place of the guanine N-7 nitrogen that disrupts G:G base pairs when the templating base adopts the syn conformation and the incoming nucleotide adopts the anti-conformation (**Figure 3-3b**). Surprisingly, our fidelity analysis revealed that the 7dG-containing template replicates with the same fidelity (93% vs 94%) observed for the all-natural DNA template (**Table 3-2**). As with the natural DNA template, the 7dG template replicated with an abundance of G-to-C transversions, suggesting that Kod-RI promotes the formation of a dG:tGTP Hoogsteen base pair in which the templating base adopts the anti conformation and the incoming tGTP residue adopts the syn conformation (**Figure 3-3b**). This prediction is contrary to our previous analysis of Therminator DNA polymerase, where 7dG-containing templates were found to inhibit manganese-induced dG-tGTP mispairing in the enzyme active site.[28]

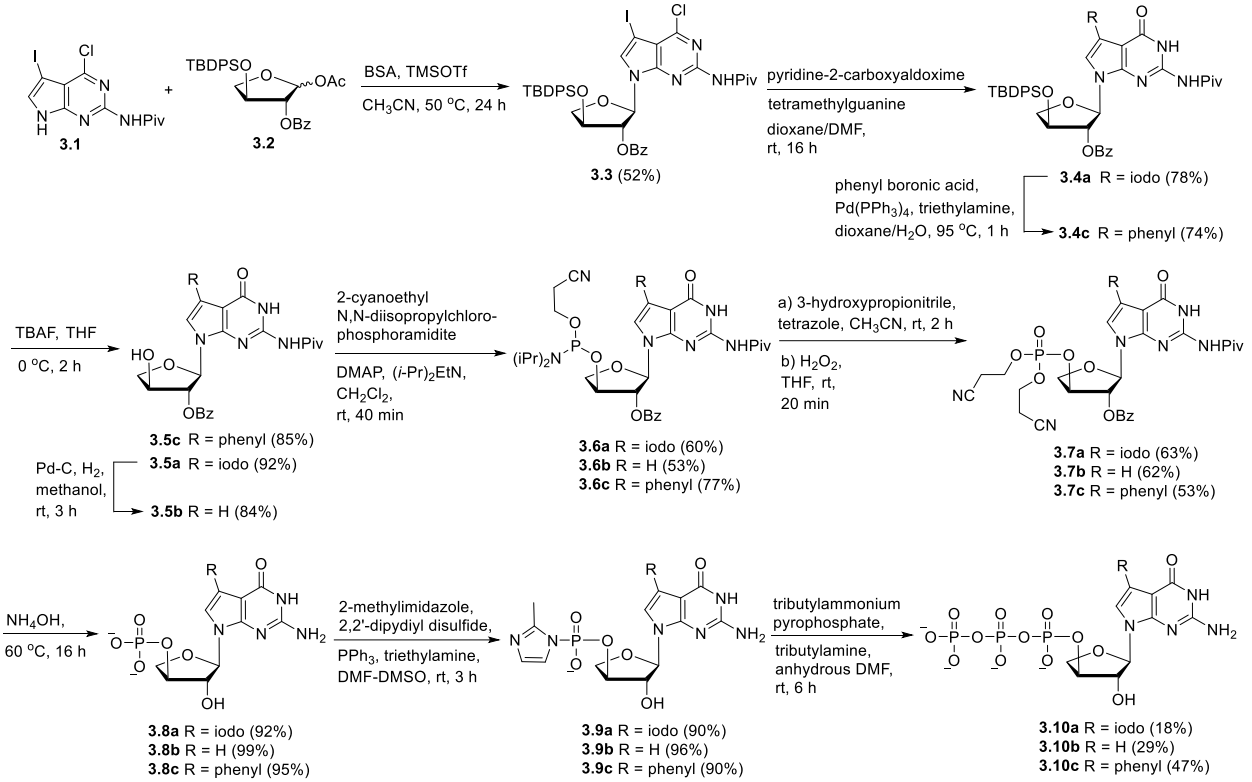
3.4.2 Synthesis of 7-Deaza-7-Modified tGTP Analogues.

To confirm that the observed G-to-C transversions were due to the formation of a dG:tGTP Hoogsteen base pair in which the incoming tGTP residue adopts a syn conformation, we chemically synthesized three different 7-deaza-7-modified tGTP analogues for testing in our fidelity assay (**Scheme 3-1**). Analogues **3.10a–c** are either unsubstituted at the N-7 nitrogen position (7-CH) or substituted with an iodo (7-I) or phenyl (7-Ph) group. These analogues allowed us to explore the fidelity and efficiency of polymerase-mediated synthesis of TNA polymers bearing diverse chemical functionality at the 7-position of the guanine nucleobase. We reasoned that faithful synthesis of chemically modified TNA polymers could

provide a new route to in vitro selected TNA molecules with enhanced chemical functionality, which is a strategy that has proven successful for many DNA aptamers.[31]

Our synthetic strategy utilized 6-chloro-7-iodo-nucleoside 3 as a synthon for 7-deazaguanine due to the limited activity of 7-deazaguanine as a nucleobase for glycosylation with the protected threose sugar.[24] Nucleoside 3 was obtained following a Vorbrüggen glycosylation in which 6-chloro-7-iodo-2-pivaloylamino-7-deazapurine (**3.1**) was conjugated to glycosyl donor **3.2**. In this reaction, the nucleobase **1.1** was first silylated with N,O-bis(trimethylsilyl)acetamide (BSA) and then treated with 2 equiv of 2 using TMSOTf as a catalyst. Subsequent treatment with pyridine-2-carboxaldoxime afforded the desired 7-deazaguanine ring structure observed in **3.4a**. A Suzuki–Miyaura cross-coupling reaction was then used to generate the 7-phenyl-substituted nucleoside **3.4c**. Key intermediates **3.5a** and **3.5c** for TNA triphosphate synthesis were obtained from **3.4a** and **3.4c**, respectively, by removal of the silyl protecting group with TBAF. Dehalogenation of **3.5a** with Pd/C and H_2 provided the methine (7-CH) derivative **3.5b**. Conversion of **3.5a–c** to the desired TNA triphosphates **3.10a–c** was performed using known methodology previously developed for TNA triphosphate synthesis.[32] Accordingly, 2′-O-benzoyl nucleosides **3.5a–c** were converted to their corresponding 3′-O-phosphoramidite derivatives **3.6a–c** upon treatment with 2-cyanoethoxy-N,N-diisopropyl chlorophosphine in the presence of Hunig's base and DMAP. Intermediates **3.6a–c** were subsequently converted to their corresponding trialkyl phosphite by replacing the N,N-diisopropylamino group with a base-labile cyanoethoxy group. The trialkyl phosphite was then oxidized with H_2O_2 to obtain the phosphate triesters **3.7a–c**. Deprotection of **3.7a–c** with 30% aqueous ammonium hydroxide afforded TNA nucleoside 3′-monophosphates **3.8a–c** as the ammonium salts. Next, TNA monophosphates

3.8a–c were converted to their 3′-phosphoro-(2-methyl)imidazolides **3.9a–c** with excess 2-methylimidazole and triphenylphosphine in the presence of 2,2′-dipyridyl disulfide. Finally, displacement of the imidazole residue in **3.9a–c** with pyrophosphate gave triphosphates **3.10a–c**. The desired triphosphates **3.10a–c** were obtained as highly pure sodium salts after HPLC purification, lyophilization, and precipitation as the sodium salt with sodium perchlorate.



Scheme 3-1. Synthesis of 7-Deaza-7-Modified Guanosine TNA Triphosphates

3.4.3 Fidelity Measurements with Modified tGTP Substrates.

To test the hypothesis that tGTP misincorporation was due to the formation of a dG:tGTP Hoogsteen base pair with the incoming tGTP nucleotide adopting a syn conformation, we measured the fidelity of TNA replication using tNTP mixtures that contained the 7-deaza-7-modified tGTP analogues **3.10a–c** in place of tGTP. In these

experiments, TNA polymers were synthesized on the four-letter DNA template composed of all natural bases. The 7-deaza-7-phenyl tGTP analogue **3.10c** functioned with the highest overall fidelity of replication (>99%), producing less than 1 G-to-C transversion per 1000 nucleotide incorporations (**Table 3-3**). The 7-deaza-7-methine tGTP analogue **3.10b** replicated with slightly lower fidelity (98% overall) due to a modest number of transitions and transversions. Surprisingly, the 7-deaza-7-iodo tGTP (7I-tGTP) analogue 10a functioned

	7-deaza-7-modified tGTP analogues		
	7H-tGTP	7I-tGTP	7-phenyl-tGTP
transitions[b]	7.1	51.1	4.3
transversions	10	2.2	0.9
insertions	–	4.4	–
deletions	2.9	–	3.5
error rate	2.0×10^{-2}	5.8×10^{-2}	8.7×10^{-3}
fidelity	98.0%	94.2%	99.1%

Table 3-3. Fidelity of TNA Synthesis Using tNTPs with Modified Bases[a]

[a]Fidelity measurements of TNA synthesis were performed using DNA templates that contained four nucleotides (dA, dC, dT, and dG) and tNTP substrates in which tGTP was substituted for the corresponding 7-deaza-7-modified-tGTP analogue. Values were normalized to 1000 sequenced bases. [b]Most of the transitions (>90%) observed with 7I-tGTP were T-to-C mutations in the DNA template. See **Table 3-4** for a detailed breakdown of specific transitions and transversions.

with the lowest overall fidelity (94%) due to a large number of T-to-C transitions. This result suggests that the 7I-tGTP analogue is prone to misincorporating opposite T in the DNA template, which leads to a rise in T-to-C transitions when the TNA strand is copied back into DNA. Despite the infidelity of 7I-tGTP, these results provide convincing evidence that the Kod-RI promotes tGTP misincorporation via a dG:tGTP Hoogsteen base pair in which the incoming tGTP nucleotide adopts a syn conformation.

3.4.4 Computational Analysis.

The basis for the unexpected observation of T-to-C transitions with 7I-tGTP was explored computationally, with the anomaly best explained by the formation of a Hoogsteen base pair between 7I-tGTP and T. Formally, such transitions could also result from wobble pairing of the 7I-G (keto form) with T or, alternatively, from Watson–Crick pairing via the enol tautomer of 7I-G. Wobble pairing is problematic due to its dissimilarity to the Watson–Crick geometry normally recognized by polymerases.[33] Watson–Crick pairing is excluded based on ab initio calculations of the keto (N^1-H) and enol (O^6-H) tautomer relative

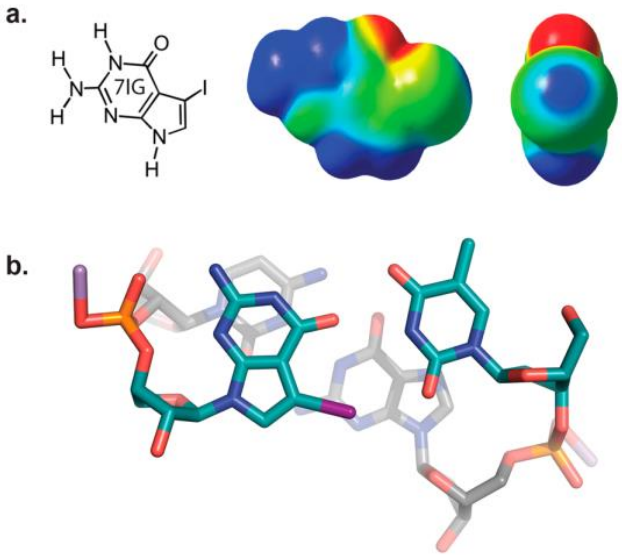


Figure 3-4. (a) Electrostatic potential energy surface for 7-deaza-7-iodo-guanine determined by ab initio calculations with Gaussian 09.

Blue color corresponds to regions of positive electrostatic potential, and red corresponds to negative potential. The image at right depicts the region of positive electrostatic potential corresponding to the sigma hole on iodine opposite the C–I bond axis. (b) Structure of the (t)7I-GC:(d)TG duplex derived from partial optimization at the B3LYP/MidiX level. This image highlights the non-covalent halogen bond formed by lone pair donation from the 2-oxygen of T to the sigma hole present on the 7-iodo group of G.

stabilities of 7I-G and G (**Table 3-4**), and the absence of any significant stabilization of the enol tautomer of 7I-G. Although the rationale for 7I-G was to impede Hoogsteen pairing of the G:G type, this need not carry over to other types of Hoogsteen pairing. Thus, while G:G

Hoogsteen pairing is anticipated and observed to be blocked by the iodo group of 7I-G due to both steric and electronic effects, the opposite is anticipated for 7I-G:T. Here the key underlying feature concerns the electro-positive "sigma-hole" found on halogen atoms, and the ability of these structural elements to interact with electronegative entities on other atoms or molecules, including lone pairs of electrons on heteroatoms, to form weak non-covalent bonding interactions.[34–36] In the present case, the electropositive nature of the iodo group terminus is displayed in **Figure 3-4a**. Partial optimization of a TNA:DNA dinucleotide duplex (**Figure 3-4b**) at the B3LYP/MidiX level resulted in a 7I-tG:dT Hoogsteen pair with an N–N distance of 9.02 Å (identical to B-DNA) and a 7-iodo---2-oxygen distance of 2.99 Å, consistent with known favorable halogen–oxygen electrostatic/bonding interaction distances. Of importance are an intermolecular hydrogen bond between 7I-G-O^6---T-N^1-H and an intramolecular hydrogen bond between the 7I-G-NH_2 and O of a nearby non-bridging phosphate. The latter is noteworthy as it may in some circumstances actively promote the syn conformation in tG residues.

3.4.5 In Vitro Selection.

The discovery that the 7-deaza-7-phenyl analogue of tGTP enables replication with high fidelity provided an opportunity to explore the functional properties of TNA by in vitro selection. To investigate this possibility, we constructed a degenerate library of ~1013 unique TNA molecules that contained a central random region of 40 nucleotide positions that was flanked on each side with fixed primer binding sites. The library was incubated with HIV reverse transcriptase (HIV RT), which was chosen as an arbitrary protein target for a model TNA selection. The selection was performed in a microtiter plate that was treated with 0.005% Tween-20 to reduce the occurrence of nonspecific binders.[37] Non-functional

molecules were removed from the pool using stringent wash conditions, and functional members were recovered, reverse transcribed back into DNA with Bst DNA polymerase, and amplified by droplet PCR. The PCR amplicons were amplified a second time in bulk solution using a PEG-modified DNA primer that allowed for strand separation by denaturing PAGE. The single-stranded DNA template was then forward transcribed back into TNA using Kod-RI to generate a new population of TNA molecules that had been enriched in HIV RT binding activity.

After three rounds of in vitro selection and amplification, the resulting pool was cloned and sequenced. Individual TNA molecules were constructed by primer extension and assayed for function in a standard electrophoretic mobility shift assay. The highest affinity aptamer identified from the selection was clone 3-2, which was found to bind HIV RT with a KD of 3 ± 0.6 nM (**Figure 3-5a**). This level of binding affinity is comparable to that of a known DNA aptamer (R1T),[38] which binds to HIVRT with a KD of 3 ± 2.1 nM (**Figure 3-5b**). Our binding affinity values for R1T are similar to the literature value of 14 ± 2 nM.[38] As expected, clone 3-2 remains functional and shows no signs of nuclease digestion after 24 h of incubation at 37 °C, while R1T rapidly degrades in the presence of snake venom phosphodiesterase (SVPE) (**Figure 3-5c, d**). This result, which is consistent with the known biological stability of TNA relative to DNA and RNA, highlights the benefit of TNA aptamers as a source of biologically stable protein capture reagents.[20]

To evaluate the functional role of the 7-phenyl groups on aptamer binding, a version of clone 3-2 was constructed in which tGTP was used in place of 7-deaza-7-phenyl-tGTP. The resulting TNA oligonucleotide exhibited 20-fold weaker binding affinity to the target protein

(Supplementary Figure 1) than the original aptamer (KD ≈ 60 nM). This result confirms the importance of the 7-phenyl group as a critical structural element for target binding.

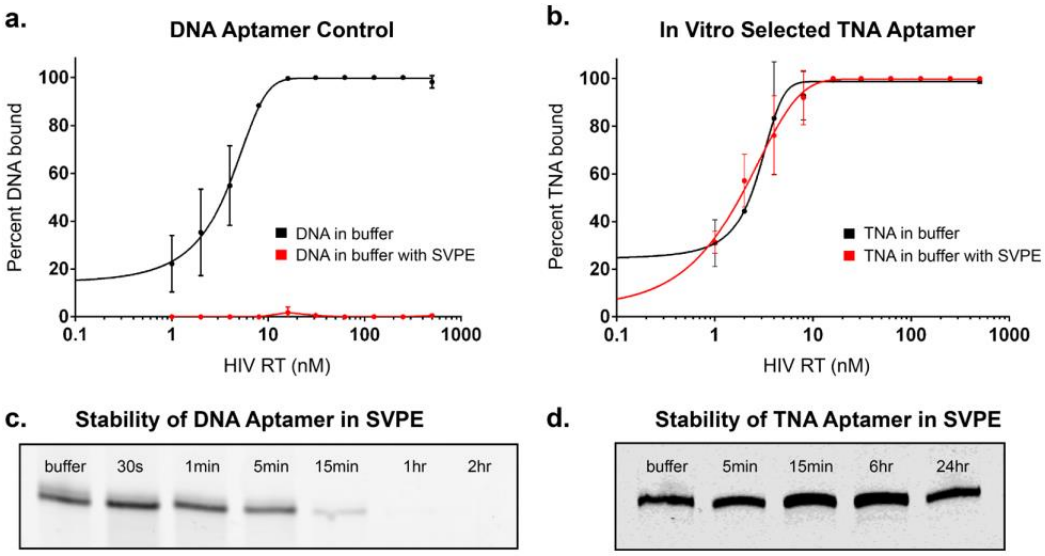


Figure 3-5. TNA aptamer binding and stability profiles.

Solution binding isotherms measured in the presence and absence of snake venom phosphodiesterase (SVPE) for DNA (a) and TNA (b) aptamers selected in vitro to bind HIV RT. Time-dependent SVPE stability profiles performed on the DNA (c) and TNA (d) aptamers.

3.5 Discussion.

We are interested in establishing a replication system that will allow for the evolution of TNA aptamers and catalysts in vitro. The critical barrier to our success has been the challenge of identifying polymerases that can transfer genetic information back and forth between DNA and TNA. All of the TNA polymerases discovered thus far function with inferior activity relative to natural DNA polymerases, many of which are able to copy entire genomes with high efficiency (>100 nt/s) and high fidelity (error rates of <10−6).[39] Thus, although our research is inspired by nature, new advances are needed to faithfully and efficiently replicate TNA polymers in a test tube.

Directed evolution approaches have revolutionized the field of synthetic genetics by providing the first examples of XNA polymerases that can recognize XNA substrates in the template and as nucleoside triphosphates.[12,40] However, these enzymes function with low catalytic efficiency and reduced fidelity when compared to their natural counterparts.[41] Efforts to overcome this problem include the elucidation of XNA polymerase structures, which provide a framework for designing new XNA polymerases,[42] and the establishment of ultra-high-throughput assays for rapidly screening polymerase variants in water-in-oil microcompartments.[21] Although these approaches are being used to accelerate the discovery of new XNA polymerases, chemical biology strategies that probe the underlying mechanism of nucleotide incorporation are often better suited for studying general problems associated with nucleotide recognition.[43]

The chemical biology strategy that we have taken has been to identify, synthesize, and test chemical analogues that could lead to enhancements in primer-extension efficiency or replication fidelity. Our early efforts in this area led to the discovery that certain commercial polymerases will incorporate tTTP residues more efficiently when the adenine base in the DNA template is replaced with diaminopurine.[26] More recently, we have found that Kod-RI will process G-rich templates with higher efficiency when tCTP substrates in the transcription mixture are replaced with a fluorescent tricyclic cytidine analogue.[44] These examples highlight the importance of using chemical analogues to overcome the inadequacies of engineered polymerases relative to natural polymerases.

The current study extends this line of research to include substrates that enable Kod-RI to faithfully synthesize TNA libraries containing all four genetic letters. By systematically analyzing the fidelity of TNA replication with natural and chemically modified bases in the

template and incoming nucleoside triphosphate, we discovered that Kod-RI-mediated tGTP misincorporation opposite dG in the DNA template occurs via a Hoogsteen base pair in which the incoming tGTP residue adopts a syn conformation with respect to the glycosidic linkage. Substitution of tGTP in the transcription mixture for the base analogue 7-deaza-7-phenyl tGTP enabled the replication of TNA polymers with >99% overall fidelity and led to the evolution of a biologically stable TNA aptamer that remains functional in the presence of SVPE.

During the course of our study, we learned that 7-deaza-7-iodo-tGTP has a strong propensity to misincorporate opposite T, which leads to a rise in T-to-C transitions when the TNA strand is copied back into DNA. This unexpected observation was rationalized using computational modeling, which reveals that 7I-tGTP forms a stable Hoogsteen base pair with thymine. The propensity for 7I-tGTP:T mispairing was attributed to the formation of an intramolecular hydrogen bond between the exocyclic amino group of 7I-tGTP and the oxygen atom of a nearby non-bridging phosphate group. We suggest that this interaction could contribute to the low fidelity of TNA replication in the presence of manganese ions by stabilizing the syn conformation of tGTP substrates in the enzyme active site. However, further studies are needed to examine this theory in greater detail.

3.6 Conclusion

In summary, we have established a faithful system for replicating TNA libraries in vitro and demonstrated the application of this system by in vitro selection. Given the growing interest in modified bases as a strategy to expand the chemical diversity of nucleic acid polymers relative to proteins,[45–48] we feel that 7-deaza-7-modified bases could lead to the discovery of new TNA aptamers and catalysts that function with superior activity relative to

analogous polymers with unmodified bases. We suggest that the ability to construct biologically stable affinity reagents will help advance the field of therapeutic aptamers by eliminating the need for costly post-SELEX modifications that are required to stabilize in vitro selected aptamers against DNA- and RNA-degrading enzymes.[49]

3.7 Experimental Details

General Methods and Materials

All non-aqueous reactions were performed using oven-dried glassware under an atmosphere of argon or nitrogen. All chemicals and solvents were purchased from Sigma-Aldrich and used without further purification. Reactions were monitored by thin layer chromatography using UV-activated TLC plates with silica gel 60 F254 and aluminium backing (Sigma-Aldrich, St. Louis, MO). Flash column chromatography was performed using SiliCycle 40-60 mesh silica gel (SiliCycle Inc., Quebec City, Canada). Yields are reported as isolated yields of pure compounds. UV quantification data are analyzed on NanoDrop 2000c using Beer's Law. ^{1}H, ^{13}C and ^{31}P NMR spectra were obtained using Bruker DRX400 and Bruker DRX500 NMR spectrometers (Bruker, Billerica, MA). ^{1}H and ^{13}C NMR values are reported in parts per million (ppm) relative to Me_4Si as internal standard or corresponding deuterium solvents as internal standard. ^{31}P NMR values are reported in ppm relative to an external standard of 85% H_3PO_4. Splitting patterns are designated as follows: s, singlet; br, broad; d, doublet; dd, doublet of doublets; t, triplet; q, quartet; m, multiplet. HPLC purification was performed with a C18 reverse-phase 250 x 21.2 mm HPLC column (Thermo Scientific, US) using a mobile phase of 0.1 M triethylammonium acetate buffer (pH 7.0)/acetonitrile. α-L-threofuranosyl nucleoside 3'-monophosphates (**3.8a-c**), 3'-phosphoro(2-methyl)imidazolides (**3.9a-c**), and 3'-triphosphates (**3.10a-c**) were analyzed

by analytical HPLC with a reverse-phase column (C18 150 × 4.6 mm, 5 μm particle size, Thermo Scientific, US) using a gradient of 0 to 15% acetonitrile in 0.1 M triethylammonium acetate buffer over 40 min. Thermo Pol buffer and Bst DNA polymerase 2.0 were purchased from New England Biolabs (Ipswich, MA). DNA oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA), purified by denaturing polyacrylamide gel electrophoresis, electro-eluted, precipitated by ethanol, resuspended in water, and quantified by UV absorbance. Recombinant Kod-RI polymerases was expressed and purified from E. coli as previously described.[48] TNA triphosphates bearing natural bases were synthesized as previously described.[32]

4-Chloro-5-iodo-7-(2'-*O*-benzoyl-3'-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl) -2-pivaloylamino-7H-pyrrolo[2,3-d]pyrimidine (3.3).

To a suspension of 4-chloro-5-iodo-2-pivaloylamino-7H-pyrrolo[2,3-d] pyrimidine (**3.2**)[49] (8.1 g, 21.4 mmol) in 150 mL anhydrous acetonitrile was added *N,O*-bis(trimethylsilyl)acetamide (5.87 mL, 23.5 mmol) and the mixture was stirred for 30 min at 50 °C. After cooling to 24 °C, TMSOTf (4.65 mL, 25.7 mmol) was added and 1-*O*-acetyl-2-*O*-benzoyl-3-*O*-tert-butyldiphenylsilyl-α-L-threofuranose (**3.1**) (26.7 g, 52.9 mmol) in 50 mL anhydrous acetonitrile was introduced in three portions (once per 8 h). In total, the reaction was stirred at 50 °C (oil bath) for 24 h, cooled to room temperature, and diluted with 200 mL of ethyl acetate. The solution was washed with aqueous saturated $NaHCO_3$ and brine, dried over $MgSO_4$. The solvent was evaporated under reduced pressure and the residue was purified by flash chromatography (silica gel, hexane/ethyl acetate, 10:1) to afford 3 (9.1 g, 52%) as a yellow foam. TLC (hexane/ethyl acetate, 3:1): Rf = 0.45. ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.07-8.06 (m, 1H), 8.02-8.01 (m, 1H), 7.96 (d, J = 8.0 Hz, 2H), 7.69-7.66 (m, 4H), 7.58-7.55

(m, 1H), 7.45-7.33 (m, 8H), 6.51 (m, 1H), 5.75 (s, 1H), 4.52 (s, 1H), 4.08-4.10 (m, 1H), 3.98-3.95 (m, 1H), 1.28 (s, 9H), 1.17-1.16 (m, 9H). ^{13}C NMR (125.8 MHz, $CDCl_3$): δ 175.3, 165.0, 153.0, 152.1, 151.6, 135.9, 135.8, 133.7, 132.7, 132.3, 131.9, 130.4, 130.3, 130.0, 128.9, 128.5, 128.2, 128.1, 113.7, 87.8, 83.1, 76.5, 75.0, 40.3, 27.5, 27.1, 19.3. HRMS (ESI-TOF) calcd for $C_{38}H_{40}N_4O_5SiClNaI$ $[M + Na]^+$ 845.1399; observed 845.1409.

5-Iodo-7-(2'-*O*-benzoyl-3'-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)-4-oxo-2-pivaloylamino-7H-pyrrolo[2,3-d]pyrimidine (3.4a).

A solution of compound **3.3** (5.75 g, 7.0 mmol), pyridine-2-carboxaldoxime (4.27 g, 35.0 mmol) and 1,1,3,3-tetramethylguanidine (4.39 mL, 35.0 mmol) in a mixture of dioxane (50 mL) and DMF (50 mL) was stirred for 16 h under argon atmosphere at room temperature. The reaction mixture was diluted with ethyl acetate (200 mL) and washed with 5% HCl (200 mL), saturated aqueous $NaHCO_3$ (200 mL), and brine (100 mL), dried over $MgSO_4$, and evaporated to dryness under reduced pressure. Purification by FC (silica gel, CH_2Cl_2/acetone, 10:1) gave **3.4a** as a light yellow solid (4.4 g, 78%). TLC (hexane/ethyl acetate, 2:1): Rf = 0.32. ^{1}H NMR (500 MHz, $CDCl_3$): δ 11.71 (s, 1H), 8.20 (s, 1H), 7.89 (d, J = 8.0 Hz, 2H), 7.68-7.65 (m, 4H), 7.62 (s, 1H), 7.56-7.54 (m, 1H), 7.45-7.33 (m, 8H), 6.18 (s, 1H), 5.64 (s, 1H), 4.49 (s, 1H), 4.07-3.98 (m, 2H), 1.24 (s, 9H), 1.17 (s, 9H). ^{13}C NMR (125.8 MHz, $CDCl_3$): δ 179.9, 164.8, 157.0, 151.2, 147.4, 146.6, 137.1, 135.8, 133.7, 132.6, 131.9, 130.3, 129.7, 128.8, 128.5, 128.1, 128.0, 127.0, 125.6, 105.5, 88.0, 83.4, 76.6, 75.2, 55.7, 40.2, 27.0, 19.2. HRMS (ESI-TOF) calcd for $C_{38}H_{41}N_4O_6SiNaI$ [M + Na]+ 827.1738; observed 827.1727.

7-(2'-*O*-Benzoyl-3'-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)-4-oxo-5-phenyl-2-pivaloylamino-7H-pyrrolo[2,3-d]pyrimidine (3.4c).

To a solution containing compound **3.4a** (1.6 g, 1.99 mmol), and phenylboronic acid (2.43 g, 19.9 mmol) in 40 mL of dioxane/H_2O (v/v, 2:1) was added triethylamine (2.77 mL, 19.9 mmol). The reaction was stirred for 1 h at 95 °C with monitoring by TLC. After complete consumption of starting material, the reaction was cooled to room temperature and was added 200 mL of ethyl acetate. The solution was sequentially washed with saturated aqueous $NaHCO_3$ (100 mL), water (100mL), brine (100 mL), and dried over $MgSO_4$. The organic layer was collected, evaporated and the crude was purified by flash column chromatography affording compound **3.4c** as a white solid (1.26 g, 74%). TLC (hexane/ethyl acetate, 3:1): Rf = 0.35. ^{1}H NMR (400 MHz, $CDCl_3$) δ 11.68 (s, 1H), 8.05 (s, 1H), 7.93 (d, J = 7.6 Hz, 2H), 7.87 (d, J = 8.4 Hz, 2H), 7.68-7.64 (m, 4H), 7.58-7.55 (m, 2H), 7.43-7.22 (m, 12H), 6.28 (s, 1H), 5.79 (s, 1H), 4.56 (s, 1H), 4.11-4.00 (m, 2H), 1.26 (s, 9H), 1.11 (s, 9H); ^{13}C NMR (125.8 MHz, $CDCl_3$) δ 179.6, 165.0, 157.6, 148.6, 146.3, 135.9, 135.8, 133.7, 133.4, 132.8, 132.1, 130.3, 130.3, 129.0, 128.6, 128.5, 128.2, 128.1, 128.1, 126.7, 122.9, 118.1, 102.7, 88.0, 83.7, 76.9, 74.9, 40.2, 27.1, 27.0, 19.3. HRMS (ESI-TOF) calcd for $C_{44}H_{46}N_4O_6SiNa$ $[M + Na]^+$ 777.3084; observed 777.3077.

5-Iodo-7-(2'-*O*-benzoyl-α-L-threofuranosyl)-4-oxo-2-pivaloylamino-7H-pyrrolo[2,3-d]pyrimidine (3.5a).

To an ice-cold (0-5 ºC) solution containing compound 4a (0.71 g, 0.88 mmol) in THF (30 mL) was added dropwise tetrabutylammonium fluoride (1 M solution in THF, 0.89 mL, 0.89 mmol) and the mixture was stirred for 2 h at 0°C. The solvent was evaporated under reduced pressure, and the residue was dissolved in 60 mL of CH_2Cl_2. The organic layer was washed twice with H_2O (50 mL) and brine (50 mL), dried over $MgSO_4$, and concentrated under reduced pressure to give a yellow syrup. The syrup was purified by flash

chromatography (silica gel, CH_2Cl_2/acetone, 10:1) to afford **3.5a** (0.46 g, 92%) as a light yellow solid. TLC (CH_2Cl_2/acetone, 9:1): Rf = 0.43. ^{1}H NMR (500 MHz, $CDCl_3$): δ 11.78 (s, 1H), 8.34 (s, 1H), 7.97 (d, J = 8.0 Hz, 2H), 7.57 (t, J = 7.5 Hz, 1H), 7.43-7.38 (m, 3H), 6.12-6.11 (m, 1H), 5.57 (s, 1H), 4.83-4.82 (m, 1H), 4.60 (s, 1H), 4.30-4.21 (m, 2H), 1.26 (s, 9H). ^{13}C NMR (125.8 MHz, $CDCl_3$): δ 180.0, 165.7, 157.1, 147.1, 146.4, 133.9, 129.8, 128.8, 128.7, 126.7, 105.7, 88.8, 83.8, 74.9, 40.3, 27.0. HRMS (ESI-TOF) calcd for $C_{22}H_{23}N_4O_6NaI$ $[M + Na]^+$ 589.0560; observed 589.0561.

7-(2'-*O*-Benzoyl-α-L-threofuranosyl)-4-oxo-2-pivaloylamino-7H-pyrrolo[2,3-d]pyrimidine (3.5b).

To a solution of **3.5a** (0.81 g, 1.43 mmol) in methanol (25 mL) was added 0.64 g Pd/C (palladium, 10% on activated carbon, 66% water wet). After being stirred at room temperature for 3 h under H_2 atmosphere, the mixture was filtered with celite and the filtrate was concentrated to dryness. The residue was purified by flash chromatography (silica gel, CH_2Cl_2/methanol, 20:1) to afford **3.5b** (0.53 g, 84%) as a white foam. TLC (CH_2Cl_2/acetone, 6:1): Rf = 0.29. ^{1}H NMR (500 MHz, $CDCl_3$): δ 11.90 (s, 1H), 8.27 (s, 1H), 8.01 (d, *J* = 7.5 Hz, 2H), 7.60 (t, *J* = 7.5 Hz, 1H), 7.45 (t, *J* = 7.5 Hz, 2H), 7.14 (d, *J* = 3.5 Hz, 1H), 6.64 (d, *J* = 3.5 Hz, 1H), 6.05 (d, *J* = 2.0 Hz, 1H), 5.60 (s, 1H), 5.36 (s, 1H), 4.55 (s, 1H), 4.27-4.19 (m, 2H), 1.29 (s, 9H). ^{13}C NMR (125.8 MHz, $CDCl_3$): δ 179.8, 165.8, 157.8, 146.7, 146.2, 133.9, 129.9, 128.9, 128.7, 122.3, 106.2, 103.9, 89.8, 84.1, 75.2, 74.6, 40.3, 27.1. HRMS (ESI-TOF) calcd for $C_{22}H_{24}N_4O_6Na$ $[M + Na]^+$ 463.1594; observed 463.1590.

7-(2'-*O*-Benzoyl-α-L-threofuranosyl)-4-oxo-5-phenyl-2-pivaloylamino-7H-pyrrolo[2,3-d]pyrimidine (3.5c).

To a cold solution (0 °C) containing compound **3.4c** (1.16 g, 1.44 mmol) in THF (14 mL) was added dropwise tetrabutylammonium fluoride (1 M solution in THF, 2.88 mL, 2.88 mmol) while stirring. After 1h, another 2.88 mL of tetrabutylammonium fluoride (1 M solution in THF, 2.88 mmol) was added to the solution and the mixture was stirred for additional 1 h at 0 °C. After complete consumption of starting material monitored by TLC, the solvent was evaporated under reduced pressure, and the residue was dissolved in 150 mL of ethyl acetate. The solution was washed with H_2O (100 mL x 2), brine (100 mL x 2), and dried over $MgSO_4$. The organic layer was collected, evaporated and purified by flash column chromatography affording **3.5c** as a white solid (0.63 g, 85%). TLC (hexane/ethyl acetate, 1:1) Rf = 0.45. ^{1}H NMR (400 MHz, $CDCl_3$) δ 11.78 (s, 1H), 8.28 (s, 1H), 7.98 (d, *J* = 7.6 Hz, 2H), 7.81 (d, *J* = 7.6 Hz, 1H), 7.56 (t, *J* = 7.2 Hz, 1H), 7.42-7.39 (m, 2H), 7.34-7.30 (m, 3H), 7.20 (t, *J* = 7.2 Hz, 1H), 6.14 (s, 1H), 5.61 (s, 1H), 4.86 (s, 1H), 4.49 (s, 1H), 4.23-4.15 (m, 2H), 1.23 (s, 9H); ^{13}C NMR (125.8 MHz, $CDCl_3$) δ 179.8, 165.9, 157.6, 147.9, 146.3, 133.9, 133.2, 129.8, 128.8, 128.7, 128.4, 128.2, 126.8, 122.3, 119.3, 103.1, 89.3, 84.1, 75.2, 74.6, 40.2, 27.0. HRMS (ESI-TOF) calcd for $C_{28}H_{28}N_4O_6Na$ $[M + Na]^+$ 539.1907; observed 539.1898.

5-Iodo-7-(2'-*O*-benzoyl-α-L-threofuranosyl)-4-oxo-2-pivaloylamino-7H-pyrrolo[2,3-d]pyrimidine 3'-(2-cyanoethyl)-*N,N'*-diisopropyl phosphoramidite (3.6a).

General Procedure for the Preparation of 6a-c.

To a stirring solution of **3.5a** (0.63 g, 1.1 mmol) and 4-dimethylaminopyridine (27 mg, 0.22 mmol, 0.2 equiv) in CH_2Cl_2 (10 mL) was added (*i*-Pr)$_2$NEt (383 µL, 2.2 mmol, 2.0 equiv) followed by the addition of 2-cyanoethyl-*N,N*-diisopropylchloro-phosphoramidite (446 µL, 2.0 mmol, 1.8 equiv). After being stirred for 40 min at 24 °C, the solution was diluted with CH_2Cl_2 (40 mL) and extracted with saturated aqueous $NaHCO_3$ (40 mL). The organic

layer was washed with brine, dried over Na_2SO_4 and evaporated. The residue was purified by flash chromatography (silica gel, CH_2Cl_2/acetone, 30:1) affording **3.6a** (0.51 g, 60%) as a light yellow foam. TLC (CH_2Cl_2/acetone, 9:1): Rf = 0.70. ^{31}P NMR (162 MHz, $CDCl_3$) δ 152.74, 151.43. HRMS (ESI-TOF) calcd for $C_{31}H_{40}N_6O_7PNaI$ $[M + Na]^+$ 789.1639; observed 789.1627.

7-(2'-*O*-Benzoyl-α-L-threofuranosyl)-4-oxo-2-pivaloylamino-7*H*-pyrrolo[2,3-d]pyrimidine 3'-(2-cyanoethyl)-*N,N'*-diisopropyl-phosphoramidite (3.6b).

As described for **3.6a**, compound **3.6b** was prepared from nucleoside **3.5b** (0.36 g, 0.82 mmol), purified by flash chromatography (silica gel, CH_2Cl_2/acetone, 15:1), and obtained as a yellowish foam (0.28 g, 53%). TLC (CH_2Cl_2/acetone, 6:1): Rf = 0.57. ^{31}P NMR (162 MHz, $CDCl_3$) δ 152.40, 151.09. HRMS (ESI-TOF) calcd for $C_{31}H_{41}N_6O_7PNa$ $[M + Na]^+$ 663.2672; observed 663.2674.

7-(2'-*O*-Benzoyl-α-L-threofuranosyl)-4-oxo-5-phenyl-2-pivaloylamino-7*H*-pyrrolo[2,3-d]pyrimidine 3'-(2-cyanoethyl)-*N,N'*-diisopropylphosphoramidite (3.6c).

As described for **3.6a**, compound **3.6c** was prepared from nucleoside **3.5c** (0.8 g, 1.55 mmol), purified by flash chromatography (silica gel, ethyl acetate/hexane from 25% to 50%), and obtained as a white solid (0.85 g, 77%). TLC (hexane/ethyl acetate, 2:1): Rf = 0.2. ^{31}P NMR (162 MHz, $CDCl_3$) δ 152.89, 151.22; HRMS (ESI-TOF) calcd for $C_{37}H_{45}N_6O_7PNa$ $[M + Na]^+$ 739.2985; observed 739.2989.

5-Iodo-7-(2'-*O*-benzoyl-α-L-threofuranosyl)-4-oxo-2-pivaloylamino-7*H*-pyrrolo[2,3-d]pyrimidine 3'-bis(2-cyanoethyl)-phosphotriester (3.7a).

General Procedure for the Preparation of **3.7a-c**. To a stirring solution of **3.6a** (0.50 g, 0.65 mmol) in acetonitrile (15 mL) was added 3-hydroxypropionitrile (67 μL, 0.98 mmol,

1.5 equiv) followed by a solution of 0.45 M tetrazole in THF (2.2 mL, 0.98 mmol, 1.5 equiv). After being stirred for 2 h at 24 °C, the solution was diluted with CH_2Cl_2 (50 mL) and washed with brine. The organic layer was dried over Na_2SO_4 and evaporated to afford a yellow foam. This material was directly used for the next step. The foam was dissolved in 6 mL of THF and H_2O_2 (30% in H_2O) (133 μL, 1.3 mmol, 2.0 equiv) was added to the solution. After being stirred at 24 °C for 20 min, the solution was diluted with CH_2Cl_2 (50 mL) and washed with brine. The organic layer was dried over Na_2SO_4 and evaporated. The residue was purified by flash chromatography (silica gel, CH_2Cl_2/acetone/methanol, 50:10:1) to afford **3.7a** (0.36 g, 63%) as a light yellow foam. TLC (CH2Cl2/acetone, 9:1): Rf = 0.26. ^{1}H NMR (500 MHz, $CDCl_3$): δ 11.84 (s, 1H), 8.40 (s, 1H), 8.05 (d, J = 7.5 Hz, 2H), 7.64 (t, J = 7.5 Hz, 1H), 7.50 (t, J = 7.5 Hz, 2H), 7.11 (s, 1H), 6.16 (s, 1H), 5.99 (s, 1H), 5.15-5.14 (m, 1H), 4.54 (d, J = 11.0 Hz, 1H), 4.37-4.30 (m, 5H), 2.86-2.81 (m, 4H), 1.30 (s, 9H). ^{13}C NMR (125.8 MHz, $CDCl_3$): δ 180.1, 165.1, 156.9, 147.2, 147.0, 134.3, 130.0, 128.9, 128.4, 124.9, 116.5, 106.1, 88.9, 80.5 (d, $J_{C,P}$ = 6.3 Hz), 80.0 (d, $J_{C,P}$ = 5.0 Hz), 73.1 (d, $J_{C,P}$ = 5.0 Hz), 63.1 (t, $J_{C,P}$ = 6.3 Hz), 55.6, 40.3, 27.0, 19.9 (dd, $J_{C,P}$ = 5.0 Hz, $J_{C,P}$ = 7.5 Hz). ^{31}P NMR (162 MHz, $CDCl_3$): δ -2.98. HRMS (ESI-TOF) calcd for $C_{28}H_{30}N_6O_9PNaI$ [M + Na]+ 775.0754; observed 775.0753.

7-(2'-*O*-Benzoyl-a-L-threofuranosyl)-4-oxo-2-pivaloylamino-7H-pyrrolo[2,3-d]pyrimidine 3'-bis(2-cyanoethyl)-phosphotriester (3.7b).

As described for **3.7a**, compound **3.7b** was prepared from nucleoside **3.6b** (0.24 g, 0.38 mmol), purified by flash chromatography (silica gel, CH_2Cl_2/methanol, 30:1), and obtained as a white foam (0.19 g, 62%). TLC (CH_2Cl_2/acetone, 6:1): Rf = 0.15. ^{1}H NMR (500 MHz, $CDCl_3$): δ 11.93 (s, 1H), 8.57 (s, 1H), 8.05-8.04 (m, 2H), 7.65-7.62 (m, 1H), 7.51-7.47 (m, 2H), 7.02-7.01 (m, 1H), 6.71-6.69 (m, 1H), 6.20-6.19 (m, 1H), 6.04-6.03 (m, 1H), 5.16-5.15

(m, 1H), 4.51 (d, J = 11.0 Hz, 1H), 4.35-4.32 (m, 5H), 2.81-2.80 (m, 4H), 1.30 (s, 9H). ^{13}C NMR (125.8 MHz, $CDCl_3$): δ 180.1, 165.2, 157.7, 147.6, 146.7, 134.2, 130.0, 128.8, 128.4, 120.2, 116.5, 105.9, 104.4, 88.6, 88.5, 80.7 (d, $J_{C,P}$ = 6.3 Hz), 80.4 (d, $J_{C,P}$ = 6.3 Hz), 72.6 (d, $J_{C,P}$ = 6.3 Hz), 62.9 (d, $J_{C,P}$ = 6.3 Hz), 62.8 (d, $J_{C,P}$ = 6.3 Hz), 40.3, 27.0, 19.8 (d, $J_{C,P}$ = 5.0 Hz). ^{31}P NMR (162 MHz, $CDCl_3$): δ -2.92. HRMS (ESI-TOF) calcd for $C_{28}H_{31}N_6O_9PNa$ $[M + Na]^+$ 649.1788; observed 649.1793.

7-(2'-*O*-Benzoyl-α-L-threofuranosyl)-oxo-5-phenyl-2-pivaloylamino-7*H*-pyrrolo[2,3-d]pyrimidine 3'-bis(2-cyanoethyl)-phosphotriester (3.7c).

As described for **3.7a**, compound **3.7c** was prepared from nucleoside **3.6c** (0.8 g mg, 1.12 mmol), purified by flash chromatography (silica gel, ethyl acetate/hexane from 33% to 75%) and obtained as a white foam (0.42 g, 53%). TLC (hexane/ethyl acetate, 1:2): Rf = 0.45. ^{1}H NMR (400 MHz, $CDCl_3$) δ 11.95 (s, 1H), 8.86 (s, 1H), 8.07 (d, *J* = 7.6 Hz, 1H), 7.93 (d, *J* = 7.2 Hz, 1H), 7.65 (t, *J* = 7.2 Hz, 1H), 7.50-7.42 (m, 4H), 7.31 (t, *J* = 7.2 Hz, 1H), 6.32 (s, 1H), 6.06 (s, 1H), 5.20 (s, 1H), 4.61 (d, *J* = 11.2 Hz, 1H), 4.40-4.26 (m, 5H), 2.76-2.26 (m, 4H), 1.32 (s, 9H); ^{13}C NMR (125.8 MHz, $CDCl_3$) δ 180.2, 165.0, 157.5, 148.5, 146.7, 134.0, 133.2, 129.8, 128.6, 128.3, 128.2, 128.1, 126.7, 122.2, 117.5, 116.5, 102.7, 88.3, 80.6 (d, $J_{C,P}$ = 6.5 Hz), 80.0 (d, $J_{C,P}$ = 6.5 Hz), 72.7 (d, $J_{C,P}$ = 6.5 Hz), 62.8 (d, $J_{C,P}$ = 6.5 Hz), 40.1, 26.8, 19.4 (d, $J_{C,P}$ = 9.7 Hz), 19.4 (d, $J_{C,P}$ = 9.7 Hz). ^{31}P NMR (162 MHz, $CDCl_3$) δ -2.61. HRMS (ESI-TOF) calcd for $C_{34}HN_6O_9PNa$ $[M + Na]^+$ 725.2101; observed 725.2092.

2-Amino-5-iodo-7-(α-L-threofuranosyl)-3,7-dihydro-pyrrolo[2,3-d]pyrimidin-4-one 3'-monophosphate (3.8a).

General Procedure for the Preparation of 8a-c.

In a sealed tube, compound **3.7a** (250 mg, 0.33 mmol) was combined with NH_3-methanol (5 mL) and NH_4OH (10 mL), and stirred for 16 h at 60 °C. The solution was cooled down to room temperature, diluted with water and washed with CH_2Cl_2 (3 x 40 mL). The aqueous layer was lyophilized to afford the product as ammonium salt. The precipitate was collected by centrifugation at 4400 rpm at room temperature for 15 min and the resulting pellet was washed twice with 30 mL of acetone and dried under high vacuum. The nucleoside 3'-monophosphate **3.8a** was obtained as the ammonium salt (light yellow solid) in near quantitative yield (150 mg, 92%). ^{1}H NMR (400 MHz, D_2O): δ 7.20 (s, 1H), 5.95 (s, 1H), 4.71-4.67 (m, 2H), 4.29 (s, 2H). ^{31}P NMR (162 MHz, D_2O): δ 0.76. HRMS (ESI-TOF) calcd for $C_{10}H_{12}N_4O_7PNaI$ $[M + Na]^+$ 480.9386; observed 480.9379.

2-Amino-7-(α-L-threofuranosyl)-3,7-dihydro-pyrrolo[2,3-d]pyrimidin-4-one 3'-monophosphate (3.8b).

As described for **3.8a**, compound **3.8b** was prepared from nucleoside **3.7b** (195 mg, 0.24 mmol) and obtained as the ammonium salt (white solid, 86 mg, 99%). ^{1}H NMR (400 MHz, D_2O): δ 7.13 (d, J = 4.0 Hz, 1H), 6.49 (d, J = 3.6 Hz, 1H), 5.98 (m, 1H), 4.73-4.69 (m, 2H), 4.26 (s, 2H). ^{31}P NMR (162 MHz, D_2O): δ 1.99. HRMS (ESI-TOF) calcd for $C_{10}H_{13}N_4O_7PNa$ [M + Na]+ 355.0420; observed 355.0422.

2-Amino-5-phenyl-7-(α-L-threofuranosyl)-3,7-dihydro-pyrrolo[2,3-d]pyrimidin-4-one 3'-monophosphate (3.8c).

As described for **3.8a**, compound **3.8c** was prepared from nucleoside **3.7c** (300 mg, 0.43 mmol) and obtained as the ammonium salt (white solid, 165 mg, 95%). ^{1}H NMR (400 MHz, DMSO-d_6) δ 7.88 (s, 2H), 7.31-7.14 (m, 4H), 5.91 (s, 1H), 4.52-4.43 (m, 2H), 4.07-4.00

(m, 2H). ^{31}P NMR (162 MHz, D_2O): δ 0.10. HRMS (ESI-TOF) calcd for $C_{16}H_{16}N_4O_7PNa_2$ [M - H + 2Na]$^+$ 453.0552; observed 453.0564.

2-Amino-5-iodo-7-(α-L-threofuranosyl)-3,7-dihydro-pyrrolo[2,3-d]pyrimidin-4-one 3'- phosphor-(2-methyl)imidazolide (3.9a).

General procedure for the preparation of 9a-c.

To a stirring solution containing nucleoside 3'-monophosphate **3.8a** (100 mg, 0.20 mmol) and 2-methylimidazole (300 mg, 3.65 mmol, 18.3 equiv) in anhydrous DMSO (1.5 mL) and DMF (1.5 mL), and was added triethylamine (0.2 mL, 1.44 mmol, 7.2 equiv), triphenylphosphine (300 mg, 1.14 mmol, 5.7 equiv) and 2,2'-dipyridyldisulfide (360 mg, 1.63 mmol, 8.2 equiv). The reaction was stirred under a nitrogen atmosphere for 3 h at room temperature with monitoring by analytical HPLC. After consumption of the starting material, the product was precipitated by the dropwise addition of the reaction mixture to a stirring solution containing 80 mL of acetone, 60 mL of diethyl ether, 5 mL of triethylamine and 2 mL of saturated $NaClO_4$ in acetone. The precipitate was collected by centrifugation at 4400 rpm for 15 min at room temperature. The pellet was washed twice with 30 mL of washing solution (acetone/diethyl ether 1:1) and dried under high vacuum to afford **3.8a** as sodium salt (white solid, 100 mg, 90%). It was used for next step directly without further purification. ^{31}P NMR (162 MHz, DMSO-d_6): δ -9.49. HRMS (ESI-TOF) calcd for $C_{14}H_{15}N_6O$-PNa_2I [M -H + 2Na]$^+$ 566.9631; observed 566.9639.

2-Amino-7-(α-L-threofuranosyl)-3,7-dihydro-pyrrolo[2,3-d]pyrimidin-4-one-3'-phosphor-(2-methyl)imidazolide (3.9b).

As described for **3.9a**, compound **3.9b** was prepared from nucleoside **3.8b** (110 mg, 0.24 mmol), and obtained as sodium salt (white solid, 121 mg, 96%). ^{31}P NMR (162 MHz, DMSO-d_6): δ -9.48. HRMS (ESI-TOF) calcd for $C_{14}H_{16}N_6O_6PNa_2$ [M -H + 2Na]$^+$ 441.0664; observed 441.0670.

2-Amino-5-phenyl-7-(a-L-threofuranosyl)-3,7-dihydro-pyrrolo[2,3-d]pyrimidin-4-one 3'-phosphor-(2-methyl)imidazolide (3.9c).

As described for **3.9a**, compound **3.9c** was prepared from nucleoside **3.8c** (60 mg, 0.148 mmol) and obtained as sodium salt (white solid, 63 mg, 90%). ^{31}P NMR (162 MHz, D_2O) δ -9.31; HRMS (ESI-TOF) calcd for $C_{20}H_{21}N_6O_6PNa$ [M - Na]$^+$ 495.1158; observed 495.1157.

2-Amino-5-iodo-7-(α-L-threofuranosyl)-3,7-dihydro-pyrrolo[2,3-d]pyrimidin-4-one 3'-triphosphate (3.10a).

General procedure for the preparation of 10a-c.

To a anhydrous DMF (2 mL) solution of compound 9a (24 mg, 44 μmol) was added tributylamine (27 μL, 0.11 mmol, 2.5 equiv) and tributylammonium pyrophosphate (62 mg, 0.11 mmol, 2.5 equiv), and the reaction was stirred for 6 h at 24 °C under N_2 with monitoring by analytical HPLC. After the reaction was finished, the reaction mixture was added dropwise to a stirring solution containing 30 mL of acetone and 5 mL of saturated $NaClO_4$ in acetone. The precipitate was collected by centrifugation at 4400 rpm for 15 min at room temperature and dried under vacuum for 1 h. The crude precipitate was dissolved in 2 ml of 0.1 M triethylammonium acetate buffer and purified by a semi-preparative HPLC. Fractions containing triphosphates were collected, concentrated, pH adjusted by triethylamine to 8.0,

and lyophilized to afford the product as a triethylammonium salt. The solid product was resuspended in 1 mL of methanol and was added dropwise to a solution containing 30 mL of acetone and 1 mL of saturated $NaClO_4$ in acetone. The solution was centrifuged at 4400 rpm for 15 min at room temperature. The supernatant was discarded, and the pellet was washed with 30 mL acetone and dried under vacuum. The desired triphosphate **3.10a** was obtained as sodium salt in a white solid form. The yield was measured by the NanoDrop 2000c, assuming an extinction coefficient[50] of 11.0 mM^{-1} cm^{-1} and found to be 18% (7.8 μmol). ^{31}P NMR (162 MHz, D_2O): δ –4.74 (d, J = 20.1 Hz), −10.65 (d, J = 20.3 Hz), −20.69 (t, J = 20.4 Hz). HRMS (ESI-TOF) calcd for $C_{10}H_{11}N_4O_{13}P_3Na_4I$ $[M -3H + 4Na]^+$ 706.8171; observed 706.8180.

2-Amino-7-(a-L-threofuranosyl)-3,7-dihydro-pyrrolo[2,3-d]pyrimidin-4-one-3'-triphosphate (**3.10b**).

As described for **3.10a**, compound **3.10b** was prepared from nucleoside **3.9b** (24 mg, 57 μmol), and obtained as sodium salt in a white solid form. The yield was measured by the NanoDrop 2000c, assuming an extinction coefficient[51] of 13.0 mM^{-1} cm^{-1} and found to be 29% (16.5 μmol). ^{31}P NMR (162 MHz, D_2O): δ –4.46 (d, J = 21.5 Hz), −10.55 (d, J = 20.4 Hz), −20.38 (t, J = 21.1 Hz). HRMS (ESI-TOF) calcd for $C_{10}H_{12}N_4O_{13}P_3Na_4$ $[M - 3H + 4Na]^+$ 580.9205; observed 580.9217.

2-Amino-5-phenyl-7-(α-L-threofuranosyl)-3,7-dihydro-pyrrolo[2,3-d]pyrimidin-4-one 3'-triphosphate (3.10c).

As described for **3.10a**, compound **3.10c** was prepared from nucleoside **3.9c** (20 mg, 0.042 mmol), and obtained as sodium salt in a white solid form. The yield was measured by

the NanoDrop 2000c, assuming an extinction coefficient[51] of 13.4 mM-1 cm-1 and found to be 47% (19.7 μmol). ^{31}P NMR (162 MHz, D_2O): δ –4.54 (d, J = 20.9 Hz), −10.51 (d, J = 19.9 Hz), -20.00 (t, J = 21.1 Hz); HRMS (ESI-TOF) calcd for $C_{16}H_{16}N_4O_{13}P_3Na_4$ $[M - 3H + 4Na]^+$ 656.9518; observed 656.9501.

TNA library preparation

For the aptamer selection, a DNA library (L16) was purchased from IDT that contained a 40-mer random region (N=A:T:C:G=35:35:15:15) that was flanked on both sides by fixed primer-binding sites that were compatible with PBS7 and PBS9 forward and reverse DNA primer sequences. The DNA library was PAGE purified, electroeluted, desalted and UV quantified. The synthetic library (1 μM) was transcribed into TNA by extending an IR800-labeled PBS9 primer annealed to the DNA library with Kod-RI TNA polymerase (1 μM) and tNTP substrates for 3 hours at 55°C. The tNTP substrates were poised at 100 μM concentration, except for 7-deaza-7-phenyl-tG, which was poised at 75 μM. The pool of TNA molecules was purified by denaturing PAGE, electro-eluted, desalted using Millipore YM-10 microcentrifugal concentrators, and quantified by UV spectroscopy.

In vitro selection

The HIV-RT protein (100 μL of 80 nM) in 100 mM NaH_2PO_4/Na_2HPO_4 buffer (pH 7.5) was immobilized onto one well of a 96-well Nunc Immobilizer Amino plate plate (ThermoFisher) at 4°C overnight. The target-coated well was washed three times with 300 μl of 1x AF buffer (20 mM Tris-HCl pH 7.4, 140 mM KCl, 5 mM NaCl, 1 mM $MgCl_2$, 1 mM $CaCl_2$, 0.005% v/v Tween 20). Next, the well was equilibrated in 300 μL 1x AF buffer at 4°C for 1

hour, which was removed right before the application of the TNA library. The ssTNA library (~150 pmole = 1013 molecules) was pre-folded in 100 µL 1x AF buffer by heating to 75°C for 5 min and cooling to 25°C in a thermocycler at 5°C/min. Upon removal of the equilibrating 1x AF buffer, the pre-folded ssTNA library was applied to the HIV-RT coated well, and incubated at room temperature for 1 hour. After incubation, the unbound library was removed and the well was sequentially washed three times with 300 µL 1x AF buffer and five times with 300 µL of stringent buffer (20 mM Tris-HCl pH 7.4, 4 M KCl, 5 mM NaCl, 1 mM $MgCl_2$, 1 mM $CaCl_2$, 0.005% v/v Tween 20). The bound aptamers were eluted three times by incubating at 70°C for 10 min with 200 µl of denaturing buffer (20 mM Tris-HCl pH 7.4, 3.5 M urea) pre-heated to 90°C. The eluate was desalted by exchanging into water (~10x volume of denaturing buffer = 6 mL), and concentrated using Millipore YM-30 microcentrifuge concentrators. The eluate was then reverse transcribed into cDNA using Bst-BF in 20 µL of 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, NEB) containing 1 µM of reverse transcription primer PBS7, 500 µM of dNTPs, and 3 mM of supplemented $MgCl_2$ at 50°C for 3 h. The reaction was treated with 0.8 U proteinase K at 50°C for 20 minutes, then held at 95°C for 10 minutes. The cDNA was amplified by PCR using Taq DNA polymerase with PBS7/PBS9 primer pair: 95°C for 5 min, N cycles of (95°C for 15 s; 58°C for 15 s; 72°C for 30 s). The number of cycles (N) was optimized for each round of selection by sampling PCR reactions every other cycle up to 20 cycles. The amplified DNA was used as template for a second PCR reaction in which PBS7 was replaced with PEGylated PBS7 (PEG-PBS7). The second PCR product was purified by denaturing PAGE, and the PEGylated strand was used as template for the next round of selection. After 3 rounds of selection, the population of DNA sequences was cloned and

sequenced. Native gel shift assays identified clone 3-2 as the highest affinity aptamer TNA sequence:

3'-CTTTTAAGAACCGGACGAACCTATCAAAAGCAGTTCCAACACAGTTATCATTCTTGGCCGAGG AATTCTCAATGCACTGACGATCC-2'

Nuclease stability assay

The all-TNA aptamer 3-2 was synthesized by primer-extension using a TNA version of the IR680-PBS9 primer as described above. Aptamers were challenged with nuclease in 75 µL solutions, prepared with 40 nM of IR680-labeled TNA 3-2 or 160 nM of FAM-labeled control DNA R1T aptamer mixed with 6.25 U/L Phosphodiesterase I from Crotalus adamanteus venom (SVPE, Sigma-Aldrich) in 40 mM Tris pH 8.4 and 10 mM $MgCl_2$. 10 µL aliquots were incubated at 37°C, quenched with stop buffer (25 mM EDTA in 95% formamide) at specified times, and analyzed by denaturing PAGE.

Kd measurement assay

Solution binding affinity constants (KD) were performed using a native gel electrophoretic mobility shift assay (EMSA) in the presence and absence of SVPE. For each KD measurement, 10 tubes of HIV-RT were prepared that spanned a concentration gradient of 1 µM to 2 nM protein in 6 µL of AF buffer. Separately, 75 µL solutions of 40 nM aptamer (TNA 3-2 or DNA R1T, with the TNA strand prepared as described for the nuclease stability assay) in 40 mM Tris pH 8.4 and 10 mM $MgCl_2$ were prepared in duplicate. SVPE (6.25 U/L) was added to the TNA and DNA aptamer solutions used for the nuclease challenge, while the second aptamer set was used to measure solution binding in the absence of SVPE (buffer

only). The aptamer solutions were incubated at 37°C for 15 minutes, and aliquoted (6 µL) into the tubes containing HIV RT. The solutions were incubated for 1 hour at room temperature, suspended in loading buffer (40 mM Tris pH 7.4, 600 mM sucrose), and analyzed by native PAGE. KD values were determined by quantifying the ratio of bound and unbound complex at each protein concentration.

	Template			7-deaza TNA triphosphates		
Mutation	3nt	4nt	7dG	tGTP (7-H)	tGTP (7-I)	tGTP (7-ph)
Transitions						
T -> C	-	0.6	2.5	5.7	48.9	4.3
C -> T	-	0.6	0.8	-	-	-
A -> G	-	-	-	1.4	-	-
G -> A	-	1.3	-	-	2.2	-
Transversions						
G -> C	-	63.9	52.5	10	-	-
G -> T	-	-	0.8	-	-	0.9
A -> T	-	-	0.8	-	-	-
T -> G	-	-	0.8	-	-	-
T -> A	-	1.9	-	-	2.2	-
A -> C	-	0.6	-	-	-	-
C -> G	-	0.6	-	-	-	-

Table 3-4. Detailed breakdown of fidelity results.

TNA fidelity assay

TNA fidelity measurements were performed using DNA templates that contained: (1) 3 nucleotides (dA, dC, and dT, denoted 3NT8/10), (2) 4 nucleotides (dA, dC, dT, and dG, enoted 4NT9G), and (3) 4 nucleotides with dG substituted for 7-deaza-7H-guanosine. In the case of (3), the 4NT9G DNA template was first PCR amplified by Taq DNA polymerase (NEB) using PBS7 and PEG-PBS8 primer pair, 7-deaza-2'-deoxyribose-GTP (7dG), and the other three normal dNTPs following standard PCR protocols. Then, the 7dG version of the 4NT9G template was separated from the complementary strand and PAGE purified. For each

reaction, the DNA primer (1 μM PBS8_extra) containing a single nucleotide mismatch was annealed to the DNA template to form the primer-template complex (1 μM). The primer-extension reaction was performed 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, NEB) supplemented with 1 mM $MnCl_2$ and tNTP mix using 1 uM Kod-RI at 55°C for 3 hours. All of the TNA synthesis reactions were performed with 100 uM tNTP, except for the reactions that contained 7-deaza-7-phenyl-tGTP, which was used at 75 uM. The fully-extended product was purified by denaturing PAGE and reverse transcribed into cDNA using 1 μM Bst-BF polymerase and a corresponding DNA primer. The cDNA was PCR amplified either with the Extra/PBS10 primer pair in the case of (1) or with the Extra/PBS7 primer pair in the case of (2) and (3), agarose purified, ligated into a TOPO vector, and subsequently cloned into E. coli DH5a competent cells following a typical chemical transformation protocol. Individual colonies were grown in liquid media and sequenced using the M13R primer by Retrogen, San Diego, CA. DNA sequences were aligned with the template used and analyzed using MEGA7 software.[52]

Computational studies of nucleobase orientation

Ab initio quantum mechanical calculations depicted in **Table 3-4** were performed using Turbomole 5.9.[53] The electrostatic potential energy surface for 7-deaza-7-iodo-guanine depicted in **Figure 3-4a** was calculated using Gaussian 09[54] at the B3LYP/MidiX level. Finally, the structure of the (t)7IGC/(d)TG duplex depicted in **Figure 3-4b** was derived from partial optimization at the B3LYP/MidiX level using Gaussian 09 with the 7IGT bases constrained to remain co-planar.

	E(h), RIDFT-bp/def2-TZVP	rel. E (kcal)	dipole (D)
G	-542.8233692210	0.00	6.92
G enol Scis	-542.8213489476	1.27	3.34
G enol Strans	-542.8208621820	1.57	3.35
7HG	-526.7691342037	0.00	4.68
7HG enol Scis	-526.7645298766	2.89	1.45
7HG enol Strans	-526.7578109052	7.11	2.59
7IG	-824.0219073372	0.00	6.40
7IG enol Scis	-824.0190783392	1.78	2.72
7IG enol Strans	-824.0167844020	3.21	3.06
7PhG	-757.9126754637	0.00	5.26
7PhG enol Scis	-757.9081022511	2.87	1.88

Table 3-5. Energies, relative energies and dipole moments of the canonical N^1-H and enol (O^2-H) tautomers for guanine (G), 7-deaza-guanine (7HG), 7-deaza-7-iodo-guanine (7IG), and 7-deaza-7-phenyl-guanine (7PhG).

References

1. Ellington, A. D.; Szostak, J. W. *Nature* **1990**, *346*, 818–822.
2. Tuerk, C.; Gold, L. *Science* **1990**, *249*, 505–510.
3. Dunn, M. R.; Jimenez, R. M.; Chaput, J. C. *Nat. Rev. Chem*. **2017**, *1*, 0076.
4. Zhou, J.; Rossi, J. Nat. Rev. Drug Discovery **2017**, *16*, 181–202.
5. Griffin, L. C.; Tidmarsh, G. F.; Bock, L. C.; Toole, J. J.; Leung, L. L. Blood **1993**, *81*, 3271–3276.
6. Bittker, J. A.; Phillips, K. J.; Liu, D. R. Curr. Opin. Chem. Biol. **2002**, *6*, 367–374.
7. Keefe, A. D.; Cload, S. T. Curr. Opin. Chem. Biol. **2008**, *12*, 448–456.
8. Cummins, L. L.; Owens, S. R.; Risen, L. M.; Lesnik, E. A.; Freier, S. M.; McGee, D.; Guinosso, C. J.; Cook, P. D. Nucleic Acids Res. **1995**, *23*, 2019–24.
9. Noronha, A. M.; Wilds, C. J.; Lok, C. N.; Viazovkina, K.; Arion, D.; Parniak, M. A.; Damha, M. J. *Biochemistry* **2000**, *39*, 7050–62.
10. Chaput, J. C.; Yu, H.; Zhang, S. Chem. Biol. **2012**, *19*, 1360–1371.
11. Houlihan, G.; Arangundy-Franklin, S.; Holliger, P. *Acc. Chem. Res*. **2017**, *50*, 1079–1087.
12. Chen, T.; Romesberg, F. E. *FEBS Lett.* **2014**, *588*, 219–229.
13. Pinheiro, V. B.; Taylor, A. I.; Cozens, C.; Abramov, M.; Renders, M.; Zhang, S.; Chaput, J. C.; Wengel, J.; Peak-Chew, S. Y.; McLaughlin, S. H.; Herdewijn, P.; Holliger, P. *Science* **2012**, *336*, 341–344.
14. Alves Ferreira-Bravo, I.; Cozens, C.; Holliger, P.; DeStefano, J. J. *Nucleic Acids Res.* **2015**, *43*, 9587–99.
15. Yu, H.; Zhang, S.; Chaput, J. C. *Nat. Chem*. **2012**, *4*, 183–187.
16. Taylor, A. I.; Pinheiro, V. B.; Smola, M. J.; Morgunov, A. S.; Peak-Chew, S.; Cozens, C.; Weeks, K. M.; Herdewijn, P.; Holliger, P. *Nature* **2015**, *518*, 427–430.
17. Schöning, K. U.; Scholz, P.; Guntha, S.; Wu, X.; Krishnamurthy, R.; Eschenmoser, A. *Science* **2000**, *290*, 1347–1351.
18. Anosova, I.; Kowal, E. A.; Dunn, M. R.; Chaput, J. C.; Van Horn, W. D.; Egli, M. *Nucleic Acids Res.* **2016**, *44*, 1007–1021.
19. Yang, Y.-W.; Zhang, S.; McCullum, E. O.; Chaput, J. C. *J. Mol. Evol*. **2007**, *65*, 289–295.
20. Culbertson, M. C.; Temburnikar, K. W.; Sau, S. P.; Liao, J.-Y.; Bala, S.; Chaput, J. C. Bioorg. Med. Chem. Lett. **2016**, *26*, 2418–2421.
21. Larsen, A. C.; Dunn, M. R.; Hatch, A.; Sau, S. P.; Youngbull, C.; Chaput, J. C. *Nat. Commun.* **2016**, *7*, 11235.
22. Tabor, S.; Richardson, C. C. *Proc. Natl. Acad. Sci. U. S. A*. **1989**, *86*, 4076–4080.
23. Yu, H.; Zhang, S.; Dunn, M. R.; Chaput, J. C. *J. Am. Chem. Soc.* **2013**, *135*, 3583–3591.
24. Sau, S. P.; Fahmi, N. E.; Liao, J.-Y.; Bala, S.; Chaput, J. C. *J. Org. Chem.* **2016**, *81*, 2302–2307.
25. Sau, S. P.; Chaput, J. C. *Org. Lett*. **2017**, *19*, 4379–4382.
26. Chaput, J. C.; Szostak, J. W. *J. Am. Chem. Soc.* **2003**, *125*, 9274–9275.
27. Dunn, M. R.; Chaput, J. C. *ChemBioChem* **2016**, *17*, 1804–1808.
28. Dunn, M. R.; Larsen, A. C.; Zahurancik, W. J.; Fahmi, N. E.; Meyers, M.; Suo, Z.; Chaput, J. C. *J. Am. Chem. Soc.* **2015**, *137*, 4014–4017.
29. Hoogsteen, K. *Acta Crystallogr*. **1963**, *16*, 907–916.
30. Skelly, J. V.; Edwards, K. J.; Jenkins, T. C.; Neidle, S. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 804–808.
31. Jager, S.; Famulok, M. *Angew. Chem., Int. Ed.* **2004**, *43*, 3337–40.
32. Bala, S.; Liao, J. Y.; Mei, H.; Chaput, J. C. *J. Org. Chem*. **2017**, *82*, 5910–5916.
33. Sousa, R. *Trends Biochem. Sci.* **1996**, *21*, 186–190.

34. Kolar, M.; Hostas, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2014**, *16*, 9987−96.
35. Ramzaeva, N.; Eickmeier, H.; Rosemeyer, H. *Chem. Biodiversity* **2014**, *11*, 532−41.
36. Auffinger, P.; Hays, F. A.; Westhof, E.; Ho, P. S. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 16789−94.
37. Ouellet, E.; Foley, J. H.; Conway, E. M.; Haynes, C. *Biotechnol. Bioeng*. **2015**, *112*, 1506−22.
38. Michalowski, D.; Chitima-Matsiga, R.; Held, D. M.; Burke, D. H. *Nucleic Acids Res.* **2008**, *36*, 7124−35.
39. Loeb, L. A.; Monnat, R. J., Jr. *Nat. Rev. Genet.* **2008**, *9*, 594−604.
40. Loakes, D.; Holliger, P. *Chem. Commun*. **2009**, 4619−4631.
41. Nikoomanzar, A.; Dunn, M. R.; Chaput, J. C. *Anal. Chem.* **2017**, *89*, 12622−12625.
42. Chim, N.; Shi, C.; Sau, S. P.; Nikoomanzar, A.; Chaput, J. C. *Nat. Commun.* **2017**, *8*, 1810.
43. Kool, E. T. *Annu. Rev. Biochem*. **2002**, *71*, 191−219.
44. Mei, H.; Shi, C.; Jimenez, R. M.; Wang, Y.; Kardouh, M.; Chaput, J. C. *Nucleic Acids Res.* **2017**, *45*, 5629−5638.
45. Kimoto, M.; Yamashige, R.; Matsunaga, K.-I.; Yokoyama, S.; Hirao, I. *Nat. Biotechnol*. **2013**, 31, 453−457.
46. Tolle, F.; Brandle, G. M.; Matzner, D.; Mayer, G. *Angew. Chem., Int. Ed*. **2015**, *54*, 10971−10974.
47. Vaught, J. D.; Bock, C.; Carter, J.; Fitzwater, T.; Otis, M.; Schneider, D.; Rolando, J.; Waugh, S.; Wilcox, S. K.; Eaton, B. E. *J. Am. Chem. Soc.* **2010**, *132*, 4141−51.
48. Nikoomanzar, A.; Dunn, M. R.; Chaput, J. C*., Curr. Protoc. Nucleic Acid Chem.* **2017**, 69, 4.75.1-4.75.20.
49. Seela, F.; Peng, X. H., *Synthesis* **2004**, 1203-1210.
50. Ramzaeva, N.; Mittelbach, C.; Seela, F., *Helv. Chim. Acta* **1997**, *80*, 1809-1822.
51. Ingale, S. A.; Seela, F., *J. Org. Chem. 2016*, **81**, 8331-8342.
52. Sudhir Kumar, Glen Stecher, Koichiro Tamura; MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets, Molecular Biology and Evolution, Volume 33, Issue 7, 1 July 2016, Pages 1870–1874, https://doi.org/10.1093/molbev/msw054
53. TURBOMOLE V5.9, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from http://www.turbomole.com.
54. Gaussian 09, Revision A.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2016.

Chapter 4

Synthesis of 2′-Deoxy-α-L-threofuranosyl Nucleoside Triphosphates

Publication note

This paper was originally published in the *Journal of Organic Chemistry.*

Saikat Bala,† Jen-Yu Liao,† Li Zhang, Chantel N. Tran, Nicholas Chim, and John C. Chaput. Synthesis of 2′-Deoxy-α-L-threofuranosyl Nucleoside Triphosphates. *J. Org. Chem*. **2018**, *83*, 8840−8850. († Equal contribution).

4.1 Contribution Statement

Liao, J.-Y. and Bala, S. designed the synthetic strategy. Liao, J. Y. and Bala, S. performed the synthesis, optimization, and compounds characterization. Li Zhang performed the primer-extension assay. Chantel N. Tran, and Nicholas Chim performed the protein crystallization. Liao, J.-Y. and Bala, S. involved in the experimental section writing and supporting information compiling. Liao, J.-Y., Bala, S. and Chaput, J. C. wrote the manuscript.

4.2 Abstract

α-L-Threofuranosyl nucleic acid (TNA) is an artificial genetic polymer in which the five-carbon ribose sugar found in RNA has been replaced with an four-carbon threose sugar. Despite a difference in sugar−phosphate backbone, TNA is capable of forming stable antiparallel Watson−Crick duplex with complementary strands of DNA and RNA and itself. This property of intersystem base pairing, coupled with the chemical simplicity of threose relative to ribose, provides support for the hypothesis that TNA is the potential candidate of RNA progenitor in the evolution of life. In an effort to evaluate the functional properties of TNA by in vitro evolution, engineered polymerases have been developed that are capable of

copying genetic information back and forth between DNA and TNA. However, the current developed TNA polymerases function with reduced activity relative to their natural counterparts, which limits the evaluation of TNA as a primordial genetic material. Here, we describe the synthesis of 2′-deoxy-α-L-threofuranosyl nucleoside 3′-triphosphates (dtNTPs) as chain-terminating substrates in a polymerase-mediated TNA synthesis reaction. The synthesis of dtNTPs should make it possible to investigate the mechanism of TNA synthesis by X-ray crystallography by trapping the polymerase in the catalytically active conformation.

4.3 Introduction

TNA (α-L-threofuranosyl nucleic acid) is an artificial genetic polymer developed by Eschenmoser and colleagues as part of an investigation into the chemical etiology of RNA.[1] Unlike most of the other RNA analogues examined in their study,[2] TNA was found to be capable of adopting stable Watson–Crick duplex structures with itself and with complementary strands of DNA and RNA.[3] This property of intersystem base pairing provides a mechanism for the transfer of genetic information between sequential genetic polymers in the evolution of life.[4] In an effort to broaden the concept of information transfer beyond the hybridization of short chemically synthesized oligonucleotides, engineered polymerases were developed to replicate TNA by copying information back and forth between TNA and DNA.[5,6] These enzymes were used to produce the first examples of TNA aptamers isolated by in vitro selection.[7–9] Since TNA is refractory to nuclease digestion,[10] TNA affinity reagents and catalysts have practical value in diagnostic and therapeutic applications that require high biological stability.[11,12]

Detailed kinetic studies measuring the rate of TNA synthesis by an engineered TNA polymerase indicate that TNA synthesis occurs at a rate that is vastly inferior to DNA

synthesis with natural DNA polymerases.[13] Insight into this problem was recently obtained when the X-ray crystal structure of a laboratory-evolved TNA polymerase was solved for the apo, binary, open and closed ternary complexes, and the translocated product postcatalysis.[14] Analysis of the enzyme active site in the closed ternary structure suggests that the slow rate of catalysis was due to a suboptimal binding geometry of the incoming TNA nucleoside triphosphate (tNTP).[14] Although this data provided a key piece of information that could be used as a framework for engineering new TNA polymerase variants, solving the structures of future TNA polymerases isolated along the same evolutionary trajectory would make it possible to uncover important structural features required to shape the design of a polymerase active site that is perfectly contoured for TNA substrates.

Although some unnatural nucleotide triphosphates are better substrates for DNA polymerases than are natural dNTPs,[15] most unnatural substrates function with reduced activity relative to their natural counterparts.[16,17] We therefore reasoned that an iterative process of directed evolution and X-ray crystallography could provide valuable insight into the general principles that govern polymerase specificity.[18] While our polymerase engineering endeavors have focused primarily on TNA, information gained from this process could be applied to other polymerase engineering efforts currently underway to develop polymerases that can replicate other types of unnatural genetic polymers.[19,20] A critical step in this process is to synthesize substrates that can trap the enzyme in the catalytic state of phosphodiester bond formation. Although our previous efforts utilized a DNA primer lacking a 2′-hydroxyl group on the terminal TNA residue, this approach was difficult to implement

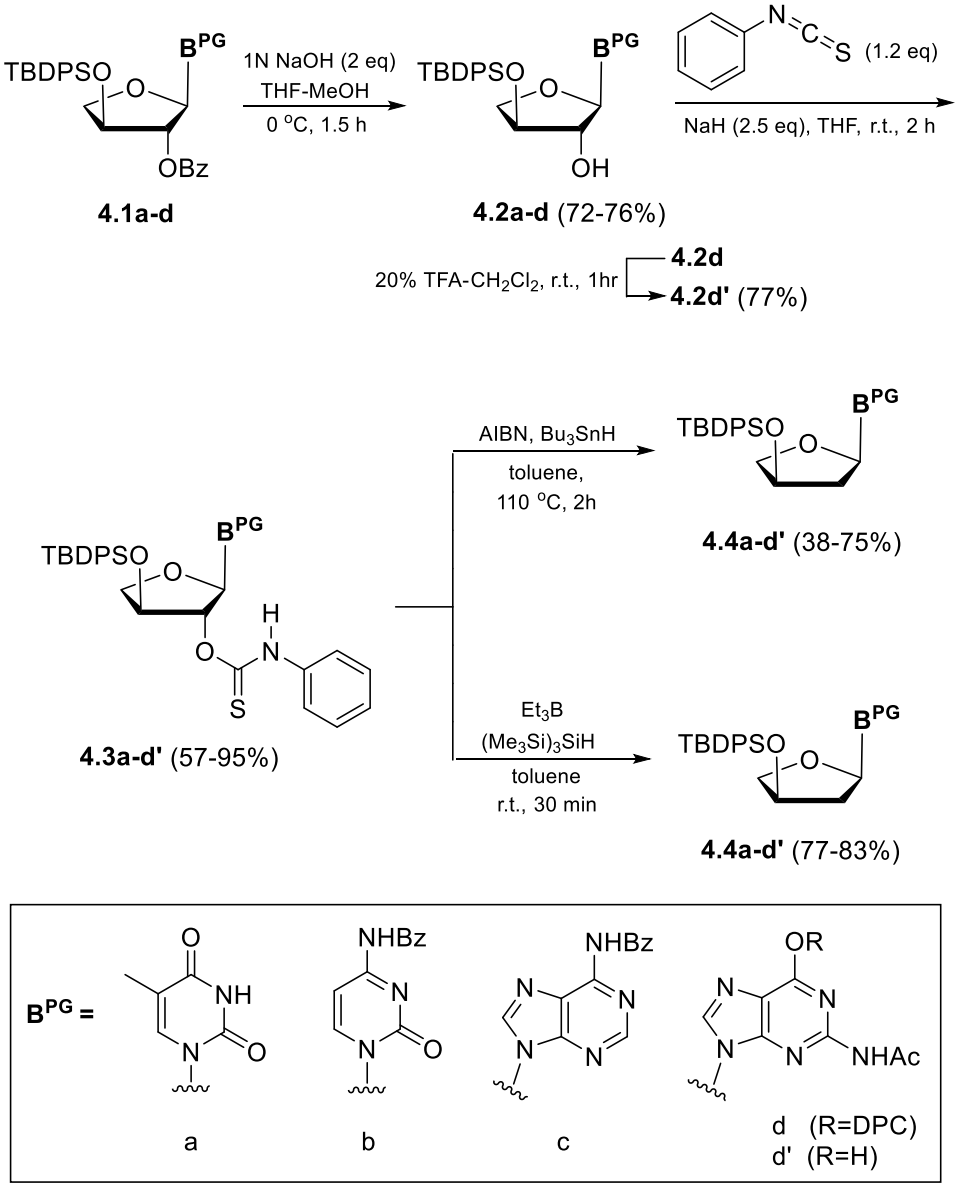
and often leads to an open ternary structure that captures the substrate binding step rather than the desired catalytic step observed in the closed ternary complex.[14]

An alternative approach is to use chain-terminating nucleoside triphosphates that undergo a cycle of catalysis prior to crystallization. This strategy has been used successfully to capture the closed ternary structures of DNA polymerases by trapping the second incoming DNA nucleoside triphosphate in the catalytically active conformation.[21,22] However, applying this approach to laboratory-evolved TNA polymerases requires chemical synthesis, as TNA substrates (amidites and triphosphates) are not available commercially. Here, we report the chemical synthesis of 2′-deoxy-α-L-threofuranosyl nucleoside 3′-triphosphates (dtNTPs) bearing the four genetic bases of adenine (A), cytosine (C), guanine (G), and thymine (T). The resulting substrates were obtained in eight chemical steps from advanced intermediates described in a previously published chemical synthesis pathway.[23] All four dtNTPs exhibit efficient site-specific chain termination in a TNA polymerization reaction and protein crystallization studies, with dtATP leading to the formation of high quality protein crystals that diffracted to 2.6–2.8 Å resolution.

4.4 Results and Discussions

We have previously described an efficient synthesis of **4.1a–d**, which involves eight chemical transformations and avoids the need for silica gel purification.[23] This strategy makes use of a highly efficient regioselective 2′-OH protection strategy developed by Herdewijn and colleagues,[24] which simplifies the synthesis of TNA monomers relative to an earlier strategy that required additional protection and deprotection steps.[25] Using this approach, protected nucleosides **4.1a–d** were generated (**Scheme 1-2**) as the starting

compounds for synthesizing the desired dtNTPs (**4.9a–d)** analogues described in the current study.

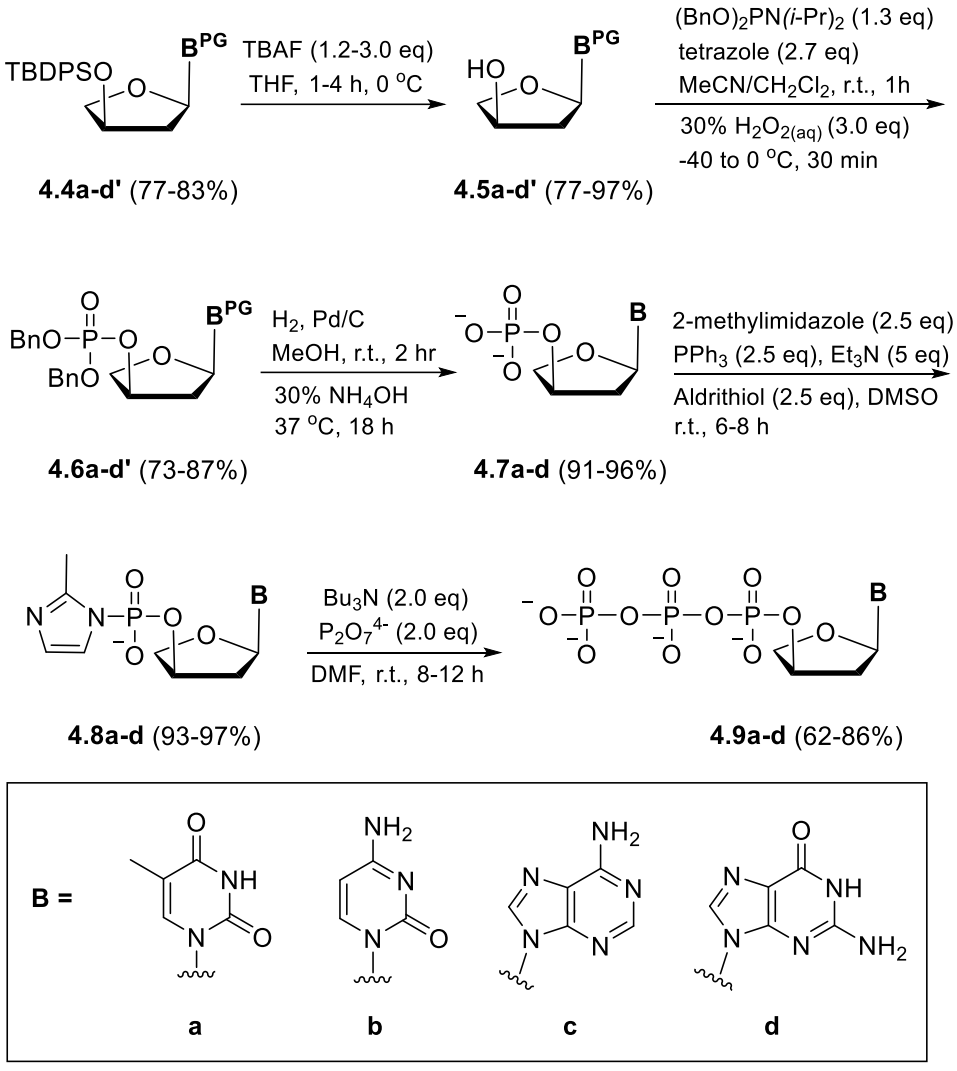


Scheme 4-1. Synthesis of 2′-Deoxy TNA Nucleosides 4.4a–d′

We began by selectively deprotecting the 2′-benzoyl group from compounds **4.1a–d** using a 1 N NaOH solution that was added dropwise to the protected TNA nucleosides (**Scheme 4-1**). Nucleosides **4.2a–d** were then treated with phenyl isothiocyanate under basic conditions to synthesize 2′-phenyl isothiocyanate derivatives **4.3a–d'**, which were

used as the precursors for a Barton deoxygenation reaction. This reaction proceeded efficiently for each nucleoside with the exception of guanosine analogue **4.2d**, which experienced significant loss of the diphenylcarbamoyl (DPC) group. To avoid the occurrence of a partially deprotected guanine base, we removed the DPC group prior to the synthesis of **4.3d′**.[27] This chemical substitution enabled us to synthesize all four TNA nucleosides as 2′-deoxy analogues in protected form. Using standard deoxygenation conditions with α,α′-azobisisobutyronitrile (AIBN) as a radical initiator and the tributylstannane as the proton source and thione scavenger, we observed efficient conversion of **4.3a–d′** to **4.4a–d′** for each nucleoside, except protected TNA cytidine derivative **4.3b**.[14,26,27] In this case, the reaction consistently produced a mixture of compounds that included **4.2b** as a side product.[28] In an effort to avoid this problem, we pursued milder reaction conditions that included the use of triethylborane and tris(trimethylsilyl)silane as alternative reagents for the homolytic deoxygenation of the 2′-hydroxyl group.[29,30] Dry oxygen was introduced into the reaction, which triggered a radical cascade and resulted in the clean synthesis of **4.4a–d′** within 30 min at room temperature.

The desilylation of **4.4a–d′** with 1 M tetrabutylammoniumfluoride (TBAF) in THF afforded TNA nucleoside derivatives **4.5a–d′** (**Scheme 4-2**). To confirm that our synthesis strategy yielded the correct nucleoside analogue, the molecular structure of **4.5c** was solved by small-molecule X-ray crystallography (**Figure 4-1**). The observed electron density and torsion angles are consistent with the formation of a 2′-deoxy analogue with the 3′-OH group occupying the β conformation on the furanose ring. Moreover, no electron density was observed for the 2′ hydroxyl moiety.



Scheme 4-2. Synthesis of 2′-Deoxy-α-L-threofuranosyl Nucleoside Triphosphates (4.9a–d)

Next, 3′-monophosphate triesters **4.6a–d′** were synthesized using a known one-pot phosphitylation–oxidation strategy.[31] Accordingly, compounds **4.5a–d′** were treated with dibenzyl phosphoramidite and tetrazole to convert the 3′-hydroxyl group into a phosphitylated P(III) derivative that was oxidized to P(V) with 30% aqueous H_2O_2. Nucleosides **4.6a–d′** were purified by flash column chromatography and hydrogenated using 0.1–0.2 mass equivalents of 10% palladium on charcoal. Following the removal of the

catalyst by filtration and solvent evaporation, the base protecting groups were removed upon treatment with 33% ammonium hydroxide (NH_4OH) for 18 h at 37 °C. Deprotection afforded desired 2′-deoxy TNA nucleoside monophosphates **4.7a–d** in >90% yield from **4.6a–d′**.

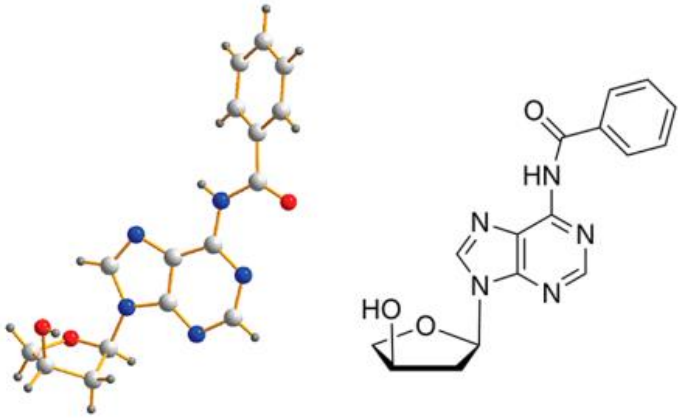


Figure 4.1 X-ray crystal structure of 6-benzoyl-2′-deoxy-α-L-threofuranosyl adenosine nucleoside 4.5c.

Although numerous strategies have been developed to synthesize nucleoside triphosphates,[32–34] we chose to follow a general strategy that has worked previously for TNA.[35] The desired dtNTPs were generated from activated monophosphates **4.8a–d** by displacing the activated 2-methylimidazole leaving group with pyrophosphate. TNA monophosphates **4.7a–d** were converted to their 3′-phosphoro-2-methylimidazole-activated derivatives using a Mitsunobu-like reaction of **4.7a–d** with excess 2-methylimidazole and triphenylphosphine in the presence of 2,2′-dipyridyl disulfide.[35] Nucleoside phosphoro-2-methylimidazolides (**4.8a–d**) were precipitated as a sodium salt with 10 equiv of sodium perchlorate in an ether/acetone solution. Compounds **4.8a–d** were then converted to the desired 2′-deoxy nucleoside triphosphates **4.9a–d** by treating **4.8a–d** with tributylammonium pyrophosphate and tributylamine in DMF. The reaction progress of

4.7a–d, **4.8a–d**, and **4.9a–d** was monitored by analytical HPLC, and compounds **4.9a–d** were purified by preparative HPLC.

Next, we used a primer extension assay to compare the ability of each 2′-deoxy TNA triphosphate to terminate TNA synthesis on a DNA primer–template complex (**Figure 4-2**). Primer–template complexes T1–T4 were designed to inhibit TNA synthesis midstream using one of four chemically synthesized dtNTP substrates **4.9a–d** (**Figure 4-2A**) that would site-specifically incorporate at a central position in the growing TNA strand. In each case, a 5′-IR labeled DNA primer annealed to a DNA template was incubated with Kod-RI TNA polymerase and the corresponding dtNTP solution for 60 min at 55 °C. Four different dtNTP solutions were prepared in which one of the four normal tNTP residues was individually replaced with the corresponding dtNTP analogue.

Figure 4-2. Sequence and primer extension results.

(A) DNA primer–template complexes were used to assess chain termination efficiency of dtNTP substrates. (B) Primer extension experiments were performed using 5′-IR labeled DNA primers annealed to DNA templates T1–T4 that were incubated with the desired TNA triphosphate solution and Kod-RI TNA polymerase for 60 min at 55 °C. Reactions were quenched and analyzed by denaturing polyacrylamide gel electrophoresis.

The chain-terminating analogues were compared to the normal tNTPs, which were used as positive control for the full-length product. An analysis of the resulting primer extension reactions by denaturing polyacrylamide gel electrophoresis revealed that all four dtNTP substrates mediate complete and efficient chain termination during TNA synthesis by producing a truncated product at the designated site in the template (**Figure 4-2B**). This result shows that the dtNTP substrates are recognized and processed by a laboratory-evolved TNA polymerase to produce a chain-terminated TNA extension product.

Next, we assessed the feasibility of using dtNTP substrates for X-ray crystallography. A binary complex of Kod-RI TNA polymerase bound to a primer–template duplex was incubated initially with excess 2′-deoxy-α-L-threofuranosyl adenosine nucleoside triphosphate **4.9c** to promote a single turnover event, followed by excess tTTP to capture the polymerase in a precatalytic state with the incoming TNA triphosphate bound in the enzyme active site. The resulting reaction mixture was utilized for crystallization trials, and the most promising crystals were selected for further optimization. A 2.6 Å resolution X-ray diffraction data set from the best crystal form (**Figure 4.3**), belonging to a space group of C2 with unit cell parameters a = 147.74 Å, b = 107.99 Å, and c = 70.71 Å, α = γ = 90°, β = 102.67, was collected. Molecular replacement using our previously solved Kod-RI ternary complex (ref 16) as a search model was unsuccessful, suggesting that the structure will have to be determined by experimental phasing. Nevertheless, this result provides confidence that dtNTPs offer a new route to ternary complexes of TNA polymerases developed by laboratory evolution.

The chemical synthesis of chain-terminating TNA triphosphates provides an opportunity to study the mechanism of TNA synthesis by laboratory-evolved TNA

polymerases. Previous structures reveal that Kod-RI recognizes the incoming tNTP substrates in a suboptimal geometry.14 This conformation is believed to account for the slow rate of TNA synthesis as compared to natural DNA polymerases. Using chain-terminating TNA substrates, it may be possible to improve our understanding of the principles that govern polymerase specificity. Such knowledge would improve our basic understanding of polymerase function and help close the gap in catalytic activity between engineered polymerases and their natural counterparts.

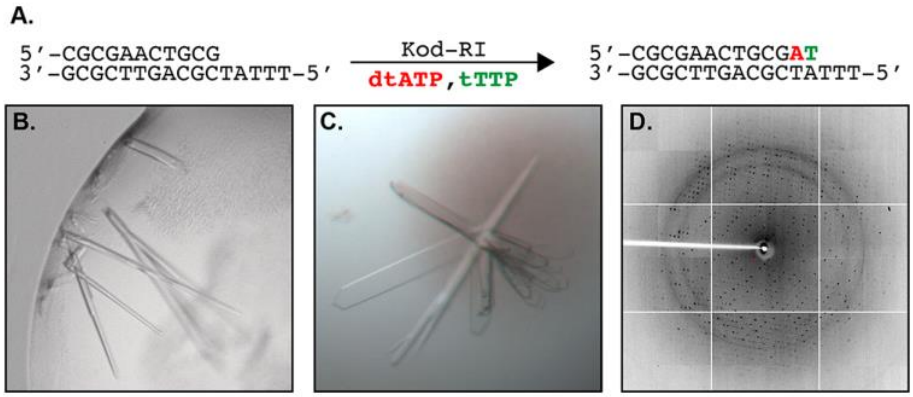


Figure 4-3. X-ray crystallography of Kod-RI bound to 4.9c.

(A) A schematic illustration showing the addition of the chain-terminating dtATP substrate (9c, red) to the primer−template duplex by Kod-RI. In the subsequent step, the tTTP (green) is trapped in the precatalytic state to form the ternary complex. (B,C) The optimization of ternary Kod-RI complex cocrystals. (D) The 2.6 Å resolution X-ray iffraction pattern of an optimized ternary Kod-RI crystal.

4.5 Conclusions

In summary, we describe the chemical synthesis of 2′-deoxy-α-L-threofuranosyl nucleoside 3′-triphosphates (dtNTPs) bearing all four genetic bases. We suggest that these substrates could find immediate use as chain-terminating reagents in applications related to the mechanism of TNA synthesis by a TNA polymerase. Structural information gained from

these studies could be used to develop new polymerase variants with greater TNA synthesis activity.

4.6 Experimental Details

General Information

All reactions were carried out in anhydrous solvents under an argon or nitrogen atmosphere unless mentioned otherwise. All reagents and solvents were obtained from commercial sources and used without further purification: Sublimed tetrazole in dry acetonitrile was obtained from Glen Research, dibenzyl diisopropylphosphoramidite was obtained from Combi-Blocks, 10% palladium on activated carbon (reduced, dry powder) was obtained from Strem Chemicals, and 5.5 M tert-butyl hydroperoxide in decane over molecular sieves was obtained from Sigma-Aldrich. Hydrogen and nitrogen gases were purchased as USP grade. Reaction progress was monitored by thin layer chromatography using glass-backed analytical SiliaPlate with UV-active F254 indicator. Flash column chromatography was performed with SiliaFlash P60 silica gel (40−63μm particle size). The ^{1}H, ^{13}C, or ^{31}P nuclear magnetic resonance (NMR) spectra were recorded at room temperature on either a Bruker DRX 400 or 500 MHz spectrometer at the University of California, Irvine, NMR Facility. The ^{1}H and ^{13}C NMR chemical shifts (δ) are reported in parts per million (ppm) with tetramethylsilane or deuterium solvent as an internal reference. The ^{31}P NMR chemical shifts are proton decoupled and reported in parts per million relative to an external standard of 85% H_3PO_4. Peaks multiplicity are designated with the following abbreviations: s, singlet; d, doublet; dd, doublet of doublets; t, triplet; m, multiplet; br, broad. HPLC analysis was performed on a reverse-phase C18 150 × 4.6 mm^2 column with 5 μm particle size, and nucleoside triphosphates were purified on a semipreparative reverse-

phase C18 250 × 9.4 mm^2 column (Thermo Scientific, USA) using a mobile phase of 100 mM triethylammonium acetate buffer (pH 7.0)/acetonitrile. Purified molecules that contain phosphate groups were converted into their corresponding Na^+ salt using the Dowex Marathon C Na+-form of resin before recording mass spectra. High resolution mass spectrometry (HRMS) data were acquired using the electrospray ionization time-of-flight (ESI-TOF) method at the University of California, Irvine, Mass Spectrometry Facility. The nucleoside monophosphates, phosphorimidazolides and triphosphates were dissolved in RNase-free water with 10 mM Tris pH 8.0 and quantified by Nanodrop 2000 (Thermo Scientific, USA) using Beer's law. (T, ε_{267}: 9600 $M^{-1}cm^{-1}$. C, ε_{271}: 91 00 $M^{-1}cm^{-1}$. G, ε_{259}: 15200 $M^{-1}cm^{-1}$. A, ε_{259}: 15400 $M^{-1}cm^{-1}$.

1-(2'-*O*-Benzoyl-3'-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)-thymine(4.1a).

Product **4.1a** was prepared as described previously.[24] The crude product was purified on silica gel with eluents (EtOAc/hexane, from 80 to 100%) to afford the 1-(2′-*O*-benzoyl-3′-*O*-t-butyldiphenylsilyl-α-L-threofuranosyl)thymine (**4.1a**) as a white solid: silica gel TLC (hexane/EtOAc, 1:1) Rf = 0.35; ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.42 (s, 1H), 7.93 (d, 2H, J = 6.4 Hz), 7.66−7.65 (m, 5H), 7.57−7.55 (m, 1H), 7.47−7.34 (m, 8H), 5.47 (s, 1H), 5.46 (s, 1H), 4.41 (d, 1H, J = 4 Hz), 4.16 (d, 1H, J = 8 Hz), 3.92 (dd, 1H, J = 5.2, 2.8 Hz), 1.91 (s, 3H), 1.11 (s, 9H); ^{13}C NMR (125.8 MHz) δ 164.8, 163.7, 150.2, 136.4, 135.8, 133.7, 132.7, 131.8, 130.5, 130.0, 129.0, 128.6, 111.1, 89.8, 82.7, 75.6, 27.0, 13.3, 12.7; HRMS (ESI-TOF) $C_{32}H_{34}N_2O_6SiNa$ $[M + Na]^+$ 593.2084; found 593.2089.

***N*⁴-Benzoyl-1-(2'-*O*-benzoyl-3'-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)cytosine (4.1b).**

Product **4.1b** was prepared as described previously.[24] The crude product was purified on silica gel with eluents (EtOAc/CH_2Cl_2, from 20 to 35%) to afford the *N*⁴-benzoyl-1-(2′-*O*-benzoyl-3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)cytosine (**4.1b**) as a white solid: silica gel TLC (hexane/EtOAc, 1:1) Rf = 0.35; ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.25 (d, 1H, *J* = 7.6 Hz), 7.97 (d, 2H, *J* = 7.6 Hz), 7.93 (d, 2H, *J* = 7.6 Hz), 7.64−7.36 (m, 18 H), 6.21 (s, 1H), 5.67 (s, 1H), 4.40 (d, 1H, *J* = 2.4 Hz), 4.17 (d, 1H, *J* = 10 Hz), 4.07 (d, 1H, *J* = 10, 3.2 Hz), 1.07 (s, 3H); ^{13}C NMR (125.8 MHz, $CDCl_3$) δ 164.5, 136.0, 135.8, 133.6, 133.3, 132.7, 131.7, 130.4, 130.4, 130.0, 129.2, 129.2, 128.5, 128.1, 128.1, 127.7, 91.4, 82.1, 77.0, 75.8, 27.0, 19.2; HRMS (ESI-TOF) $C_{38}H_{37}N_3O_6SiNa$ $[M + Na]^+$ 682.2349; found 682.2339.

N^6-Benzoyl-9-(2'-*O*-benzoyl-3'-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)adenine (4.1c).

Product **4.1c** was prepared as described previously.[24] The crude product was purified on silica gel with eluents (EtOAc/hexane, from 33 to 50%) to afford the *N*⁶-benzoyl-9-(2′-*O*-benzoyl-3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)adenine (**4.1c**) as a white solid: silica gel TLC (hexane/EtOAc, 1:1) Rf = 0.40; ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.25 (d, 1H, *J* = 7.6 Hz), 7.97 (d, 2H, *J* = 7.6 Hz), 7.93 (d, 2H, *J* = 7.6 Hz), 7.64−7.36 (m, 18 H), 6.21 (s, 1H), 5.67 (s, 1H), 4.40 (d, 1H, *J* = 2.4 Hz), 4.17 (d, 1H, *J* = 10 Hz), 4.07 (d, 1H, *J* = 10, 3.2 Hz), 1.07 (s, 3H); ^{13}C NMR (125.8 MHz, $CDCl_3$) δ 162.7, 145.5, 136.2, 136.0, 136.0, 135.9, 133.4, 132.8, 131.8, 130.5, 130.4, 129.2, 129.1, 128.2, 128.2, 128.1, 127.8, 126.0, 75.8, 66.0, 27.2, 19.3; HRMS (ESI-TOF) $C_{39}H_{37}N_5O_5SiNa$ $[M + Na]^+$ 706.2462; found 706.2471.

N^2-Acetyl-O^6-diphenylcarbamoyl-9-(2′-*O*-benzoyl-3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)guanine (4.1d).

Product **4.1d** was prepared as described previously.[24] The crude product was purified using flash column chromatography on a silica gel with eluents (EtOAc/hexane, from 80 to 100%) to afford *N*^{2}-acetyl-*O*^{6}-diphenylcarbamoyl-9-(2′-*O*-benzoyl-α-L-threofuranosyl)guanine (**4.1d**) as a yellow foam: silica gel TLC (hexane/EtOAc, 1:1) Rf = 0.65; ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.42 (s, 1H), 7.95–7.90 (m, 3H), 7.62–7.55 (m, 5H), 7.44–7.34 (m, 15H), 7.31–7.23 (m, 5H), 6.16 (d, 1H, *J* = 1.2 Hz), 5.81 (s, 1H), 4.57 (m, 1H), 4.22 (dd, 1H, *J* = 12, 7.6 Hz), 4.11 (dd, 1H, *J* = 14.4, 5.6 Hz), 2.50 (s, 3H), 1.03 (s, 9H); ^{13}C NMR (125.8 MHz, $CDCl_3$) δ 164.8, 156.5, 154.6, 152.4, 150.4, 142.4, 135.8, 134.0, 132.5, 131.9, 130.5, 130.0, 128.7, 128.1, 120.9, 88.8, 82.5, 76.2, 75.9, 27.0, 25.3, 19.1; HRMS (ESI-TOF) $C_{47}H_{44}N_6O_7SiNa$ $[M + Na]^+$ 855.2938; found 855.2960.

1-(3′-*O*-tert-Butyldiphenylsilyl-α-L-threofuranosyl)thymine (4.2a).

To a solution containing 3.73 g (6.54 mmol) of 1-(2′-*O*-benzoyl-3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)thymine (**4.1a**) in 12 mL of THF and 5 mL of MeOH at 0 °C was added 6.54 mL (6.54 mmol) of ice-cold 1 N $NaOH_{(aq)}$ dropwise while stirring. After 15 min, additional 3.27 mL (3.27 mmol) of ice-cold 1 N $NaOH_{(aq)}$ was added to the reaction for an additional 1 h of stirring, at which time TLC (hexane/EA, 1:1) showed the complete consumption of the starting material, the reaction was quenched with the dropwise addition of 5 mL of 1 N $HCl_{(aq)}$ at 0 °C. The solution was condensed to a 20 mL volume under reduced pressure, and the crude product was diluted with 50 mL of EtOAc. The organic layer was washed with 50 mL of H_2O and then 50 mL of brine, dried over $MgSO_4$, and evaporated under reduced pressure. The residue was purified on silica gel with eluents (MeOH/CH_2Cl_2, from 0 to 5%). Product **4.2a** was obtained as a white foam: yield 2.17 g (72.1%); silica gel TLC (CH_2Cl_2/MeOH, 20:1) Rf = 0.65; ^{1}H NMR (400 MHz, $CDCl_3$) δ 10.76 (s, 1H), 7.60 (dd, 2H, *J* =

5.2, 4.0 Hz), 7.59 (d, 1H, *J* = 1.2 Hz), 7.52 (dd, 2H, *J* = 8.0, 1.2 Hz), 7.43−7.41 (m, 1H), 7.39−7.32 (m, 5 H), 5.74 (s, 1H), 5.28 (d, 1H, *J* = 7.2 Hz), 4.39 (s, 1H), 4.28 (d, 1H, *J* = 2 Hz), 4.17 (d, 2H, *J* = 3.2 Hz), 1.87 (s, 3H), 1.01 (s, 9H); ^{13}C NMR (125.8 MHz, $CDCl_3$) *δ* 164.9, 151.1, 136.8, 135.8, 132.9, 130.3, 128.1, 109.9, 94.2, 81.5, 77.7, 26.9, 19.1, 12.8; HRMS (ESI-TOF) $C_{25}H_{30}N_2O_5SiNa$ $[M + Na]^+$ 489.1822; found 489.1822.

***N*^{4}-Benzoyl-1-(3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)cytosine (4.2b)**.

A stirring solution containing 6.8 g (10.31 mmol) of *N*^{4}-benzoyl-1-(2′-*O*-benzoyl-3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)cytosine (**4.1b**) in 51 mL of THF and 51 mL of MeOH was cooled to 0 °C. The solution was added to 10.3 mL (10.3 mmol) of ice-cold 1 N $NaOH_{(aq)}$ dropwise while stirring. After 30 min, another 10.3 mL (10.3 mmol) of ice-cold 1 N $NaOH_{(aq)}$ was added dropwise to the solution for an additional 30 min of stirring, at which time the TLC showed the reaction was finished; the solution was quenched by adding 10.3 mL of 1 N $HCl_{(aq)}$ at 0 °C. The crude product was condensed to 50 mL of solution under reduced pressure, and the solution was poured into 200 mL of EtOAc. The organic layer was washed with 200 mL of brine then 200 mL of H2O, dried over $MgSO_4$, and evaporated under reduced pressure. The crude product was purified by flash column chromatography with eluents ($MeOH/CH_2Cl_2$, from 1.6 to 2.5%) to afford *N*^{4}-benzoyl-1-(3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)cytosine (**4.2b**) as a white solid: yield 3.84 g (67.1%); silica gel TLC ($MeOH/CH_2Cl_2$, 1:40) Rf = 0.45; ^{1}H NMR (400 MHz, $CDCl_3$) *δ* 8.19 (d, 1H, *J* = 7.6 Hz), 7.94 (d, 2H, *J* = 7.2 Hz), 7.65−7.56 (m, 5H), 7.52−7.48 (m, 4H), 7.44−7.32 (m, 8H), 5.73 (s, 1H), 4.39 (s, 1H), 4.30 (m, 1H), 4.15 (m, 2H), 0.98 (s, 9H); ^{13}C NMR (125.8 MHz, $CDCl_3$) *δ* 162.7, 156.0, 145.0, 135.8, 135.7, 133.3, 133.1, 132.9, 132.6, 130.3, 130.2, 129.1, 128.0, 128.0,

127.9, 127.8, 96.4, 95.1, 81.9, 77.1, 26.9, 19.1; HRMS (ESI-TOF) $C_{31}H_{33}N_3OSiNa$ $[M + Na]^+$ 578.2087; found 578.2092.

N^6-Benzoyl-9-(3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)adenine (4.2c).

A stirring solution containing 4.0 g (5.85 mmol) of N^6-benzoyl-9-(2′-*O*-benzoyl-3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)adenine (**4.1c**) in 29 mL of THF and 29 mL of MeOH was cooled to 0 °C. To the solution was added 5.8 mL (5.8 mmol) of ice-cold 1 N $NaOH_{(aq)}$ dropwise while stirring. After 30 min, another 5.8 mL (5.8 mmol) of ice-cold 1 N $NaOH_{(aq)}$ was added dropwise to the solution for an additional 1 h of stirring, at which time the TLC showed the reaction was finished; the solution was quenched by adding 5.8 mL of 1 N $HCl_{(aq)}$ at 0 °C. The crude product was condensed to 25 mL of solution under reduced pressure, and the solution was poured into 200 mL of EtOAc. The organic layer was washed with 200 mL of brine and then 200 mL of H_2O, dried over $MgSO_4$, and evaporated under reduced pressure. The crude product was purified by flash column chromatography with eluents (EtOAc/ CH_2Cl_2, from 30 to 60%) to afford N^6-benzoyl-9-(3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)adenine (2c) as a white solid: yield 2.6 g (76.7%); silica gel TLC (EtOAc/CH_2Cl_2, 2:1) Rf = 0.33; ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.23 (s, 1H), 8.78 (s, 1H), 8.42 (s, 1H), 8.01 (d, 2H, J = 6.4 Hz), 7.60−7.55 (m, 3H), 7.49 (t, 2H, J = 6.4 Hz), 7.44−7.43 (m, 3H), 7.39−7.35 (m, 3H), 7.28−7.25 (m, 2H), 6.05 (s, 1H), 4.65 (s, 1H), 4.23−4.21 (m, 1H), 4.16 (dd, 1H, J = 4.4, 3.2 Hz), 0.94 (s, 9H); ^{13}C NMR (125.8 MHz, $CDCl_3$) δ 164.7, 152.6, 150.9, 149.7, 141.5, 135.8, 135.7, 133.8, 132.9, 132.6, 132.5, 130.3, 130.3, 129.0, 128.1, 128.0, 128.0, 123.2, 92.1, 81.6, 77.6, 76.3, 26.9, 19.0; HRMS (ESI-TOF) $C_{32}H_{33}N_5O_4SiNa$ $[M + Na]^+$ 602.2200; found 602.2216.

N^2-Acetyl-O^6-diphenylcarbamoyl-9-(3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl) guanine (4.2d).

To a solution containing 6.0 g (7.20 mmol) of N^2-acetyl-O^6-diphenylcarbamoyl-9-(2′-*O*-benzoyl-α-L-threofuranosyl) guanine (**4.1d**) in 25 mL of THF and 5 mL MeOH at 0 °C was added 8.6 mL of ice-cold 1 N $NaOH_{(aq)}$ dropwise while stirring. After 20 min of stirring, an additional 4.3 mL of ice-cold 1 N $NaOH_{(aq)}$ was added dropwise to the reaction for another 1.5 h of stirring, at which TLC (hexane/EA, 1:1) showed the complete consumption of the starting material. The solution was quenched by adding 8 mL of 1 N $HCl_{(aq)}$ at 0 °C. The crude product was condensed to 15 mL of solution under reduced pressure, and the solution was poured into 150 mL of EtOAc. The organic layer was washed with 100 mL of brine and then 100 mL of H_2O, dried over $MgSO_4$, and evaporated under reduced pressure. The crude product was purified by flash column chromatography with eluents (EtOAc/hexane, from 25 to 55% with 5% CH_2Cl_2) to afford product **4.2d** as a yellow solid: yield 3.8 g (72.4%); silica gel TLC (EtOAc/CH_2Cl_2, 2:1) Rf = 0.33; ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.46 (s, 1H), 8.26 (s, 1H), 7.58 (d, 4H, J = 6.4 Hz), 7.41−7.30 (m, 14H), 7.25−7.23 (m, 2H), 5.86 (s, 1H), 4.56 (s, 1H), 4.41 (s, 1H), 3.97 (m, 2 H), 2.18 (s, 3H), 1.00 (s, 9H); ^{13}C NMR (125.8 MHz, CD_3OD) δ 156.1, 154.2, 151.6, 150.7, 142.4, 135.9, 135.8, 133.2, 132.6, 130.1, 129.4, 128.0, 128.0, 121.6, 92.3 82.5, 77.3, 74.9, 27.0, 25.1, 19.2; HRMS (ESI-TOF) $C_{40}H_{40}N_6O_6SiNa$ $[M + Na]^+$ 751.2676; found 751.2681.

N^2-Acetyl-9-(3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)-guanine (4.2d′).

To a cooled solution of 27 mL of 20% trifluoroacetic acid in CH_2Cl_2 was added 3.59 g (4.93 mmol) of N^2-acetyl-O^6-diphenylcarbamoyl-9-(3′-O-tert-butyldiphenylsilyl-α-L-threofuranosyl)guanine (**4.2d**), and it was stirred at 0 °C for 1 h. The reaction mixture was

evaporated to dryness under reduced pressure, and the resulting crude product was purified by silica gel column chromatography with eluents (MeOH/CH2Cl2) to afford product **4.2d′** as a pale-yellow solid: yield 2.4 g (77%); TLC (MeOH/CH_2Cl_2, 1:10) Rf = 0.45; ^{1}H NMR (400 MHz, $CDCl_3$) δ 12.07 (s, 1H), 9.27 (s, 1H), 8.30 (s, 1H), 7.59–7.53 (m, 3H), 7.45–7.34 (m, 6H), 7.26 (d, 2H, J = 8 MHz), 6.09 (s, 1H), 4.57 (s, 1H), 4.01–3.99 (m, 1H), 2.47–2.44 (m, 1H), 2.28 (s, 3H), 1.02 (s, 9H); ^{13}C NMR (125.8 MHz, $CDCl_3$) δ 172.1, 148.4, 147.6, 138.5, 136.1, 133.2, 133.1, 130.6, 130.5, 128.4, 128.3, 121.6, 84.3, 80.0, 72.5, 41.5, 27.2, 24.9, 19.3; HRMS (ESI-TOF) $C_{27}H_{31}N_5O_5SiNa$ $[M + Na]^+$ 556.1992; found 556.2003.

1-(2′-*O*-Phenylthioxocarbamoyl-3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl) thymine (**4.3a**).

To a solution containing 3.72 g (8.08 mmol) of 1-(3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)thymine (**4.2a**) and 0.52 g (21.8 mmol) of NaH in 60 mL of anhydrous THF was added dropwise 1.40 mL (11.3 mmol) of phenyl isothiocyanate, and the mixture was stirred for 1.5 h in room temperature. When TLC (hexane/EtOAc, 3:2) showed the reaction was completed, the reaction was quenched with the addition of 10 mL of water at 0 °C. The volatile solution was evaporated under reduced pressure, and the residue was dissolved in 100 mL of EtOAc. The organic layer was washed with 200 mL of brine and then 200 mL of H_2O, dried over $MgSO_4$, and evaporated under reduced pressure. The residue was purified on silica gel, eluting with 60–80% EtOAc in hexane in steps of 5% increase in EtOAc for every 200 mL. Product **4.3a** was obtained as a white solid: yield 4.01 g (83%); silica gel TLC (hexane/EtOAc, 1:1) Rf = 0.35; ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.47 (s, 1H), 7.67 (m, 5H), 7.51–7.34 (m, 6H), 7.25–7.24 (m, 2H), 7.12–7.02 (m, 1H), 6.08 (s, 1H), 5.92 (s, 1H), 4.45 (s, 1H), 4.05 (d, 1H, J = 10 Hz), 3.84 (m, 1H), 3.78 (m, 1H), 1.86 (s, 3H), 1.10 (s, 9H); ^{13}C NMR

(125.8 MHz, $CDCl_3$) δ 164.2, 150.7, 136.0, 135.7, 132.8, 130.4, 129.0, 126.0, 128.1, 111.4, 89.5,75.0, 27.0, 19.3, 12.5; HRMS (ESI-TOF) $C_{32}H_{35}N_3O_5SSiNa$ $[M + Na]^+$ 624.1964; found 624.1959.

N^4-Benzoyl-1-(2′-*O*-phenylthioxocarbamoyl-3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)cytosine (**4.3b**).

To a stirring solution containing 2.60 g (4.68 mmol) of N^4-benzoyl-1-(3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)cytosine (**4.2b**) and 319 mg (12.64 mmol) of NaH in 47 mL of anhydrous THF was added dropwise 813 μL (6.56 mmol) of phenyl isothiocyanate at 0 °C. The reaction was stirring at room temperature for 1.5 h, at which time the TLC shows that the reaction was finished. The mixture was quenched slowly at 0 °C by adding dropwise 2 mL of water followed by 8 mL of 1 N $HCl_{(aq)}$. The crude mixture was condensed to 20 mL of solution under reduced pressure, and the solution was poured into 200 mL of EtOAc. The organic layer was washed with 200 mL of brine and then 200 mL of H_2O, dried over $MgSO_4$, and evaporated under reduced pressure. The crude product was purified by flash column chromatography with eluents (EtOAc/CH2Cl2, from 13 to 25%) to afford product **4.3b** as a white solid: yield 3.1 g (95.9%); silica gel TLC (EtOAc/CH_2Cl_2, 1:4) Rf = 0.4; ^1H NMR (400 MHz, $CDCl_3$) δ 8.24 (d, 1H, J = 6.8 Hz), 7.93 (d, 2H, J = 7.6 Hz), 7.64−7.26 (m, 18H), 7.13 (t, 1H, J = 7.2 Hz), 5.99 (s, 1H), 4.50 (d, 1H, J = 2.4 Hz), 4.12 (d, 1H, J = 10 Hz), 1.04 (s, 9H); ^{13}C NMR (125.8 MHz, $CDCl_3$) 165.0, 152.8, 151.8, 149.6, 141.9, 135.9, 135.7, 133.7, 132.7, 132.5, 131.7, 130.3, 128.8, 128.1, 128.0, 126.1, 123.0, 87.9, 75.3, 29.7, 26.9, 19.1; HRMS (ESI-TOF) $C_{38}H_{38}N_4O_5SSiNa$ $[M + Na]^+$ 713.2230; found 713.2205.

N^6-Benzoyl-9-(2′-*O*-phenylthioxocarbamoyl-3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)adenine (4.3c).

To a stirring solution containing 3.3 g (5.69 mmol) of N^6-benzoyl-9-(3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)adenine (**4.2c**) and 368 mg (15.36 mmol) of NaH in 57 mL of anhydrous THF was added dropwise 917 μL (7.40 mmol) of phenyl isothiocyanate at 0 °C. The reaction was stirred at room temperature for 1.5 h, at which time the TLC shows that the reaction was finished. The mixture was quenched slowly at 0 °C by adding dropwise 3 mL of water followed by 10 mL of 1 N $HCl_{(aq)}$. The crude mixture was condensed to 30 mL of solution under reduced pressure, and the solution was poured into 200 mL of EtOAc. The organic layer was washed with 200 mL of brine and then 200 mL of H_2O, dried over $MgSO_4$, and evaporated under reduced pressure. The crude product was purified with flash column chromatography with eluents (EtOAc/CH2Cl2, from 13 to 25%) to afford **4.3c** as a yellowish solid: yield 3.72 g (91.4%); silica gel TLC (EtOAc/CH_2Cl_2, 1:4) Rf = 0.45; ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.71 (s, 1H), 8.59 (s, 1H), 8.00 (d, 2H, J = 6.8 Hz), 7.65−7.04 (m, 18H), 6.37 (s, 2H), 4.58 (s, 1H), 4.13 (s, 1H), 3.94 (s, 1H), 1.07 (s, 9H); ^{13}C NMR (100.6 MHz, $CDCl_3$) δ 165.0, 152.8, 151.8, 149.6, 141.9, 135.9, 135.7, 133.7, 132.7, 132.5, 131.7, 130.3, 128.8, 128.1, 128.0, 126.1, 123.0, 87.9, 75.3, 29.7, 26.9, 19.1; HRMS (ESI-TOF) $C_{39}H_{38}N_6O_4SSiNa$ $[M + Na]^+$ 737.2343; found 737.2358.

N^2-Acetyl-9-(2′-*O*-phenylthioxocarbamoyl-3′-O-tert-butyldiphenylsilyl-α-L-threofuranosyl)guanine (4.3d′).

To a solution containing 1.32 g (2.47 mmol) of **4.2d′** and 0.27 g (6.68 mmol) of NaH in 20 mL of anhydrous THF was added dropwise 0.37 mL (2.97 mmol) of phenyl isothiocyanate, and the mixture was stirred for 1.5 h at room temperature. When TLC

(MeOH/$CHCl_3$, 1:10) showed the complete consumption of starting material, the reaction was quenched with the addition of 10 mL of water. The volatile solution was evaporated under reduced pressure, and the residue was dissolved in 100 mL of EtOAc. The organic layer was washed with 200 mL of brine and then 200 mL of H_2O, dried over $MgSO_4$, and evaporated under reduced pressure. The residue was purified by silica gel chromatography with eluents (MeOH/DCM, from 0 to 6%) to afford the product **4.3d′** as a pale-yellow solid: yield 920 mg (57%); silica gel TLC (CH_2Cl_2/MeOH, 10:1) Rf = 0.65; ^{1}H NMR (400 MHz, $CDCl_3$) δ 12.33–11.91 (m, 1H), 9.74 (s, 1H), 8.23–8.16 (m, 1H), 7.57 (m, 5H), 7.41–7.25 (m, 7H), 7.20–7.17 (m, 2H), 7.06 (s, 1H), 6.50 (s, 1H), 3.79 (s, 1H), 3.75–3.72 (m, 2H), 2.16 (d, 3H, *J* = 12 Hz), 0.93 (d, 9H, *J* = 18 Hz); ^{13}C NMR (100.6 MHz, $CDCl_3$) δ 156.1, 148.6, 147.9, 137.9, 135.9, 135.7, 132.5, 131.8, 130.2, 129.8, 129.0, 128.0, 127.8, 125.3, 122.8, 121.4, 115.2, 88.8, 87.5, 77.3, 75.5, 26.9, 26.8, 24.2, 19.0; HRMS (ESI-TOF) $C_{34}H_{36}N_6O_5SSiNa$ $[M + Na]^+$ 691.2135; found 691.2120.

1-(2′-Deoxy-3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)thymine (4.4a).

1-(2′-*O*-Phenylthioxocarbamoyl-3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl) thymine (**4.3a**) was coevaporated twice with 3 mL of anhydrous toluene under reduced pressure and dried overnight under vacuum. To a solution of 1.09 g (2.14 mmol) of **4.3a** in 15 mL of anhydrous toluene was added 2.08 mL (5.35 mmol) of tris(trimethylsilyl)silane followed by 2.57 mL of triethylborane (1 M in THF) under an argon atmosphere. The reaction was flushed with dry air several times, and the resulting mixture was stirred at room temperature for 45 min. When TLC (CH_2Cl_2/MeOH, 10:1) showed the reaction to be complete, the reaction was quenched with the addition of 7 mL of water. The solvent was evaporated under reduced pressure, and the residue was then purified by silica gel column

chromatography eluent (MeOH/CH_2Cl_2, from 0 to 8%) to afford the product **4.4a:** yield 760 mg (77.8%); silica gel TLC (CH_2Cl_2/MeOH, 20:1) Rf = 0.65; ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.48 (s, 1H), 7.70 (d, 1H, *J* = 1.2 Hz), 7.63 (dd, 2H, J = 9.2, 6.8 Hz), 7.55 (dd, 2H, *J* = 9.6, 6.4 Hz), 7.46−7.36 (m, 6H), 6.12 (dd, 1H, *J* = 10, 5.6 Hz), 4.47−4.35 (m, 1H), 4.14 (m, 1H), 3.75 (dd, 1H, *J* = 13.6, 6.2 Hz), 2.43−2.36 (m, 1H), 2.12 (d, 1H, J = 16 Hz), 1.93 (d, 3H, *J* = 1.2 Hz), 1.61 (s, 1H), 1.05 (s, 9H); ^{13}C NMR (125.8 MHz, $CDCl_3$) δ 164.0, 150.5, 136.7, 135.8, 133.0, 130.8, 128.1, 110.3, 86.1, 77.6, 72.3, 41.8, 27.0, 19.2, 12.8; HRMS (ESI-TOF) $C_{25}H_{30}N_2O_4SiNa$ $[M + Na]^+$ 473.1873; found 473.1868.

N^4-Benzoyl-1-(2′-deoxy-3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)cytosine (4.4b).

N^4-Benzoyl-1-(2′-*O*-phenylthioxocarbamoyl-3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)cytosine (**4.3b**) was coevaporated with 10 mL of anhydrous toluene twice under reduced pressure. To a solution containing 2.0 g (2.90 mmol) of **4.3b** in 58 mL of anhydrous toluene was added 2.18 mL (7.25 mmol) of trimethylsilylsilane followed by adding dropwise 3.48 mL (3.48 mmol) of triethylborane (1 M in THF) under an argon atmosphere. The reaction was flushed with dry air several times, and the resulting mixture was stirred at room temperature for 30 min. When TLC showed that the reaction was finished, the crude solution was brought to 150 mL with EtOAc, washed with 200 mL of brine and then 200 mL of H_2O, dried over $MgSO_4$, and evaporated under reduced pressure. The crude product was purified by flash column chromatography with eluents (MeOH/CH_2Cl_2, from 1 to 3%) to afford **4.4b** as a white solid: yield 1.21 g (77.4%); silica gel TLC (MeOH/CH_2Cl_2, 1:40) Rf = 0.28; ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.27 (d, 1H, J = 7.6 Hz), 7.94 (d, 2H, J = 7.2 Hz), 7.60−7.57 (m, 4H), 7.51−7.48 (m, 4H), 7.45−7.33 (m, 6H), 6.13 (d, 1H, *J* = 5.6

Hz), 4.47 (s, 1H), 4.23 (d, 1H, *J* = 10 Hz), 3.91 (dd, 1H, *J* = 10, 3.6 Hz), 2.39 (m, 2H) 0.99 (s, 9H); ^{13}C NMR (125.8 MHz, $CDCl_3$) δ 162.3, 145.3, 135.7, 135.6, 133.1, 132.8, 132.6, 130.1, 130.1 129.0, 128.0, 127.9, 127.7, 96.0, 88.4, 78.5, 72.2, 42.0, 26.8, 18.9; HRMS (ESI-TOF) $C_{31}H_{33}N_3O_4SiNa$ $[M + Na]^+$ 562.2138; found 562.2142.

N^6-Benzoyl-9-(2′-deoxy-3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)adenine (4.4c).

N^6-Benzoyl-9-(2′-*O*-phenylthioxocarbamoyl-3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)adenine (**4.3c**) was coevaporated with 10 mL of anhydrous toluene twice under reduced pressure. To a solution containing 1.8 g (2.52 mmol) of **4.3c** in 25 mL of anhydrous toluene was added 1.95 mL (6.30 mmol) of trimethylsilylsilane followed by adding dropwise 3.02 mL (3.02mmol) of triethylborane (1 M in THF) under an argon atmosphere. The reaction was flushed with dry air several times, and the resulting mixture was stirred at room temperature for 30 min. When TLC showed that the reaction was finished, the crude solution was poured into 150 mL of EtOAc, washed with 200 mL of brine and then 200 mL of H_2O, dried over $MgSO_4$, and evaporated under reduced pressure. The crude product was purified by flash column chromatography with eluents ($MeOH/CH_2Cl_2$, from 1 to 3%) to afford product **4.4c** as a white solid: yield 1.12g (78.9%); silica gel TLC ($MeOH/CH_2Cl_2$, 1:40) Rf = 0.3; ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.65 (s, 1H), 8.74 (s, 1H), 8.57 (s, 1H), 8.00 (d, 2H, *J* = 7.2 Hz), 7.56 (d, 2H, *J* = 7.2 Hz), 7.51−7.27 (m, 11H), 6.41 (d, 1H, *J* = 6.8 Hz), 4.58 (s, 1H), 4.12 (d, 1H, *J* = 9.6 Hz), 3.90 (dd, 1H, *J* = 9.6, 3.6 Hz), 2.54 (m, 2H), 0.95 (s, 9H); ^{13}C NMR (125.8 MHz, $CDCl_3$) δ 164.9, 152.3, 151.4, 149.4, 135.5, 135.4, 133.7, 132.6, 132.4, 130.0, 128.6, 127.8, 127.8, 123.3, 84.6, 77.1, 72.1, 41.3, 26.7, 18.7; HRMS (ESI-TOF) $C_{32}H_{33}N_5O_3SiNa$ $[M + Na]^+$ 586.2250; found 586.2251.

N^2-Acetyl-9-(2′-deoxy-3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)guanine (4.4d′).

N^2-Acetyl-9-(2′-*O*-phenylthioxocarbamoyl-3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)guanine (**4.3d′**) was coevaporated twice with 3 mL of anhydrous toluene under reduced pressure and dried overnight under vacuum. To a solution of 0.82 g (1.28 mmol) of 3d′ in 15 mL of anhydrous toluene was added 1.3 mL (3.12 mmol) of tris(trimethylsilyl)silane followed by 1.5 mL of triethylborane (1 M in THF) under an argon atmosphere. The reaction was flushed with dry air several times, and the resulting mixture was stirred at room temperature for 45 min, at which time TLC (CH_2Cl_2/MeOH, 10:1) showed the reaction to be complete. The reaction was quenched with the addition of 3 mL of water. The volatile solution was evaporated under reduced pressure, and the crude product was purified by silica gel column chromatography with eluents (MeOH/CH_2Cl_2, from 0 to 7%) to afford the product **4.4d'** as a white solid: yield 540 mg (83.2%); silica gel TLC (CH_2Cl_2/MeOH, 20:1) Rf = 0.55; ^{1}H NMR (400 MHz, $CDCl_3$) δ 11.88 (s, 1H), 9.09 (s, 1H), 8.11 (s, 1H), 7.43−7.37 (m, 4H), 7.26−7.19 (m, 7 H), 7.17−7.07 (m, 1H), 5.91 (dd, 1H, J = 7.6, 3.2 Hz), 4.38 (s, 1H), 3.82 (d, 1H, J = 8 Hz), 3.66 (dd, 1H, J = 11.6, 4.0 Hz), 2.28−2.26 (m, 1H), 2.09 (s, 3H), 0.84 (s, 9H); ^{13}C NMR(125.8 MHz) δ 171.8, 148.2, 147.3, 138.2, 135.8, 132.9, 130.3, 128.1, 121.3, 84.1, 72.3, 41.3, 27.0, 24.6, 19.1; HRMS (ESI-TOF) $C_{27}H_{31}N_5O_4SiNa$ $[M + Na]^+$ 540.2064; found 540.2064.

1-(2′-Deoxy-α-L-threofuranosyl)thymine (4.5a).

To a cold solution (0−5 °C) of 750 mg (1.664 mmol) of 1-(2′-deoxy-3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)thymine (**4.4a**) in 12 mL of THF was added dropwise

1.9 mL (1.9 mmol) of tetrabutylammonium fluoride (1 M solution in THF), and the mixture was stirred for 2 h at 0−5 °C. The solvent was evaporated under reduced pressure, and the crude product was purified by silica gel column chromatography with eluents ($MeOH/CH_2Cl_2$, from 0 to 12%). 1-(2′-Deoxy-α-L-threofuranosyl)thymine (**4.5a**) was obtained as a white solid: yield 305 mg (86.3%); silica gel TLC (CH_2Cl_2/MeOH, 10:1) Rf = 0.35; ^{1}H NMR (400 MHz, DMSO-d_6) δ 6.89 (s, 1H), 5.18 (dd, 1 H, *J* = 8.4, 4.8 Hz), 3.50 (s, 1H), 3.13 (d, 1H, *J* = 7.6 Hz), 2.91 (dd, 1H, *J* = 10.4, 4.4 Hz), 1.64−1.60 (m, 2H), 1.02 (d, 1H, *J* = 11.6 Hz), 0.91 (s, 3H); ^{13}C NMR (125.8 MHz, DMSO-d_6) *δ* 163.9, 150.5, 137.1, 108.5, 84.6, 76.3, 68.8, 40.07, 12.3; HRMS (ESI-TOF) $C_9H_{12}N_2O_4Na$ $[M + Na]^+$ 235.0797; found 235.0791.

***N*^{4}-Benzoyl-1-(2′-deoxy-α-L-threofuranosyl)cytosine (4.5b).**

To a cold solution (0−5 °C) containing 1.2 g (2.23 mmol) of *N*^{4}-benzoyl-1-(2′-deoxy-3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)cytosine (**4.4b**) in 22 mL of THF was added dropwise 6.69 mL (6.69 mmol) of tetrabutylammonium fluoride (1 M solution in THF). The mixture was stirred for 4 h at 0 °C. The solvent was evaporated under reduced pressure, and the crude product was purified by flash column chromatography with eluents ($MeOH/CH_2Cl_2$, from 2 to 5%). Product **5.5b** was obtained as a white solid: yield 650 mg (97.0%); silica gel TLC ($MeOH/CH_2Cl_2$, 1:20) Rf = 0.31; ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.08 (d, 1H, *J* = 7.6 Hz), 7.88 (d, 2H, *J* = 8.0 Hz), 7.58 (t, 1H, *J* = 7.2 Hz), 7.50−7.43 (m, 3H), 6.00 (d, 1H, *J* = 6.8 Hz), 4.63 (m, 1H), 4.34 (d, 1H, *J* = 10 Hz), 4.11 (dd, 1H, *J* = 10, 3.6 Hz), 2.72 (d, 1H, *J* = 14.8 Hz), 2.47 (m, 1H); ^{13}C NMR (125.8 MHz, $CDCl_3$) *δ* 168.9, 164.6, 157.9, 147.1, 134.7, 134.0, 129.8, 129.0, 97.5, 90.0, 79.4, 71.0, 42.3; HRMS (ESI-TOF) $C_{15}H_{15}N_3O_4Na$ $[M + Na]^+$ 324.0960; found 324.0972.

***N*⁶-Benzoyl-9-(2′-deoxy-α-L-threofuranosyl)adenine (4.5c).**

To a cold solution (0−5 °C) containing 1.3 g (2.30 mmol) of *N*⁶-benzoyl-9-(2′-deoxy-3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)adenine (**4.4b**) in 23 mL of THF was added dropwise 4.6 mL (4.6 mmol) of tetrabutylammonium fluoride (1 M solution in THF). The mixture was stirred for 4 h at 0 °C. The solvent was evaporated under reduced pressure, and the crude product was purified by flash column chromatography with eluents ($MeOH/CH_2Cl_2$, from 2 to 4%) to afford the product **4.5b** as a white solid: yield 678 mg (90.3%); silica gel TLC ($MeOH/CH_2Cl_2$, 1:20) Rf = 0.28; ^1H NMR (400 MHz, $CDCl_3$) δ 8.72 (d, 2H, J = 3.6 Hz), 8.09 (d, 2H, J = 7.6 Hz), 7.66 (t, 1H, J = 7.6 Hz), 7.57 (t, 2H, J = 7.6 Hz), 6.53 (dd, 1H, J = 8.0, 1.6 Hz), 4.62 (t, 1H, J = 4.4 Hz), 4.20 (d, 1H, J = 9.6 Hz), 4.10 (dd, 1H, J = 9.6, 4.0 Hz), 2.78 (m, 1H), 2.49 (d, 1H, J = 14.8 Hz); ^{13}C NMR (125.8 MHz, CD_3OD) δ 168.1, 153.0, 151.0, 145.0, 135.0, 133.9, 129.8, 129.4, 125.1, 86.2, 78.6, 71.3, 41.8; HRMS (ESI-TOF) $C_{16}H_{15}N_5O_3Na$ $[M + Na]^+$ 348.1073; found 348.1076.

***N*²-Acetyl-9-(2′-deoxy-α-L-threofuranosyl)guanine (4.5d′).**

To a cold solution (0−5 °C) of 530 mg (1.02 mmol) of *N*²-acetyl-9-(2′-deoxy-3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl)guanine (**4.4d**) in 12 mL of THF was added dropwise 1.3 mL (1.3 mmol) of tetrabutylammonium fluoride (1 M solution in THF), and the mixture was stirred for 1 h at 0−5 °C. The solvent was evaporated under reduced pressure, and the resulting crude product was purified by silica gel column chromatography with eluents ($MeOH/CH_2Cl_2$, from 0 to 10%) to afford the product **5.5d** as a white solid: yield 220 mg (77%); silica gel TLC (CH_2Cl_2/MeOH, 10:1) Rf = 0.35; ^1H NMR (400 MHz, DMSO-d_6) δ 11.14 (s, 1H), 10.83 (s, 1H), 7.32 (s, 1H), 5.24 (dd, 1H, J = 10, 6 Hz), 4.52 (s, 1H), 3.62 (s, 1H), 3.08−3.07 (m, 2H), 1.77−1.73 (m, 1H), 1.41−1.37 (m, 1H), 1.29 (s, 3H); ^{13}C NMR (125.8 MHz,

DMSO-d_6) δ 173.5, 154.9, 148.2, 147.8, 138.2, 119.9, 83.3, 76.5, 69.0, 40.3,23.8; HRMS (ESI-TOF) C11H13N5O4Na [M + Na]$^+$ 302.0865; found 302.0854.

General Procedure A for the Preparation of 4.6a-d′.

To a suspension of **4.5a–d** in sublimed tetrazole in anhydrous acetonitrile (1.8 equiv, 0.45 M) was added $(BnO)_2PN(i\text{-}Pr)_2$ (1.3 equiv), and the reaction mixture stirred at room temperature for 1 h. The reaction mixture was cooled to 0 °C, treated with 30% H_2O_2 in water (3–10 equiv), and stirred for an additional 30 min. The reaction mixture was then evaporated under reduced pressure and purified by silica gel chromatography.

1-(2′-Deoxy-α-L-threofuranosyl)thymine-3′-dibenzylmonophosphate (4.6a).

General procedure A was used with the following: 1.00 g (4.71 mmol) of **4.5a**, 20 mL (9 mmol) of tetrazole, 1.80 g (5.18 mmol) of $(BnO)_2PN(i\text{-}Pr)_2$, and 2 mL of H_2O_2. Characterization included the following: column chromatography with eluents (EtOAc/MeOH, 100–98/2–95/5%); yield 1.95 g (87%); silica gel TLC (MeOH/CH_2Cl_2, 1:20) Rf = 0.55; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.36–7.26 (m, 10 H), 6.09 (d, 1H, J = 8.6 Hz), 5.04–4.94 (m, 4H), 4.28 (m, 1H), 3.85 (d, 1H, J = 12.6 Hz), 2.48 (m, 1H), 2.16 (d, 1H, J = 15.6 Hz), 1.79 (s, 3H); ^{13}CNMR (125.8 MHz, $CDCl_3$) δ 135.9, 129.1, 128.9, 128.2, 110.4, 85.5, 77.5 (d, $J_{C,P}$ = 6.4 Hz), 76.9, 75.3 (d, $J_{C,P}$ = 7.4 Hz), 70.0 (d, $J_{C,P}$ = 5.9 Hz), 39.7, 12.7; ^{31}P NMR (162 MHz, $CDCl_3$) δ 0.65; HRMS (ESI-TOF) $C_{23}H_{25}N_2O_7PNa$ [M + Na]$^+$ 495.1297; found 495.1299.

***N*^{4}-Benzoyl-1-(2′-deoxy-α-L-threofuranosyl)cytosine-3′-dibenzyl-monophosphate (4.6b)**.

General procedure A was used with the following: 300 mg (1 mmol) of **4.5b**, 4 mL (1.8 mmol) of tetrazole in acetonitrile, 316 μL (1.3 mmol) of $(BnO)_2PN(i\text{-}Pr)_2$, and 2 mL of H_2O_2.

Characterization included the following: silica gel chromatography with eluents ($MeOH/CH_2Cl_2$, from 1.0 to 2.5%); yield 410 mg (73.3%); silica gel TLC ($MeOH/CH_2Cl_2$, 1:40) Rf = 0.33; ^1H NMR (400 MHz, CD_3OD) δ 8.07 (d, 1H, J = 7.2 Hz), 7.93 (d, 1H, J = 6.8 Hz), 7.65–7.61 (m, 1H), 7.53 (m, 3H), 7.32–7.25 (m, 11H), 6.03 (d, 1H, J = 6.4 Hz), 5.05 (s, 1H), 4.96–4.88 (m, 4H), 4.42 (d, 1H, J = 10.8 Hz), 4.14 (d, 1H, J = 10.8 Hz), 2.63–2.49 (m, 2H); ^{13}C NMR (125.8 MHz, CD_3OD) δ 168.9, 164.7, 157.7, 146.0, 136.9 (d, $J_{C,P}$ = 6.2 Hz), 136.7 (d, $J_{C,P}$ = 5.9 Hz), 134.7, 134.0, 129.8, 129.8, 129.7, 129.7, 129.2, 129.1, 129.1, 97.8, 89.8, 78.8 (d, $J_{C,P}$ = 5.4 Hz), 77.5 (d, $J_{C,P}$ = 6.4 Hz), 71.1 (d, $J_{C,P}$ = 6.0 Hz), 70.9 (d, $J_{C,P}$ = 5.9 Hz), 40.8 (d, $J_{C,P}$ = 3.8 Hz); ^{31}P NMR (162 MHz, CD_3OD) δ −1.48; HRMS (ESI-TOF) $C_{29}H_{28}N_3O_7PNa$ $[M + Na]^+$ 584.1563; found 584.1561.

N^6-Benzoyl-9-(2′-deoxy-α-L-threofuranosyl)adenine-3′-dibenzyl-monophosphate (4.6c).

General procedure A was used with the following: 180 mg (0.55 mmol) of **4.5c**, 2.2 mL (0.99 mmol) of tetrazole in acetonitrile, 228 μL (0.72 mmol) of $(BnO)_2PN(i\text{-}Pr)_2$, and 2 mL of H_2O_2. Characterization included the following: silica gel chromatography with eluents ($MeOH/CH_2Cl_2$, from 1 to 2%); yield 242 mg (74.7%); silica gel TLC ($MeOH/CH_2Cl_2$, 1:40) Rf = 0.35; ^1H NMR (400 MHz, CD_3OD) δ 8.61 (s, 1H), 8.34 (s, 1H), 8.00 (d, 2H, J = 7.2 Hz), 7.59 (t, 1H, J = 7.6 Hz), 7.51–7.48 (m, 2H), 7.28–7.20 (m, 8H), 7.15–7.13 (m, 2H), 6.41 (dd, 1H, J = 5.6, 2.8 Hz), 5.10 (s, 1H), 4.88–4.76 (m, 4H), 4.32 (d, 1H, J = 11.2 Hz), 4.09 (m, 1H), 2.70–2.68 (m, 2H); ^{13}C NMR (125.8 MHz, CD_3OD) δ 167.9, 153.0, 152.6, 150.9, 143.3, 136.8 (d, $J_{C,P}$ = 6.0 Hz), 136.7 (d, $J_{C,P}$ = 6.2 Hz), 135.0, 133.9, 129.8, 129.8, 129.7, 129.7, 129.6, 129.4, 129.2, 129.1, 124.9, 86.9, 78.9 (d, $J_{C,P}$ = 5.4 Hz), 76.7 (d, $J_{C,P}$ = 6.0 Hz), 71.0 (d, $J_{C,P}$ = 6.0 Hz), 70.9 (d, $J_{C,P}$ =

5.8 Hz), 49.9, 40.5 (d, $J_{C,P}$ = 4.2 Hz); ^{31}P NMR (162 MHz, CD_3OD) δ −1.37; HRMS (ESI-TOF) $C_{29}H_{28}N_3O_7PNa$ $[M + Na]^+$ 584.1563; found 584.1561.

***N*²-Acetyl-9-(2′-deoxy-α-L-threofuranosyl)guanine-3′-dibenzylmonophosphate (4.6d′).**

General procedure A was used with the following: 600 mg (2.15 mmol) of **4.5d'**, 10 mL (4.5 mmol) of tetrazole, 0.87 mL (2.58 mmol) of $(BnO)_2PN(i\text{-}Pr)_2$, 2 mL of H_2O_2. Characterization included the following: column chromatography with eluents (MeOH/EtOAc, 2–5%); yield 0.81 g (76%); silica gel TLC (CH_2Cl_2/MeOH, 20:1) Rf = 0.65; ^{1}H NMR (400 MHz, $CDCl_3$) δ 12.14 (s, 1H), 10.75 (s, 1H), 7.85 (s, 1H), 7.36–7.25 (m, 10H), 6.02 (dd, 1H, J = 8.4, 4.4 Hz), 5.02–4.89 (m, 4H), 4.16–4.11 (m, 1H), 3.89–3.86 (m, 1H), 2.64 (d, 2H, J = 12 Hz), 2.27(s, 3H), 1.25 (s, 1H); ^{13}C NMR (125.8 MHz, $CDCl_3$) δ 172.9, 155.9, 148.2, 147.7, 128.9, 128.8, 128.0, 121.7, 84.1, 77.1 (d, $J_{C,P}$ = 6.0 Hz), 76.9, 69.9 (d, $J_{C,P}$ = 6.0 Hz), 69.7, 38.8 (d, $J_{C,P}$ = 4.2 Hz), 24.2; ^{31}P NMR (162 MHz, $CDCl_3$) δ −1.20; HRMS (ESI-TOF) $C_{25}H_{26}N_5O_7PNa$ $[M + Na]^+$ 562.1467; found 562.1476.

General Procedure B for the Preparation of 4.7a–d.

Trialkyl monophosphates **4.6a–d** were dissolved in methanol and purged with nitrogen gas three times. The solution was added to 0.1–0.2 mass equivalent of 10% Pd/C and repurged with nitrogen gas, followed by hydrogen gas. The mixture was stirred at room temperature for 1–2 h under hydrogen gas, at which time TLC showed that the reaction was finished. The reaction was purged with nitrogen gas three times and poured on a pad of Celite for filtration. The filtrate was collected and evaporated under reduced pressure to dryness. The crude product was added to 15 mL of 30% NH_4OH for deprotection at 37 °C for 18 h. The mixture was cooled to room temperature, and the solvent was evaporated to dryness. The

residue was resuspended in 3 mL of methanol at 40 °C with stirring, and the solution was added dropwise to 40 mL of acetone to precipitate the product as an ammonium salt. The precipitate was collected by centrifugation at 4400 rpm at room temperature for 15 min, and the resulting pellet was washed twice with 30 mL of acetone and dried under high vacuum. Products **4.7a–d** were obtained as ammonium salts.

1-(2′-Deoxy-α-L-threofuranosyl)thymine-3′-monophosphate (4.7a).

General procedure B was used with the following: 0.39 g (0.79 mmol) of **4.6a**, 8 mL of MeOH, and 130 mg of 10% Pd/C; ^{1}H NMR (400 MHz, CD_3OD) δ 7.57 (s, 1H), 6.12 (dd, 1H, J = 9.2, 6.0 Hz), 4.93 (s, 1H), 4.37 (d, 1H, J = 10.4 Hz), 3.96 (d, 1H, J = 8.4 Hz), 2.63 (m, 1H), 2.30 (d, 1H, J = 15.2 Hz), 1.88 (s, 3H); ^{13}C NMR (125.8 MHz, CD_3OD) δ 165.1, 151.0, 136.8, 109.5, 85.5, 75.5, 47.1, 39.1, 11.1; ^{31}P NMR −0.11; HRMS (ESI-TOF) $C_9H_{13}N_2O_7PNa$ $[M + Na]^+$ 315.0356; found 315.0356.

1-(2′-Deoxy-α-L-threofuranosyl)cytosine-3′-monophosphate (4.7b).

General procedure B was used with the following: 350 mg (0.62 mmol) of **4.6b**, 10 mL of MeOH, and 70 mg of 10% Pd/C; yield 165 mg (96.5%, ε_{271}: 13100); ^{1}H NMR (400 MHz, DMSO-d_6) δ 7.72 (d, 1H, J = 7.6 Hz), 5.91 (dd, 2H, J = 29.2, 7.6 Hz), 4.68 (s, 1H), 4.21 (d, 1H, J = 10.4 Hz), 3.87 (d, 1H, J = 10.4 Hz), 2.42 (m, 1H), 2.16 (d, 1H, J = 15.6 Hz); ^{13}C NMR (125.8 MHz, DMSO-d6) δ 164.7, 154.4, 142.0, 93.7, 85.6, 75.3, 72.9, 40.4; ^{31}P NMR (162 MHz, DMSO-d_6) δ −0.51; HRMS (ESI) $C_8H_{12}N_3O_6PNa_2$ $[M - H + 2Na]^+$ 322.0181; found 322.0185.

9-(2′-Deoxy-α-L-threofuranosyl)adenine-3′-monophosphate (4.7c).

General procedure B was used with the following: 1.3 g (2.22 mmol) of **4.6c**, 15 mL of MeOH, and 200 mg of 10% Pd/C; yield 650 mg (97.8%, ε_{259}: 15200); ^{1}H NMR (400 MHz,

DMSO-d_6) δ 8.34 (s, 1H), 8.15 (s, 1H), 7.22 (s, 1H), 6.32 (d, 1H, J = 7.2 Hz), 4.91 (s, 1H), 4.17 (d, 1H, J = 9.6 Hz), 3.96 (dd, 1H, J = 9.2, 3.6 Hz), 2.75 (m, 1H), 2.52 (m, 1H); ^{13}C NMR (125.8 MHz, DMSO-d_6) δ 155.9, 152.6, 149.4, 139.4, 118.5, 82.5, 74.6 (d, $J_{C,P}$ = 4.2 Hz), 73.2 (d, $J_{C,P}$ = 3.7 Hz), 38.8; ^{31}P NMR (162 MHz, DMSO-d_6) δ −0.20; HRMS (ESI) $C_9H_{12}N_5O_5PH$ $[M + H]^+$ 302.0654; found 302.0668.

9-(2′-Deoxy-α-L-threofuranosyl)guanine-3′-monophosphate (**4.7d**).

In the first step, general procedure B was used with the following: 0.73 g (1.353 mmol, 1 equiv) of **4.6a**, 15 mL of MeOH, and 180 mg of 10% Pd/C); silica gel TLC (CH_2Cl_2/MeOH 10:1) Rf = 0.35. In the second step, the conditions of general procedure B and 5 mL 30% NH_4OH in 37 °C for 18 h were followed: 1H NMR (400 MHz, D_2O) δ 8.17 (s, 1H), 6.31 (d, 1H, J = 9.6 Hz), 5.06 (s, 1H), 4.36 (d, 1H, J = 8.0 Hz), 4.19 (d, 1H, J = 8.4 Hz), 2.86 (s, 2H), 2.68 (d, 2H, J = 8.0 Hz), 2.27 (s, 3H); ^{13}C NMR (125.8 MHz, DMSO-d_6) δ 156.8, 153.6, 150.9, 136.1, 115.9, 81.8, 74.5, 72.1, 48.5, 40.0, 30.5; ^{31}P NMR (162 MHz, DMSO-d_6) δ 0.84; HRMS (ESI) $C_9H_{11}N_5O_6PNa_2$ $[M - H + 2Na]^+$ 362.0242; found 362.0250.

General Procedure C for the Preparation of 4.8a–d.

To a solution containing 1 equiv of **4.7a–d** in anhydrous DMSO was added 5 equiv of trimethylamine followed by 2.5 equiv of 2-methylimidazole, 2.5 equiv of triphenylphosphine, and 2.5 equiv of 2, 2′-dipyridyldisulfide sequentially. The reaction was stirred under a nitrogen atmosphere for 6–8 h at room temperature with monitoring by analytical HPLC. When the HPLC showed that the reaction was finished, the product was precipitated by adding dropwise the reaction mixture to a stirring solution containing 80 mL of acetone, 60 mL of diethyl ether, 5 mL of triethylamine, and 5 mL of saturated $NaClO_4$ in

acetone. The precipitate was collected by centrifugation at 4400 rpm for 15 min at room temperature. The pellet was washed twice with 30 mL of washing solution (acetone/diethyl ether 1:1) and dried under high vacuum to afford white solid products **4.8a–d** as sodium salts.

1-(2′-Deoxy-α-L-threofuranosyl)thymine-3′-phosphor-2-methylimidazolide (4.8a).

General procedure C was used with the following: 100 mg (0.36 mmol) of **4.7a**, 5 mL of anhydrous DMSO, 250 μL (1.80 mmol) of trimethylamine, 236 mg (0.9 mmol) of triphenylphosphine, and 198 mg (0.9 mmol) of 2, 2′-dipyridyldisulfide; product yield 115 mg (97.4%, ε_{267} = 9600); ^{31}P NMR (162 MHz, D_2O) δ −7.81; HRMS (ESI-TOF) $C_{13}H_{16}N_4O_6P$ [M − H]$^-$ 355.0787; found 355.0800.

1-(2′-Deoxy-α-L-threofuranosyl)cytosine-3′-phosphor-2-methylimidazolide (4.8b).

General procedure C was used with the following: 100 mg (0.36 mmol) of **4.7b**, 2 mL of anhydrous DMSO, 250 μL (1.80 mmol) of trimethylamine, 236 mg (0.9 mmol) of triphenylphosphine, and 198 mg (0.9 mmol) of 2, 2′-dipyridyldisulfide; product yield 118 mg (95.4%, ε_{280}: 13100), HPLC retention time; ^{31}P NMR (162 MHz, D_2O) δ −7.99; HRMS (ESI-TOF) $C_{12}H_{15}N_5O_5PNa_2$ [M + 2Na]$^+$ 386.0606; found 386.0617.

9-(2′-Deoxy-α-L-threofuranosyl)adenine-3′-phosphor-2-methylimidazolide (4.8c).

General procedure C was used with the following: 40 mg of 4.7c (0.15 mmol), 1 mL of anhydrous DMSO, 104 μL (0.75 mmol) of trimethylamine, 79 mg (0.3 mmol) of triphenylphosphine, and 66 mg (0.3 mmol) of 2, 2′-dipyridyldisulfide; product yield 45 mg (96.5%, ε_{259}: 15200), HPLC retention time; ^{31}P NMR (162 MHz, D_2O) δ −7.88; HRMS (ESI-TOF) $C_{13}H_{15}N_7O_4P$ [M − H]$^-$ 364.0923; found 364.0936.

9-(2′-Deoxy-α-L-threofuranosyl)guanine-3′-phosphor-2-methylimidazolide (4.8d).

General procedure C was used with the following: 50 mg of **4.7d** (0.13 mmol), 1 mL of anhydrous DMSO, 104 μL (0.75 mmol) of trimethylamine, 79 mg (0.3 mmol) of triphenylphosphine, and 66 mg (0.3 mmol) of 2, 2′-dipyridyldisulfide; product yield 60 mg (93.5%, ε_{259}: 13700); ^{31}P NMR (162 MHz, D_2O) δ −8.15; HRMS (ESI-TOF) $C_{13}H_{16}N_7O_5PNa$ $[M + Na]^+$ 404.0848; found 404.0847.

General Procedure C for the Preparation of 4.9a–d.

To a solution of **4.8a–d** in anhydrous DMF was added 2 equiv of tributylamine and 2 equiv of tributylammonium pyrophosphate. The reaction mixture was stirred under a nitrogen atmosphere for 8–12 h at room temperature with monitoring by analytical HPLC. When HPLC showed that the reaction was finished, the reaction mixture was added dropwise to a stirring solution containing 30 mL of acetone and 5 mL of saturated $NaClO_4$ in acetone. The precipitate was collected by centrifugation at 4400 rpm for 15 min at room temperature and dried under vacuum for 1 h. The crude precipitate was dissolved in 3 mL of 0.1 M triethylammonium acetate buffer and purified by a semipreparative HPLC column. Fractions containing triphosphates were collected, concentrated under reduced pressure, pH adjusted with triethylamine to 7.0, and lyophilized to afford the product as a triethylammonium salt. The solid product was resuspended in 3 mL of methanol and was added dropwise to a solution containing 40 mL of acetone and 5 mL of saturated $NaClO_4$ in acetone. The solution was centrifuged at 4400 rpm for 15 min at room temperature. The supernatant was discarded, and the pellet was washed with 30 mL of acetone and dried under vacuum for 1

h. The resulting white solid was dissolved in RNase-free water containing 10 mM Tris pH 8.0 to afford the **4.9a–d** solution.

1-(2′-Deoxy-α-L-threofuranosyl)thymine-3′-triphosphate (4.9a).

Product yield after HPLC purification was: 18.3 mg (86%, ε_{267}: 9600); ^{1}H NMR (600 MHz, D_2O) δ 7.76 (s, 1H), 6.19 (d, 1H, *J* = 7.8 Hz), 5.09 (s, 1H), 4.45 (d, 1H, *J* = 12.6 Hz), 4.04 (d, 1H, *J* = 10.8 Hz), 2.73 (m, 1H), 2.23 (d, 1H, *J* = 15.6 Hz); ^{13}C NMR (150 MHz, D_2O): *δ* 169.5, 154.6, 141.0, 113.7, 88.5, 78.3 (d, *J* = 3.5 Hz), 77.6 (d, *J* = 4.1 Hz), 41.3, 14.5; ^{31}P NMR (162 MHz, D_2O): *δ* −5.80 (d, *J* = 19.4 Hz), −10.86 (d, *J* = 19.4 Hz), −20.59 (t, *J* = 19.4 Hz); HRMS (ESI-TOF) $C_9H_{12}N_2O_{13}P_3Na_2$ [M − 3H + 2Na]− 494.9348; found 494.9351.

1-(2′-Deoxy-α-L-threofuranosyl)cytosine-3′-triphosphate (4.9b).

Product yield after HPLC purification was: 15.3 mg (78%, ε_{280}: 13100); ^{1}H NMR (400 MHz, D_2O) *δ* 7.99 (d, 1H, *J* = 7.6 Hz), 6.25 (d, 1H, *J* = 8.0 Hz), 6.13 (d, 1H, *J* = 7.2 Hz), 5.13 (s, 1H), 4.55 (d, 1H, *J* = 10.4 Hz), 4.12 (dd, 1H, *J* = 10.8, 2.8 Hz), 2.74 (m, 1H), 2.41 (d, 1H, *J* = 15.6 Hz); 13C NMR (150 MHz, D_2O) δ 166.4, 157.9, 142.6, 96.0, 88.7, 75.9 (d, J = 5.6 Hz), 75.1 (d, J = 6.5 Hz), 39.2 (d, *J* = 5.9 Hz); ^{31}P NMR (162 MHz, D_2O) *δ* −5.46 (d, *J* = 19.4 Hz), −11.00 (d, *J* = 19.4 Hz), −20.89 (t, *J* = 19.4 Hz); HRMS (ESI-TOF) $C_8H_{11}N_3O_{12}P_3Na_4$ $[M - 3H + 4Na]^+$ 525.9146; found 525.9126.

9-(2′-Deoxy-α-L-threofuranosyl)adenine-3′-triphosphate (4.9c).

Product yield after HPLC purification was: 7.4 mg (56%, ε_{259}: 15200); ^{1}H NMR (600 MHz, D_2O) *δ* 8.48 (s, 1H), 8.21 (s, 1H), 6.43 (dd, 1H, *J* = 7.8, 1.2 Hz), 5.20 (d, 1H, *J* = 8.4 Hz), 4.38 (d, 1H, *J* = 10.2 Hz), 4.17 (dd, 1H, *J* = 10.2, 3.6 Hz), 2.89 (m, 1H), 2.77 (d, 1H, *J* = 15.6 Hz); ^{13}C NMR (150 MHz, D_2O) *δ* 155.6, 152.7, 148.8, 140.9, 118.6, 83.8, 75.7 (d, *J* = 4.7 Hz), 75.3

(d, *J* = 5.4 Hz), 38.7; ^{31}P NMR (162 MHz, D_2O) δ −8.25 (d, *J* = 19.4 Hz), −10.98 (d, *J* = 19.4 Hz), −21.50 (t, *J* = 19.4 Hz); HRMS (ESI-TOF) $C_9H_{13}N_5O_{11}P_3$ $[M - H]^-$ 459.9825; found 459.9842.

9-(2′-Deoxy-α-L-threofuranosyl)guanine-3′-triphosphate (4.9d).

Product yield after HPLC purification was: 9.3 mg (62%, ε_{259}: 15400); ^{1}H NMR (400 MHz, D_2O) δ 8.17 (s, 1H), 6.27 (d, 1H, *J* = 7.2 Hz), 5.26 (m, 1H), 4.42 (d, 1H, *J* = 10.4 Hz), 4.20 (dd, 1H, *J* = 10.4, 2.8 Hz), 2.94 (m, 1H), 2.73 (m, 1H); ^{13}C NMR (150 MHz, D_2O) δ 161.4, 156.3, 151.4, 138.0, 116.5, 83.2, 75.3 (d, *J* = 4.8 Hz), 75.2 (d, *J* = 5.6 Hz), 38.4 (d, *J* = 4.8 Hz); ^{31}P NMR (162 MHz, D_2O) δ −9.75 (d, *J* = 19.4 Hz), −11.27 (d, *J* = 19.4 Hz), −22.16 (t, *J* = 19.4 Hz); HRMS (ESI-TOF) $C_9H_{11}N_5O_{12}P_3Na_4$ $[M - 3H + 4Na]^+$ 565.9208; found 565.9213.

Primer Extension.

Primer extension reactions were performed in a final volume of 10 μL using 10 pmol of primer labeled with a 5′-IR680 dye and 15 pmol of template. Each reaction contained a primer−template complex, 1× 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], 1 μL of Kod-RI polymerase, 100 μL of each TNA triphosphate, and 1 mM $MnCl_2$. Reactions were incubated for 1 h at 55 °C, quenched with stop buffer (8 M urea, 45 mM EDTA), and analyzed by 20% denaturing urea PAGE. Gels were imaged using a Li-Cor imager.

Protein Crystallization.

Kod-RI for X-ray crystallography was expressed and purified using a previously published protocol.[15] An initial binary complex was prepared by incubating Kod-RI (5 mg mL−1) with 1.5 molar equivalents of the primer−template duplex (primer, 5′-CGCGAACTGCG-3′ and template, 5′-AAATTCGCAGTTCGCG-3′) at 37 °C for 30 min in a buffer

supplemented with 20 mM $MgCl_2$. Five molar excess of the 2′-deoxy tATP monomer was added to the binary complex, and the solution was incubated at 37 °C for 30 min. The resulting binary complex was further incubated with 5 M excess tTTP at 37 °C for another 30 min. The ternary complex initially cocrystallized in 0.1 M 2-(N-morpholino)ethanesulfonic acid pH 6.0, 0.15 M ammonium sulfate, and 15% polyethylene glycol 4000. Crystals were further optimized in 24-well trays over a range of MES pH and PEG concentration. Trays were stored in the dark at room temperature. Crystals typically grew between 1 and 4 weeks.

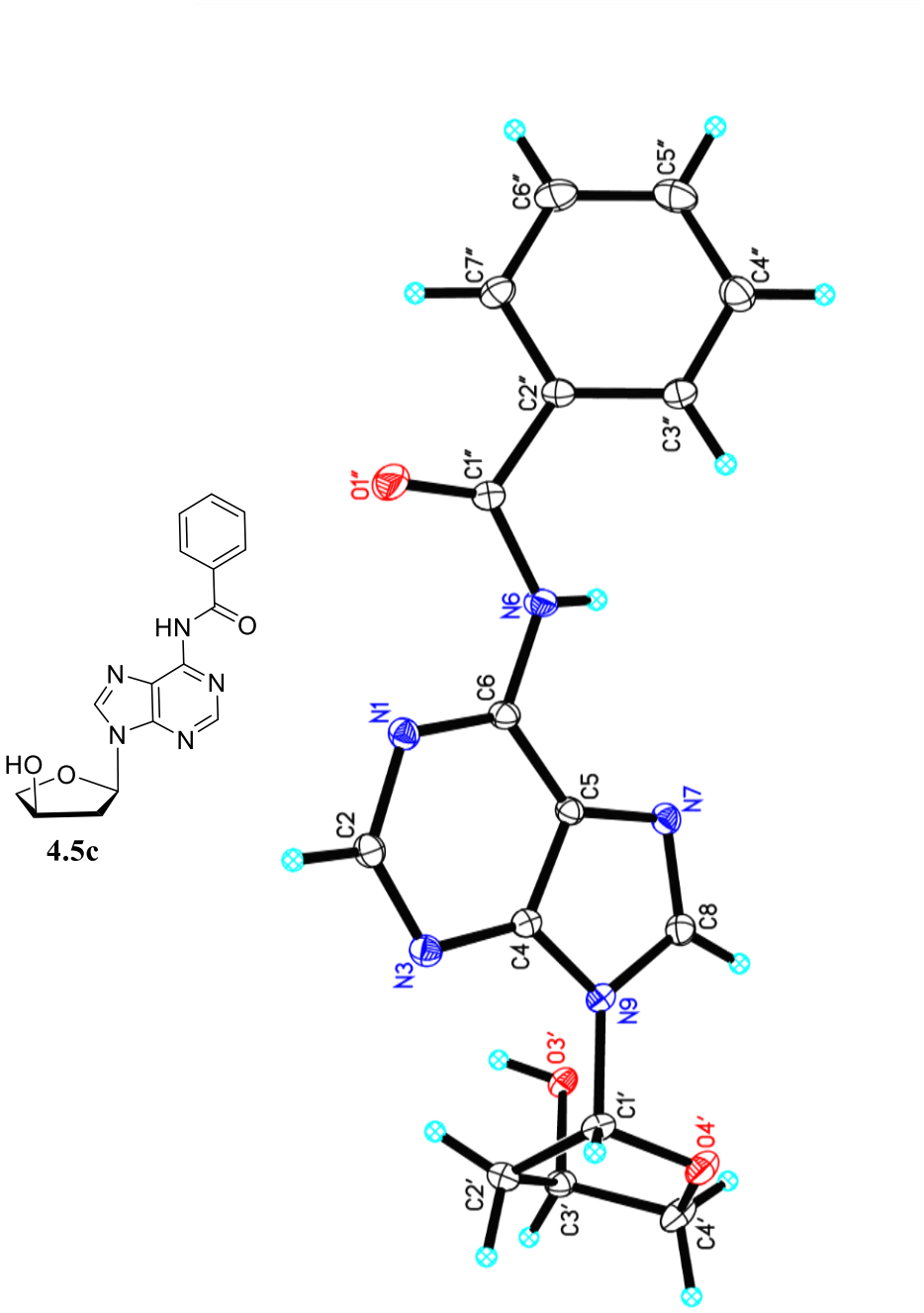


Figure 4-4. ORTEP representation of 6-benzoyl-2′-deoxy-α-L-threofuranosyl adenosine nucleoside 4.5c.

ORTEP representation of compound **4.5c** with thermal ellipsoids at the 50% probability level. The saturated compound **4.5c** was dissolved in methanol with slow crystallization via solvent evaporation.

Table 1. Crystal data and structure refinement for jcc5.

Identification code	jcc5	
Empirical formula	C_{16} H_{15} N_5 O_3	
Formula weight	325.33	
Temperature	88(2) K	
Wavelength	0.71073 Å	
Crystal system	Monoclinic	
Space group	$P2_1$	
Unit cell dimensions	a = 6.817(4) Å	α = 90°.
	b = 7.077(4) Å	β = 101.960(7)°.
	c = 15.985(9) Å	γ = 90°.
Volume	754.4(8) $Å^3$	
Z	2	
Density (calculated)	1.432 Mg/m^3	
Absorption coefficient	0.103 mm^{-1}	
F(000)	340	
Crystal color	colorless	
Crystal size	0.596 x 0.251 x 0.138 mm^3	
Theta range for data collection	3.884 to 30.601°	
Index ranges	$-9 \le h \le 9$, $-10 \le k \le 10$, $-22 \le l \le 22$	
Reflections collected	11727	
Independent reflections	4580 [R(int) = 0.0184]	
Completeness to theta = 25.500°	99.5 %	
Absorption correction	Semi-empirical from equivalents	
Max. and min. transmission	1.0000 and 0.9635	
Refinement method	Full-matrix least-squares on F^2	
Data / restraints / parameters	4580 / 1 / 277	
Goodness-of-fit on F^2	1.048	
Final R indices [I>2sigma(I) = 4425 data]	R1 = 0.0311, wR2 = 0.0791	
R indices (all data, 0.70 Å)	R1 = 0.0324, wR2 = 0.0801	
Largest diff. peak and hole	0.428 and -0.164 $e.Å^{-3}$	

Table 2. Atomic coordinates (x 10^4) and equivalent isotropic displacement parameters ($Å^2$x 10^3) for jcc5. U(eq) is defined as one third of the trace of the orthogonalized U^{ij} tensor.

	x	y	z	U(eq)
O(1")	3192(2)	5573(2)	2094(1)	24(1)
O(3')	9646(2)	2828(2)	7037(1)	14(1)
O(4')	11061(2)	6955(2)	7110(1)	16(1)
N(1)	4149(2)	6607(2)	3820(1)	12(1)
N(3)	5228(2)	7171(2)	5336(1)	13(1)
N(6)	6289(2)	5304(2)	2979(1)	15(1)
N(7)	9457(2)	5259(2)	4623(1)	13(1)
N(9)	8716(2)	6318(2)	5853(1)	11(1)
C(1')	8997(2)	7014(2)	6740(1)	13(1)
C(1")	4970(2)	5214(2)	2203(1)	13(1)
C(2')	7969(2)	5773(2)	7299(1)	14(1)
C(2)	3900(2)	7159(2)	4599(1)	13(1)
C(2")	5904(2)	4627(2)	1474(1)	12(1)
C(3')	9567(2)	4286(2)	7637(1)	14(1)
C(3")	7891(2)	5036(2)	1449(1)	14(1)
C(4')	11475(2)	5444(2)	7738(1)	18(1)
C(4)	7012(2)	6535(2)	5237(1)	11(1)
C(4")	8619(2)	4579(2)	723(1)	18(1)
C(5)	7496(2)	5883(2)	4480(1)	11(1)
C(5")	7379(3)	3690(3)	32(1)	22(1)
C(6)	5941(2)	5937(2)	3753(1)	12(1)
C(6")	5405(3)	3271(2)	61(1)	20(1)
C(7")	4666(2)	3745(2)	777(1)	16(1)
C(8)	10118(2)	5563(2)	5448(1)	13(1)

Table 3. Bond lengths [Å] and angles [°] for jcc5.

O(1")-C(1")	1.215(2)
O(3')-C(3')	1.4177(18)
O(3')-H(3')	0.86(3)
O(4')-C(1')	1.4093(19)
O(4')-C(4')	1.4537(19)
N(1)-C(6)	1.3349(19)
N(1)-C(2)	1.3481(19)
N(3)-C(2)	1.3294(19)
N(3)-C(4)	1.3371(19)
N(6)-C(1")	1.3739(18)
N(6)-C(6)	1.3827(19)
N(6)-H(6)	0.79(3)
N(7)-C(8)	1.3199(19)
N(7)-C(5)	1.3809(19)
N(9)-C(4)	1.3676(18)
N(9)-C(8)	1.3692(19)
N(9)-C(1')	1.4747(19)
C(1')-C(2')	1.523(2)
C(1')-H(1A)	1.00(2)
C(1")-C(2")	1.498(2)
C(2')-C(3')	1.531(2)
C(2')-H(2A)	0.96(2)
C(2')-H(2B)	0.97(3)
C(2)-H(2)	0.93(2)
C(2")-C(3")	1.394(2)
C(2")-C(7")	1.398(2)
C(3')-C(4')	1.517(2)
C(3')-H(3A)	0.93(2)
C(3")-C(4")	1.391(2)
C(3")-H(3")	0.99(2)
C(4')-H(4A)	0.99(2)
C(4')-H(4B)	0.97(2)
C(4)-C(5)	1.3971(19)
C(4")-C(5")	1.395(2)

C(4")-H(4")	1.00(2)
C(5)-C(6)	1.4014(18)
C(5")-C(6")	1.388(3)
C(5")-H(5")	1.00(2)
C(6")-C(7")	1.385(2)
C(6")-H(6")	0.97(2)
C(7")-H(7")	0.95(2)
C(8)-H(8)	0.97(2)

C(3')-O(3')-H(3')	108.1(17)
C(1')-O(4')-C(4')	110.22(11)
C(6)-N(1)-C(2)	117.99(12)
C(2)-N(3)-C(4)	111.18(12)
C(1")-N(6)-C(6)	128.38(13)
C(1")-N(6)-H(6)	115.2(17)
C(6)-N(6)-H(6)	116.3(17)
C(8)-N(7)-C(5)	103.43(12)
C(4)-N(9)-C(8)	105.96(12)
C(4)-N(9)-C(1')	125.21(12)
C(8)-N(9)-C(1')	128.35(12)
O(4')-C(1')-N(9)	108.25(11)
O(4')-C(1')-C(2')	106.11(12)
N(9)-C(1')-C(2')	112.77(12)
O(4')-C(1')-H(1A)	108.9(12)
N(9)-C(1')-H(1A)	106.6(12)
C(2')-C(1')-H(1A)	114.0(12)
O(1")-C(1")-N(6)	124.56(13)
O(1")-C(1")-C(2")	121.22(13)
N(6)-C(1")-C(2")	114.22(13)
C(1')-C(2')-C(3')	102.90(12)
C(1')-C(2')-H(2A)	114.7(14)
C(3')-C(2')-H(2A)	112.4(15)
C(1')-C(2')-H(2B)	109.5(16)
C(3')-C(2')-H(2B)	106.0(16)
H(2A)-C(2')-H(2B)	111(2)
N(3)-C(2)-N(1)	128.89(14)

N(3)-C(2)-H(2)	117.3(12)
N(1)-C(2)-H(2)	113.8(12)
C(3")-C(2")-C(7")	119.94(13)
C(3")-C(2")-C(1")	122.53(12)
C(7")-C(2")-C(1")	117.40(14)
O(3')-C(3')-C(4')	108.25(13)
O(3')-C(3')-C(2')	112.65(12)
C(4')-C(3')-C(2')	101.42(13)
O(3')-C(3')-H(3A)	110.9(15)
C(4')-C(3')-H(3A)	111.5(14)
C(2')-C(3')-H(3A)	111.7(14)
C(4")-C(3")-C(2")	119.55(13)
C(4")-C(3")-H(3")	119.7(13)
C(2")-C(3")-H(3")	120.8(13)
O(4')-C(4')-C(3')	106.53(12)
O(4')-C(4')-H(4A)	107.8(15)
C(3')-C(4')-H(4A)	113.2(14)
O(4')-C(4')-H(4B)	108.9(13)
C(3')-C(4')-H(4B)	111.0(14)
H(4A)-C(4')-H(4B)	109.2(19)
N(3)-C(4)-N(9)	127.60(13)
N(3)-C(4)-C(5)	126.75(12)
N(9)-C(4)-C(5)	105.61(12)
C(3")-C(4")-C(5")	120.21(16)
C(3")-C(4")-H(4")	118.3(13)
C(5")-C(4")-H(4")	121.5(13)
N(7)-C(5)-C(4)	110.86(12)
N(7)-C(5)-C(6)	133.35(12)
C(4)-C(5)-C(6)	115.77(13)
C(6")-C(5")-C(4")	120.17(15)
C(6")-C(5")-H(5")	119.2(14)
C(4")-C(5")-H(5")	120.6(14)
N(1)-C(6)-N(6)	121.12(12)
N(1)-C(6)-C(5)	119.40(12)
N(6)-C(6)-C(5)	119.48(13)
C(7")-C(6")-C(5")	119.83(14)

C(7")-C(6")-H(6")	120.4(14)
C(5")-C(6")-H(6")	119.7(14)
C(6")-C(7")-C(2")	120.29(15)
C(6")-C(7")-H(7")	121.3(14)
C(2")-C(7")-H(7")	118.4(14)
N(7)-C(8)-N(9)	114.13(13)
N(7)-C(8)-H(8)	125.6(12)
N(9)-C(8)-H(8)	120.3(12)

Table 4. Anisotropic displacement parameters ($Å^2 x\ 10^3$) for jcc5. The anisotropic displacement factor exponent takes the form: $-2\ ^2[\ h^2\ a^{*2} U^{11} + ...\ + 2\ h\ k\ a^*\ b^*\ U^{12}\ ]$

	U^{11}	U^{22}	U^{33}	U^{23}	U^{13}	U^{12}
O(1")	13(1)	41(1)	18(1)	-8(1)	1(1)	4(1)
O(3')	13(1)	12(1)	15(1)	-2(1)	1(1)	-2(1)
O(4')	16(1)	17(1)	14(1)	2(1)	-2(1)	-5(1)
N(1)	12(1)	13(1)	13(1)	0(1)	3(1)	1(1)
N(3)	13(1)	12(1)	13(1)	-1(1)	3(1)	1(1)
N(6)	12(1)	24(1)	10(1)	0(1)	1(1)	6(1)
N(7)	11(1)	15(1)	14(1)	2(1)	3(1)	2(1)
N(9)	11(1)	11(1)	11(1)	0(1)	1(1)	-1(1)
C(1')	15(1)	12(1)	11(1)	-1(1)	1(1)	-1(1)
C(1")	14(1)	13(1)	11(1)	-1(1)	2(1)	0(1)
C(2')	16(1)	15(1)	12(1)	-1(1)	4(1)	1(1)
C(2)	12(1)	12(1)	14(1)	0(1)	4(1)	2(1)
C(2")	14(1)	11(1)	11(1)	0(1)	2(1)	2(1)
C(3')	16(1)	14(1)	10(1)	1(1)	2(1)	-2(1)
C(3")	15(1)	14(1)	12(1)	0(1)	1(1)	1(1)
C(4')	18(1)	15(1)	16(1)	1(1)	-4(1)	-3(1)
C(4)	13(1)	9(1)	10(1)	1(1)	2(1)	-1(1)
C(4")	17(1)	22(1)	17(1)	0(1)	6(1)	2(1)
C(5)	11(1)	12(1)	10(1)	1(1)	3(1)	1(1)
C(5")	28(1)	25(1)	15(1)	-3(1)	8(1)	3(1)
C(6)	13(1)	13(1)	10(1)	1(1)	2(1)	1(1)
C(6")	25(1)	20(1)	14(1)	-6(1)	2(1)	0(1)
C(7")	16(1)	16(1)	15(1)	-3(1)	1(1)	-1(1)
C(8)	12(1)	14(1)	13(1)	2(1)	2(1)	0(1)

Table 5. Hydrogen coordinates (x 10^4) and isotropic displacement parameters ($Å^2$x 10^3) for jcc5.

	x	y	z	U(eq)
H(3')	8480(40)	2320(40)	6903(15)	25(6)
H(6)	7370(40)	4880(40)	2985(15)	23(6)
H(1A)	8540(30)	8350(30)	6704(13)	13(5)
H(2A)	6720(30)	5220(40)	7007(14)	23(5)
H(2B)	7780(40)	6490(40)	7793(16)	32(6)
H(2)	2610(30)	7560(30)	4608(12)	9(4)
H(3A)	9410(30)	3790(30)	8160(14)	17(5)
H(3")	8770(40)	5690(40)	1931(15)	25(6)
H(4A)	11880(40)	6030(40)	8309(15)	24(6)
H(4B)	12580(30)	4680(30)	7623(14)	20(5)
H(4")	10030(40)	4950(40)	704(14)	23(5)
H(5")	7890(30)	3350(40)	-488(15)	24(6)
H(6")	4580(40)	2590(40)	-407(14)	25(6)
H(7")	3310(40)	3510(40)	804(15)	21(5)
H(8)	11460(30)	5310(30)	5763(13)	12(5)

Table 6. Torsion angles [°] for jcc5.

C(4')-O(4')-C(1')-N(9)	-108.16(13)
C(4')-O(4')-C(1')-C(2')	13.14(15)
C(4)-N(9)-C(1')-O(4')	-165.27(12)
C(8)-N(9)-C(1')-O(4')	5.70(19)
C(4)-N(9)-C(1')-C(2')	77.65(17)
C(8)-N(9)-C(1')-C(2')	-111.39(16)
C(6)-N(6)-C(1")-O(1")	-5.5(3)
C(6)-N(6)-C(1")-C(2")	174.32(14)
O(4')-C(1')-C(2')-C(3')	-30.58(14)
N(9)-C(1')-C(2')-C(3')	87.76(14)
C(4)-N(3)-C(2)-N(1)	0.2(2)
C(6)-N(1)-C(2)-N(3)	1.4(2)
O(1")-C(1")-C(2")-C(3")	149.34(16)
N(6)-C(1")-C(2")-C(3")	-30.53(19)
O(1")-C(1")-C(2")-C(7")	-26.5(2)
N(6)-C(1")-C(2")-C(7")	153.64(14)
C(1')-C(2')-C(3')-O(3')	-80.41(14)
C(1')-C(2')-C(3')-C(4')	35.10(14)
C(7")-C(2")-C(3")-C(4")	0.7(2)
C(1")-C(2")-C(3")-C(4")	-175.05(14)
C(1')-O(4')-C(4')-C(3')	9.93(16)
O(3')-C(3')-C(4')-O(4')	90.50(14)
C(2')-C(3')-C(4')-O(4')	-28.21(15)
C(2)-N(3)-C(4)-N(9)	-178.76(13)
C(2)-N(3)-C(4)-C(5)	-1.5(2)
C(8)-N(9)-C(4)-N(3)	177.83(14)
C(1')-N(9)-C(4)-N(3)	-9.5(2)
C(8)-N(9)-C(4)-C(5)	0.07(14)
C(1')-N(9)-C(4)-C(5)	172.71(12)
C(2")-C(3")-C(4")-C(5")	-1.0(2)
C(8)-N(7)-C(5)-C(4)	-0.75(15)
C(8)-N(7)-C(5)-C(6)	-178.72(15)
N(3)-C(4)-C(5)-N(7)	-177.36(13)
N(9)-C(4)-C(5)-N(7)	0.43(15)

N(3)-C(4)-C(5)-C(6)	1.0(2)
N(9)-C(4)-C(5)-C(6)	178.79(12)
C(3")-C(4")-C(5")-C(6")	0.5(3)
C(2)-N(1)-C(6)-N(6)	178.47(13)
C(2)-N(1)-C(6)-C(5)	-1.84(19)
C(1")-N(6)-C(6)-N(1)	-1.4(2)
C(1")-N(6)-C(6)-C(5)	178.95(14)
N(7)-C(5)-C(6)-N(1)	178.67(14)
C(4)-C(5)-C(6)-N(1)	0.77(19)
N(7)-C(5)-C(6)-N(6)	-1.6(2)
C(4)-C(5)-C(6)-N(6)	-179.54(13)
C(4")-C(5")-C(6")-C(7")	0.3(3)
C(5")-C(6")-C(7")-C(2")	-0.6(2)
C(3")-C(2")-C(7")-C(6")	0.1(2)
C(1")-C(2")-C(7")-C(6")	176.08(14)
C(5)-N(7)-C(8)-N(9)	0.81(16)
C(4)-N(9)-C(8)-N(7)	-0.58(16)
C(1')-N(9)-C(8)-N(7)	-172.91(13)

Table 7. Hydrogen bonds for jcc5 [Å and °].

D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
O(3')-H(3')...O(1")#1	0.86(3)	2.48(3)	3.0553(19)	125(2)
O(3')-H(3')...N(1)#1	0.86(3)	1.99(3)	2.802(2)	158(2)
N(6)-H(6)...O(4')#2	0.79(3)	2.35(3)	3.001(2)	140(2)

Symmetry transformations used to generate equivalent atoms:
#1 -x+1,y-1/2,-z+1 #2 -x+2,y-1/2,-z+1

Table 4-1. Crystal data and structure refinement for 4.5c.

References

1. Eschenmoser, A. *Science* **2000,** *290*, 1347−1351.
2. Eschenmoser, A. *Science* **1999**, *284*, 2118−2124.
3. Schöning, K.-U.; Scholz, P.; X, W.; Guntha, S.; Delgado, G.; Krishnamurthy, R.; Eschenmoser, A. *Helv. Chim. Acta* **2002**, *85*, 4111−4153.
4. Anosova, I.; Kowal, E. A.; Dunn, M. R.; Chaput, J. C.; Van Horn, W. D.; Egli, M. *Nucleic Acids Res.* **2016**, *44*, 1007−1021.
5. Larsen, A. C.; Dunn, M. R.; Hatch, A.; Sau, S. P.; Youngbull, C.; Chaput, J. C. *Nat. Commun.* **2016**, *7*, 11235.
6 Dunn, M. R.; Otto, C.; Fenton, K. E.; Chaput, J. C. *ACS Chem. Biol.* **2016**, *11*, 1210−1219.
7. Yu, H.; Zhang, S.; Chaput, J. C. *Nat. Chem.* **2012**, *4*, 183−187.
8. Dunn, M. R.; Chaput, J. C. *ChemBioChem* **2016**, *17*, 1804−1808.
9. Mei, H.; Liao, J. Y.; Jimenez, R. M.; Wang, Y.; Bala, S.; McCloskey, C.; Switzer, C.; Chaput, J. C. *J. Am. Chem. Soc.* **2018**, *140*, 5706−5713.
10. Culbertson, M. C.; Temburnikar, K. W.; Sau, S. P.; Liao, J.-Y.; Bala, S.; Chaput, J. C. *Bioorg. Med. Chem. Lett.* **2016**, *26*, 2418−2421.
11. Dunn, M. R.; Jimenez, R. M.; Chaput, J. C. *Nat. Rev. Chem.* **2017**, *1*, 0076.
12. Zhou, J.; Rossi, J. *Nat. Rev. Drug Discovery* **2017**, *16*, 181−202.
13. Nikoomanzar, A.; Dunn, M. R.; Chaput, J. C. *Anal. Chem.* **2017**, *89*, 12622−12625.
14. Chim, N.; Shi, C.; Sau, S. P.; Nikoomanzar, A.; Chaput, J. C. *Nat. Commun.* **2017**, *8*, 1810.
15. Kielkowski, P.; Fanfrlik, J.; Hocek, M. *Angew. Chem., Int. Ed.* **2014,** *53*, 7552−7555.
16. Chen, T.; Romesberg, F. E. *FEBS Lett.* **2014**, *588*, 219−229.
17. Loakes, D.; Holliger, P. *Chem. Commun.* **2009**, 4619−4631.
18. Kool, E. T. *Annu. Rev. Biochem.* **2002**, *71*, 191−219.
19. Houlihan, G.; Arangundy-Franklin, S.; Holliger, P. *Acc. Chem. Res.* **2017**, *50*, 1079−1087.
20. Aschenbrenner, J.; Marx, A. *Curr. Opin. Biotechnol*. **2017**, *48*, 187−195.
21. Doublie, S.; Tabor, S.; Long, A. M.; Richardson, C. C.; Ellenberger, T. *Nature* **1998**, *391*, 251−258.

22. Kropp, H. M.; Betz, K.; Wirth, J.; Diederichs, K.; Marx, A. *PLoS One* **2017**, 12, e0188005.
23 Sau, S. P.; Fahmi, N. E.; Liao, J.-Y.; Bala, S.; Chaput, J. C. *J. Org. Chem.* **2016**, *81*, 2302−2307.
24. Dumbre, S. G.; Jang, M.-Y.; Herdewijn, P. *J. Org. Chem.* **2013**, *78*, 7137−7144.
25. Zou, K.; Horhota, A.; Yu, B.; Szostak, J. W.; McLaughlin, L. W. *Org. Lett.* **2005**, *7*, 1485−1487.
26. Barton, D. H. R.; McCombie, S. W. *J. Chem. Soc., Perkin Trans.* **1975**, 1574−1585.
27. Robins, M. J.; Wilson, J. S. *J. Am. Chem. Soc.* **1981**, *103*, 932−933.
28. Robins, M. J.; Madej, D.; Hansske, F.; Wilson, J. S.; Gosselin, G.; Bergogne, M. C.; Imbach, J. L.; Balzarini, J.; Declercq, E. Nucleic-Acid Related-Compounds 0.53. *Can. J. Chem.* **1988**, *66*, 1258−1262.
29. Barton, D. H. R.; Jang, D. O.; Jaszberenyi, J. C. *Tetrahedron Lett*. **1990**, *31*, 4681−4684.
30. Oba, M.; Nishiyama, K. *Tetrahedron* **1994**, 50, 10193−10200.
31. Sau, S. P.; Chaput, J. C. *Org. Lett.* **2017**, *19*, 4379−4382.
32. Burgess, K.; Cook, D. *Chem. Rev.* **2000**, *100*, 2047−2060.
33. Hollenstein, M. *Molecules* **2012**, *17*, 13569−13591.
34. Kore, A. R.; Srinivasan, B. *Curr. Org. Synth*. **2013**, *10*, 903−934.
35. Bala, S.; Liao, J. Y.; Mei, H.; Chaput, J. C. *J. Org. Chem.* **2017**, 82, 5910−5916.

Chapter 5

Conclusions and Future Perspectives

The research accomplishments presented in this thesis have described optimal synthetic strategies for α-L-threofuranosyl (TNA) monomers, nucleoside phosphoramdites, and nucleoside triphosphates synthesis in scalable scale. Those molecular building blocks are used as immediate substrates to construct the TNA oligonucleotides chemically or enzymatically for functional property evaluation and therapeutic application toward the selected biological targets.

In chapter 1, we have developed a ten-step synthetic strategy for the synthesis of TNA nucleoside monomers and phosphoramidites in quantitative scale. Key challenges have been resolved and purification steps were minimized to increase the yield of product synthesis. Several TNA nucleoside intermediates were also prepared to be used as the immediate precursors for the following modified TNA nucleoside and nucleotide synthesis. In chapter 2, stepwise synthetic strategy was developed to construct the α-L-threofuranosyl nucleoside monophosphates and nucleoside triphosphates. This method overcame the challenge of poor phosphorylation efficiency on 3'-OH of TNA nucleoside by introducing the combined phosphitylation-oxidation reaction. The resulting TNA nucleoside monophosphate precursors were subjected to a silica gel chromatography and hydrogenation reaction to generated high purity of TNA nucleoside monophosphates that preclude the HPLC purification. Then the stepwise high yield conversion from α-L-threofuranosyl nucleoside 3'-monophosphates to nucleoside 3'-monophosphate derivatives to nucleoside 3'-triphosphates facilitated a throughput HPLC chromatography of the final products. This synthetic strategy was extensively applied to the 2'-deoxy TNA nucleoside triphosphates

(sugar-modified, chapter 4) and 7-deaza-7-phenyl TNA nucleoside triphosphates (nucleobase-modified, chapter 3) synthesis.

In chapter 3, we have established the faithful system for replicating the four-nucleotide TNA sequence libraries in vitro by introducing the 7-deaza-7-phenyl TNA nucleoside triphosphates. This molecule minimized the G to C mutation in TNA replication cycle (DNA->TNA->DNA) using engineered polymerases. The system was demonstrated to be able to evolve the TNA folded sequences with nano-molar binding affinity against the HIV-RT in-vitro.

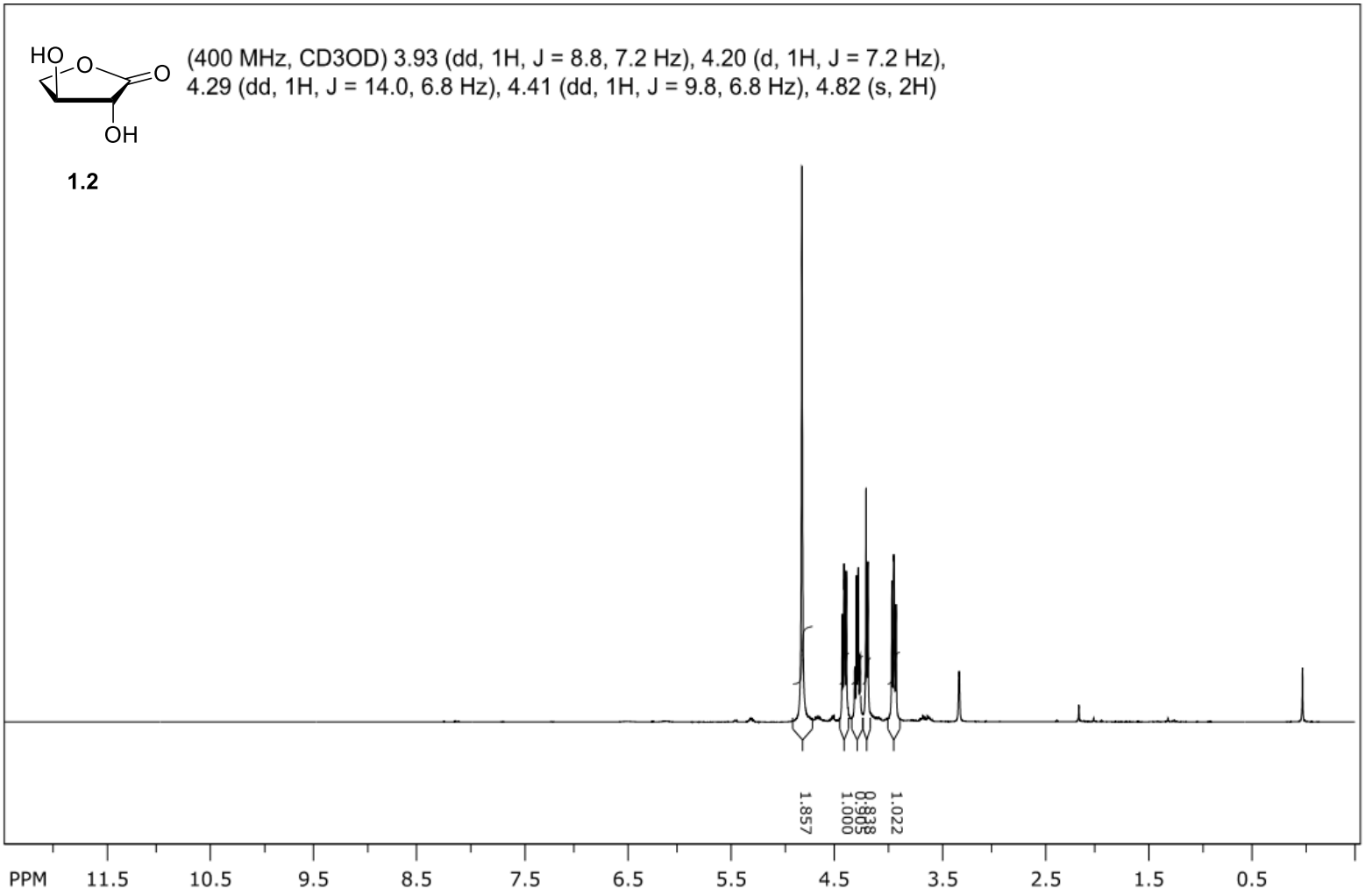
Chapter 4 described the chemical synthesis of 2′-deoxy-α-L-threofuranosyl nucleoside 3′-triphosphates (dtNTPs). The key synthetic challenge was resolved by introducing the triethylborane as the radical initiator for room temperature 2'-deoxygenation reaction of TNA nucleosides. Stepwise phosphorylation strategy developed in chapter 2 allows the synthesis of 2′-deoxy-α-L-threofuranosyl nucleoside 3′-triphosphates bearing all four genetic bases. Looking into future, we suggested these molecules can be used as the chain terminators to study the mechanism of TNA synthesis by a TNA polymerase. The information gained from this studies will allow us to engineer the polymerase variants with better efficiency and faithful genetic information delivery.

Current progress in the development of a novel class of pyrophosphate reagents in our lab has made it possible to synthesize the gram scale of natural/modified nucleoside triphosphate using a silica gel chromatography instead of a classic HPLC purification. Extensive reagent applications on the other phosphorylated molecules are under consideration. Look into future, support-on monophosphate or pyrophosphate reagents are feasible idea having the great potential to be pursed to improve the purification process of

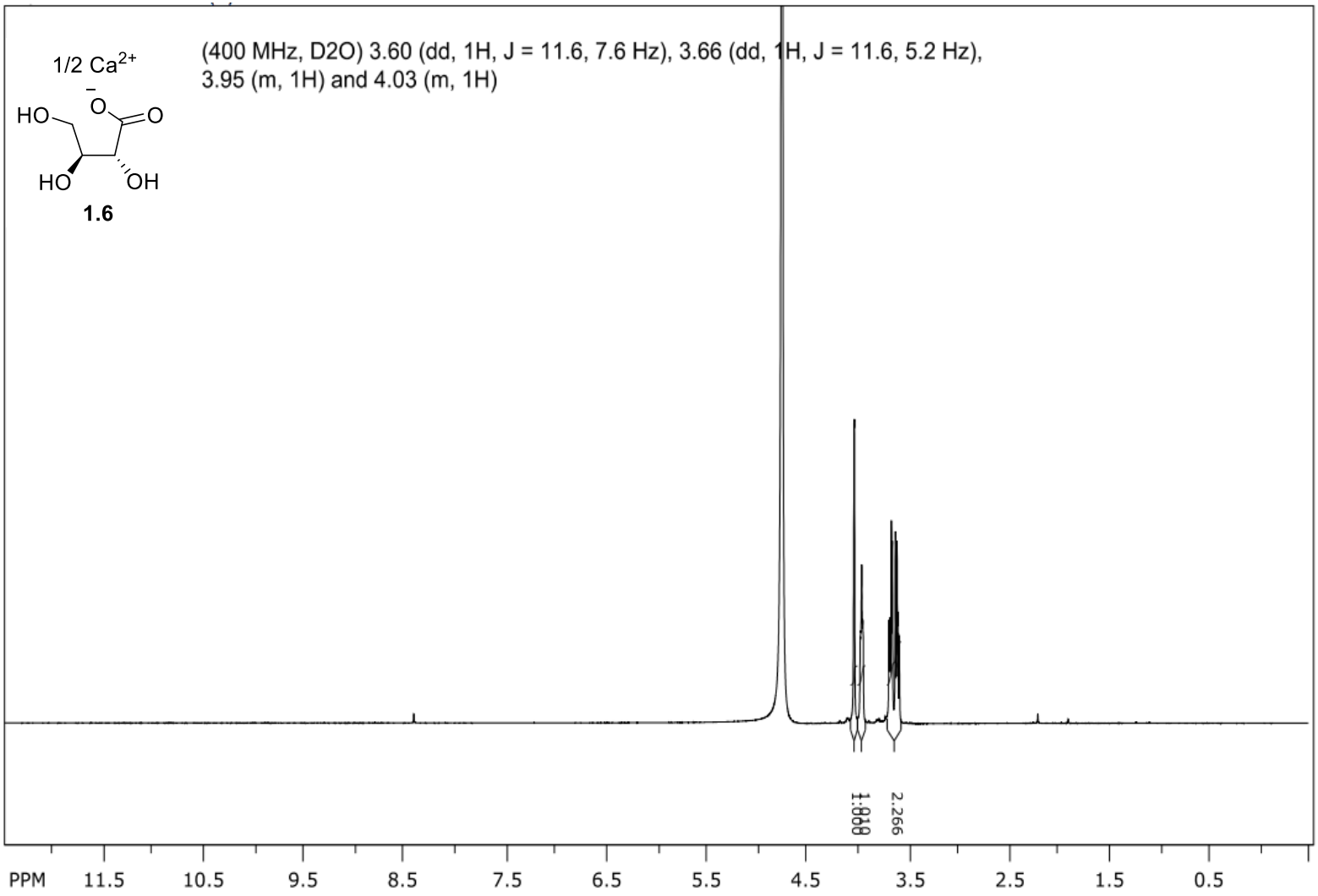
phosphorylated molecules in the synthetic chemistry field.

APPENDIX A

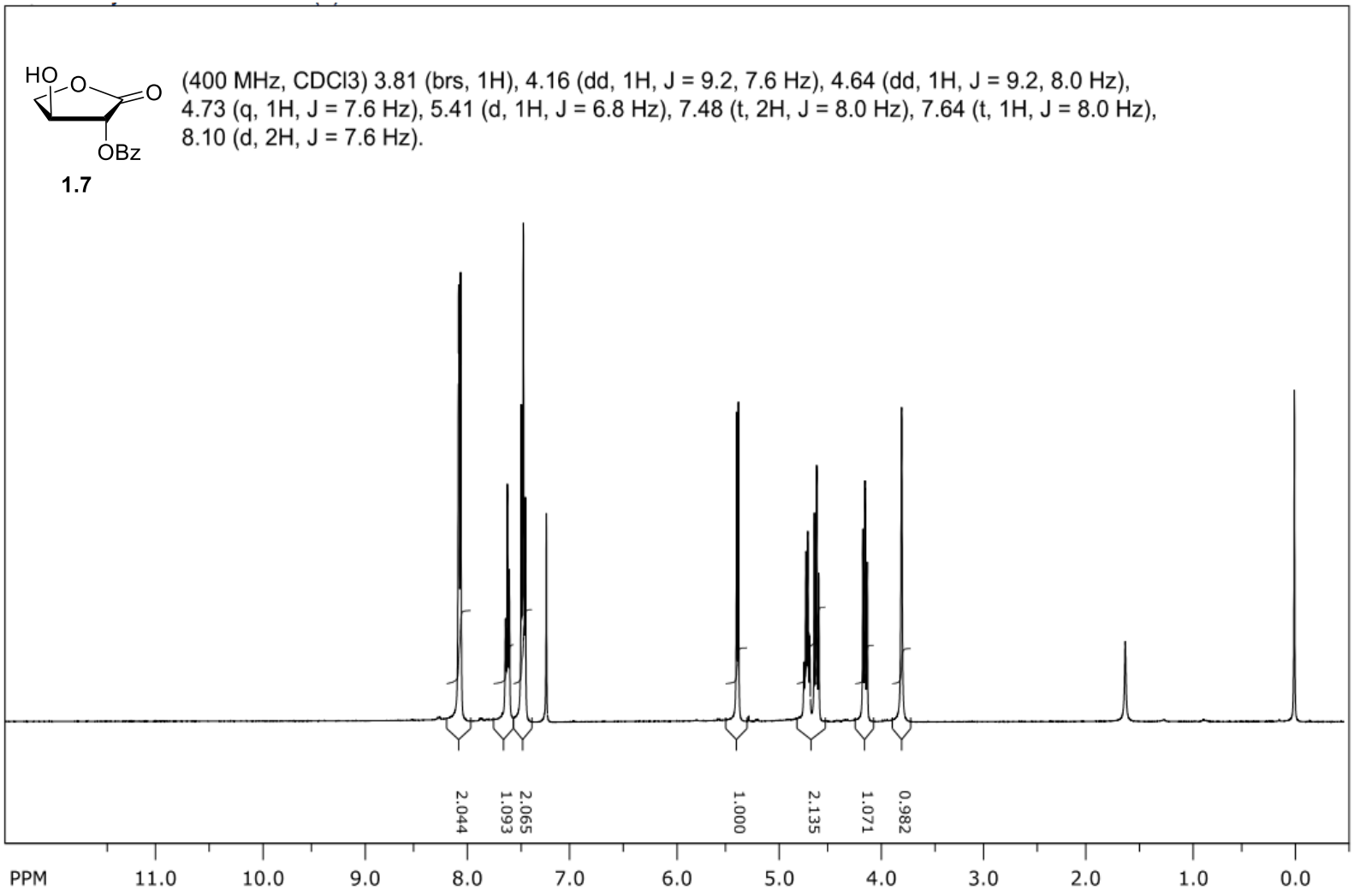
SUPPLEMENTAL NMR SPECTRA



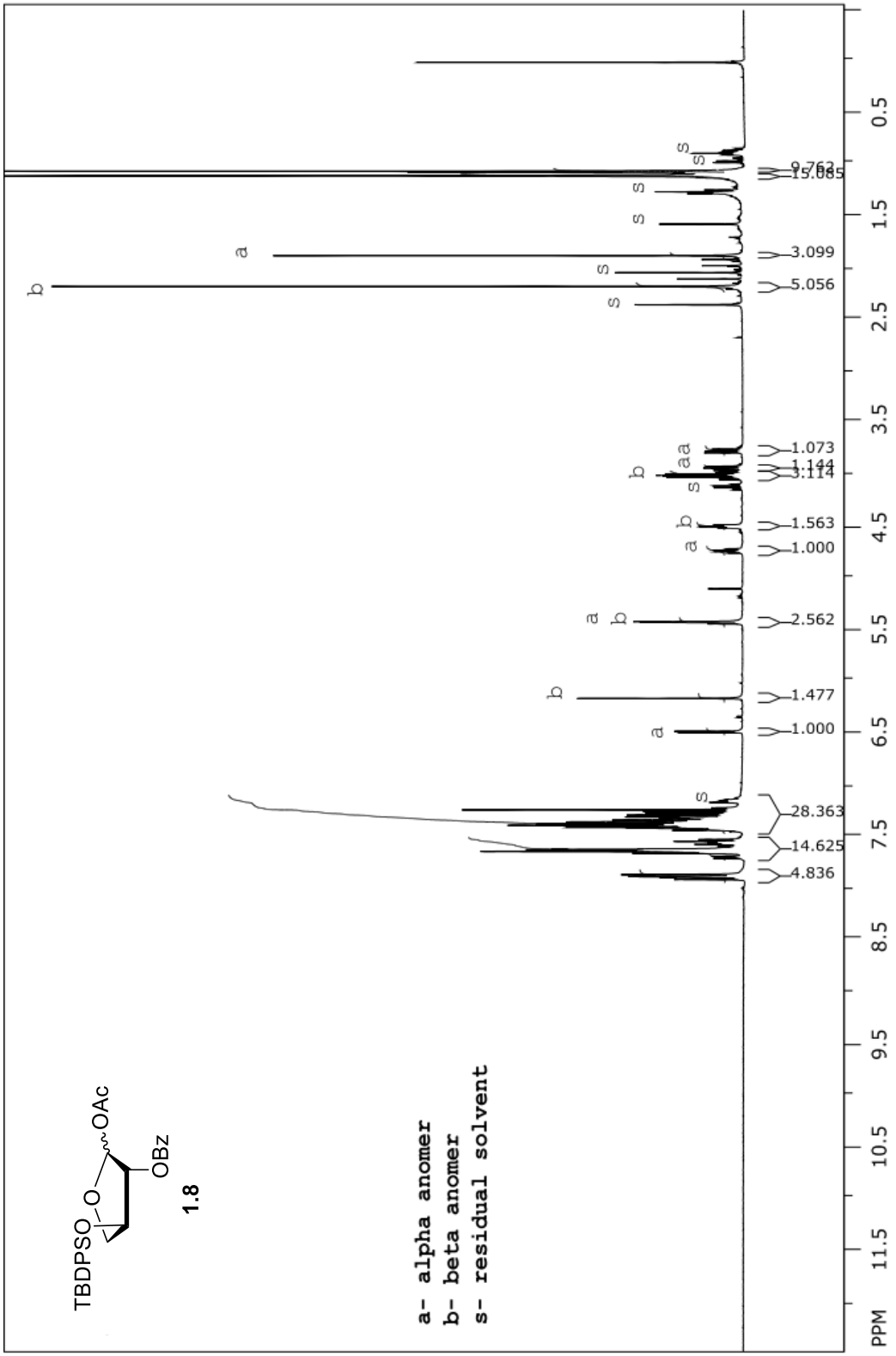
^{1}H NMR spectrum of compound **1.2** (400 MHz, $CDCl_3$)



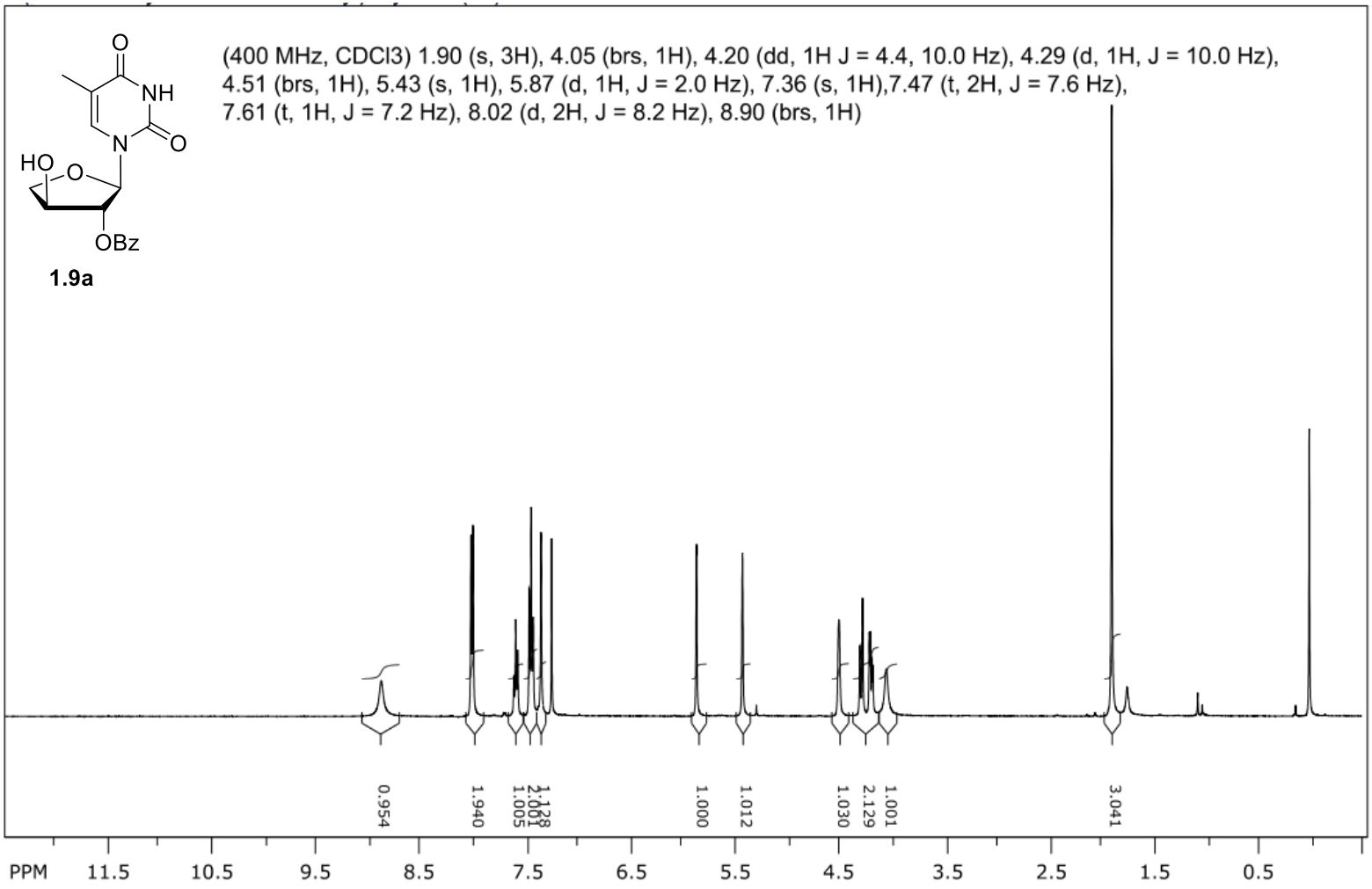
^{1}H NMR spectrum of compound **1.6** (400 MHz, D_2O)



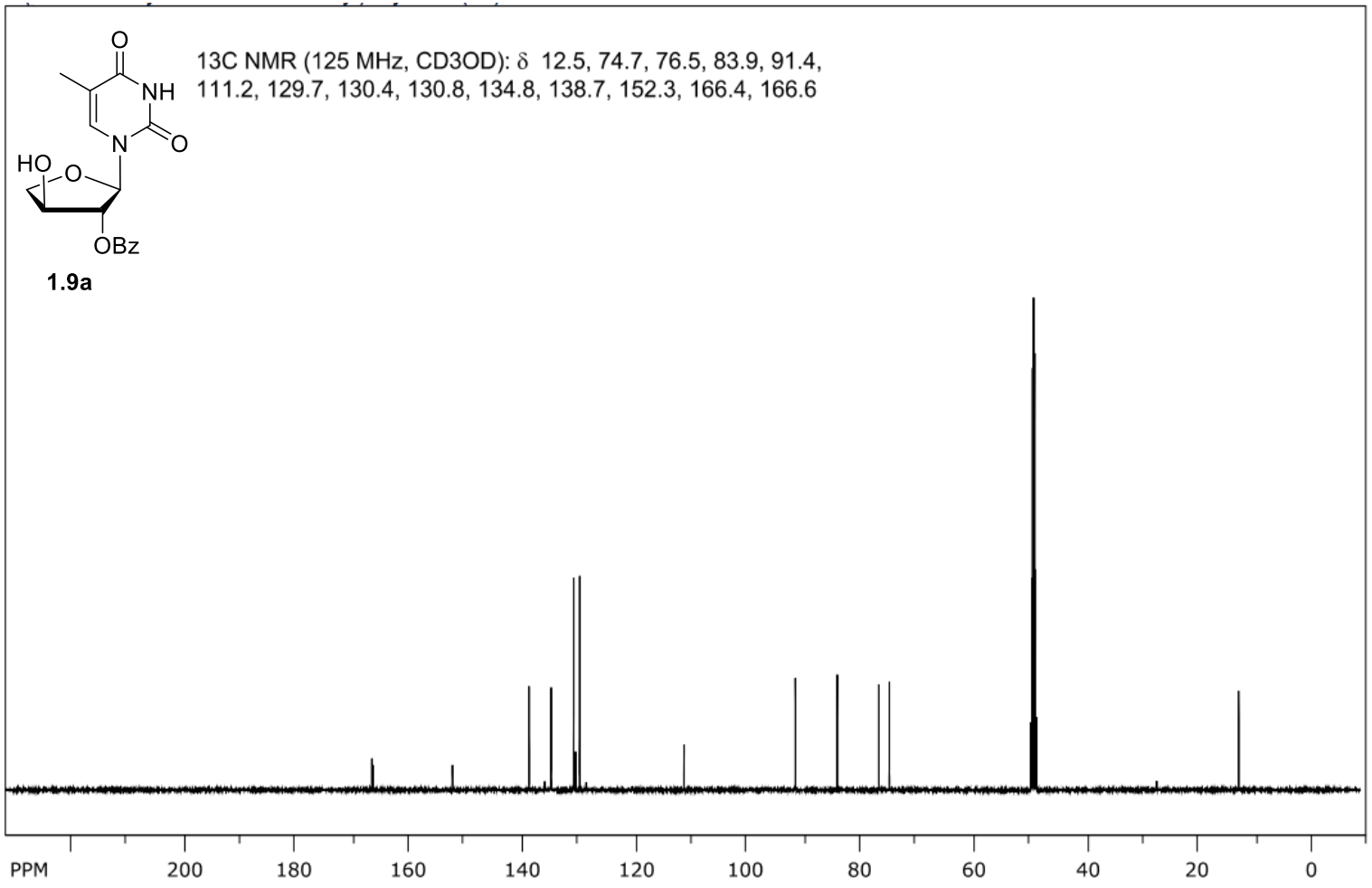
^{1}H NMR spectrum of compound **1.7** (400 MHz, $CDCl_3$)



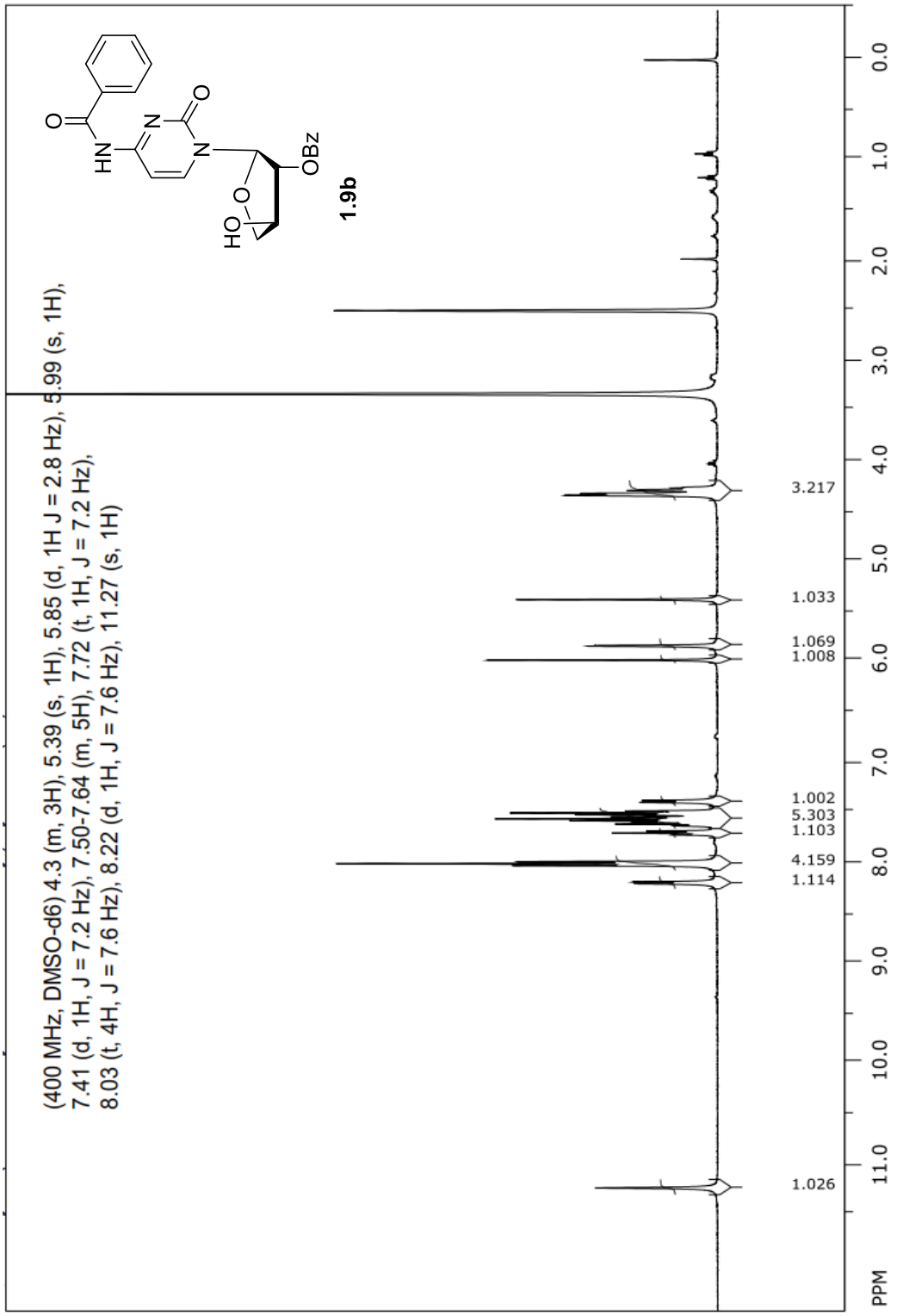
^{1}H NMR spectrum of compound **1.8** (400 MHz, $CDCl_3$)



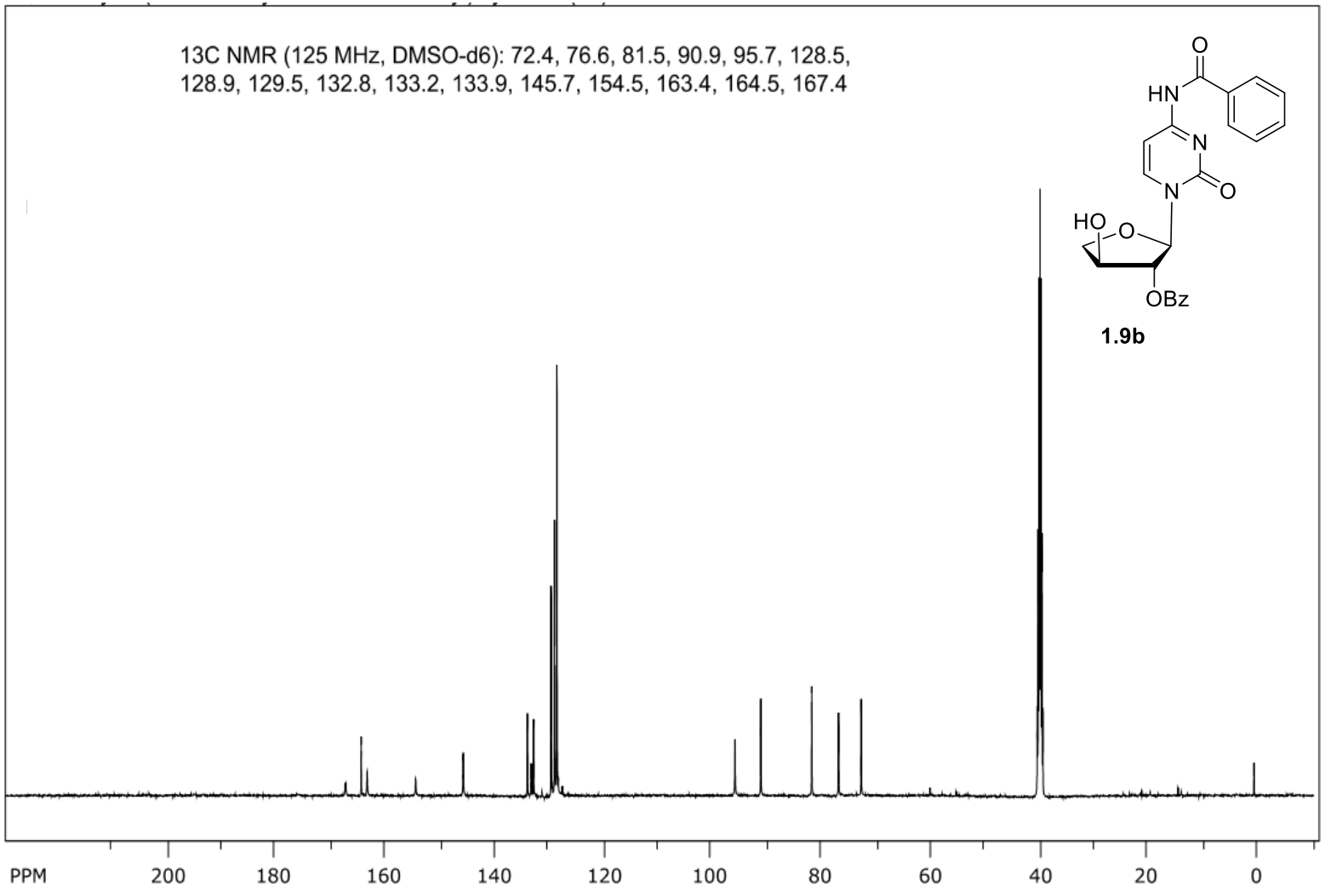
^{1}H NMR spectrum of compound **1.9a** (400 MHz, $CDCl_3$)



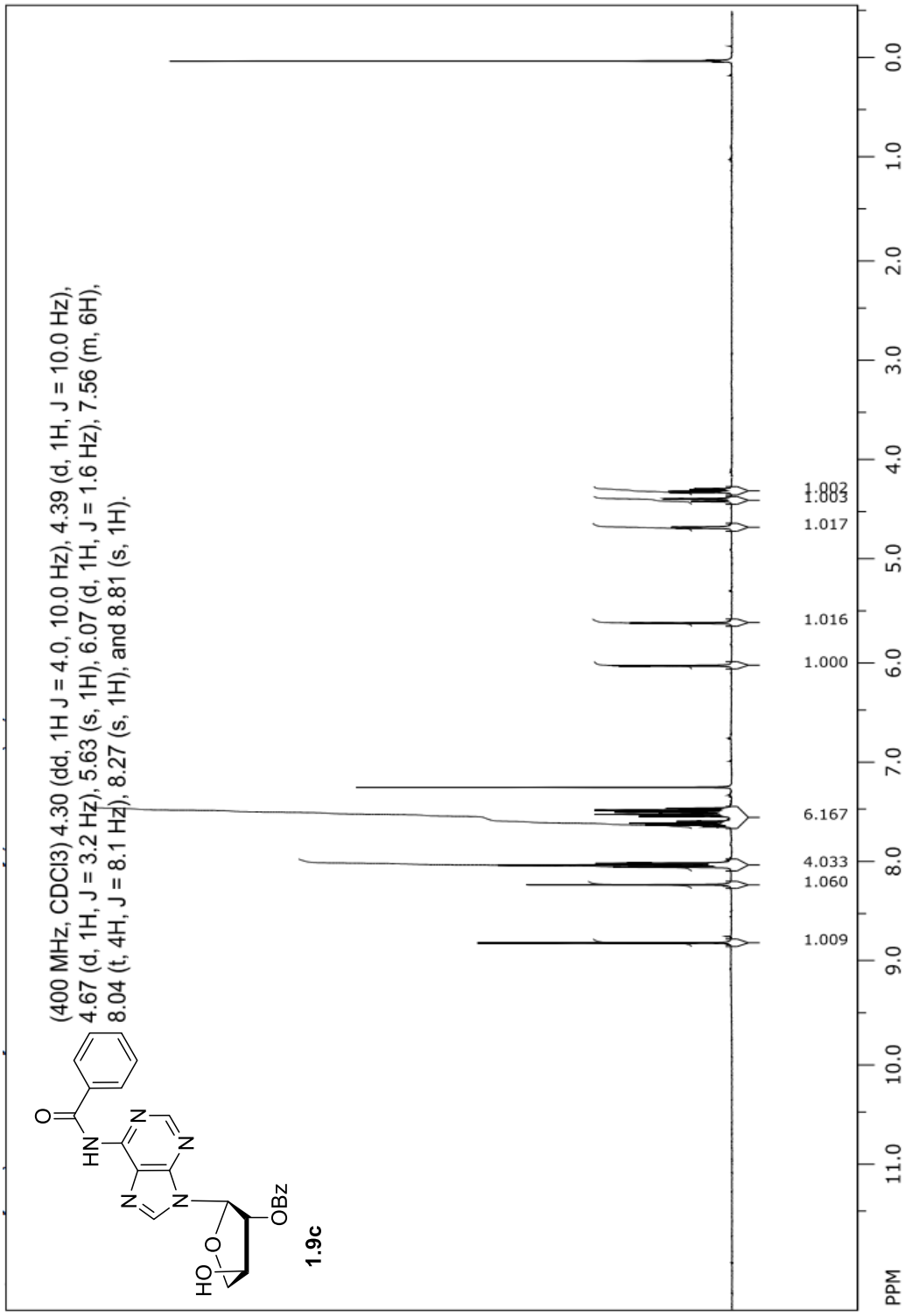
^{13}C NMR spectrum of compound **1.9a** (125 MHz, $CDCl_3$)



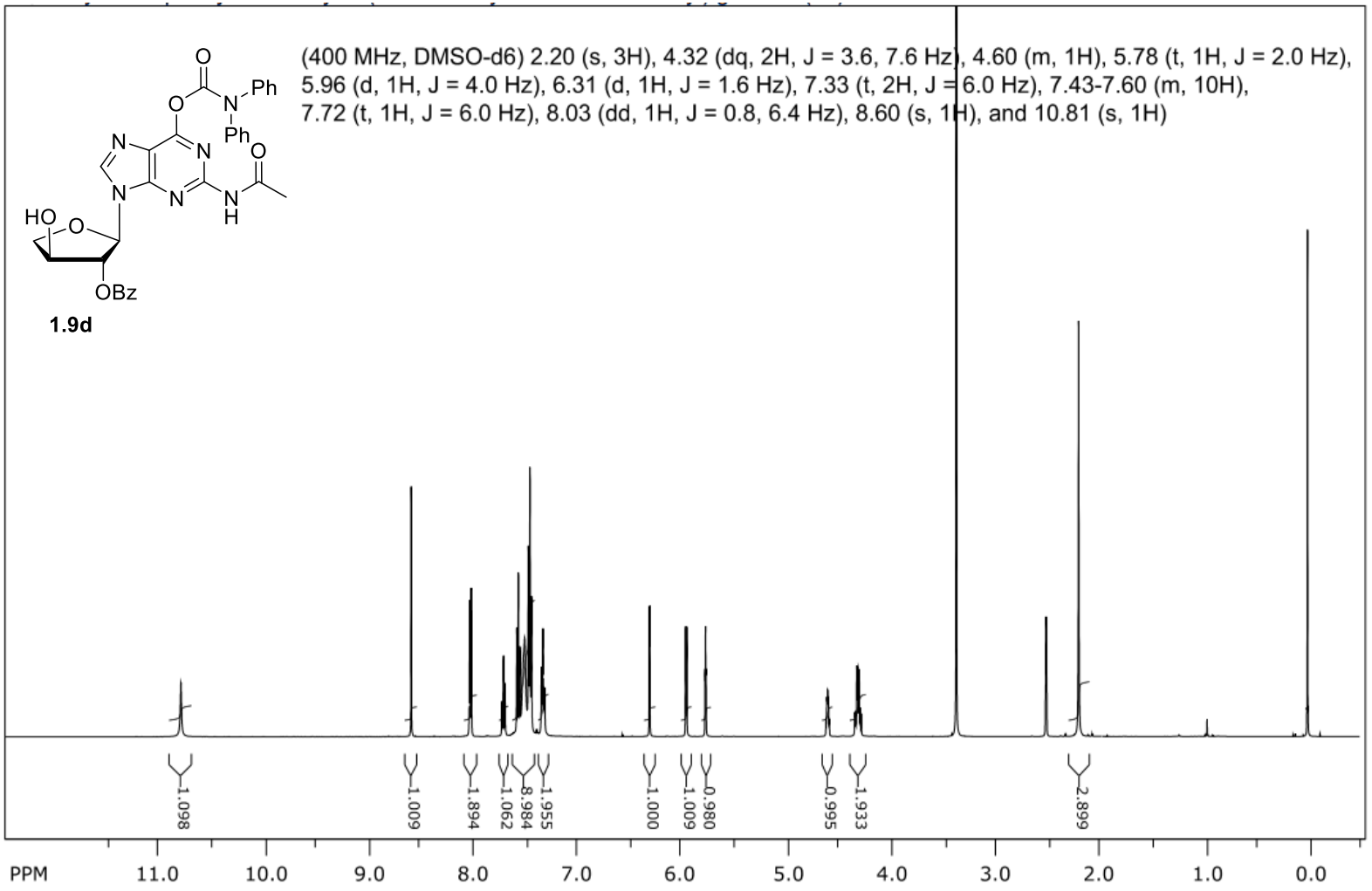
^{1}H NMR spectrum of compound **1.9b** (400 MHz, DMSO-d_6)



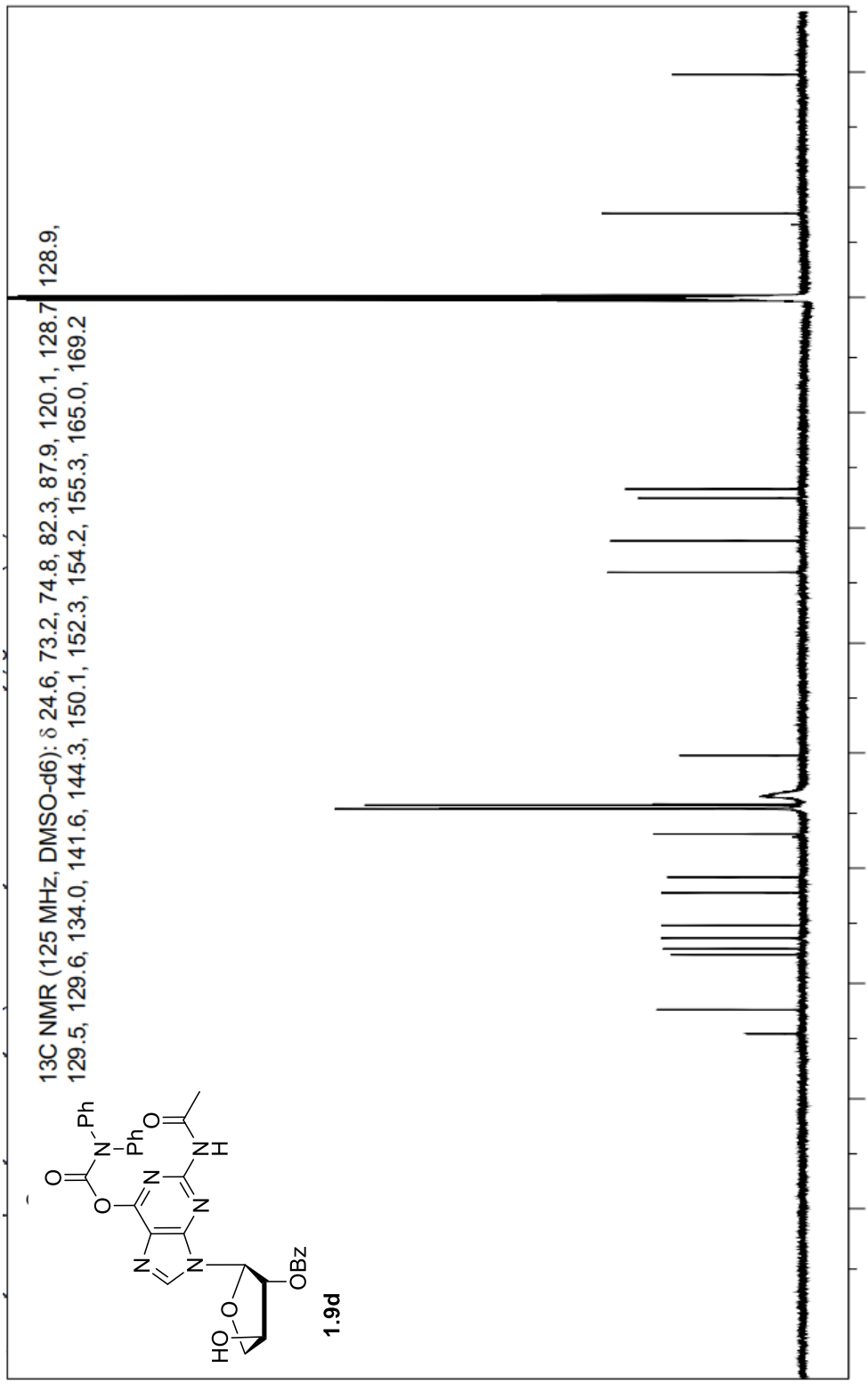
^{13}C NMR spectrum of compound **1.9b** (125 MHz, DMSO-d_6)



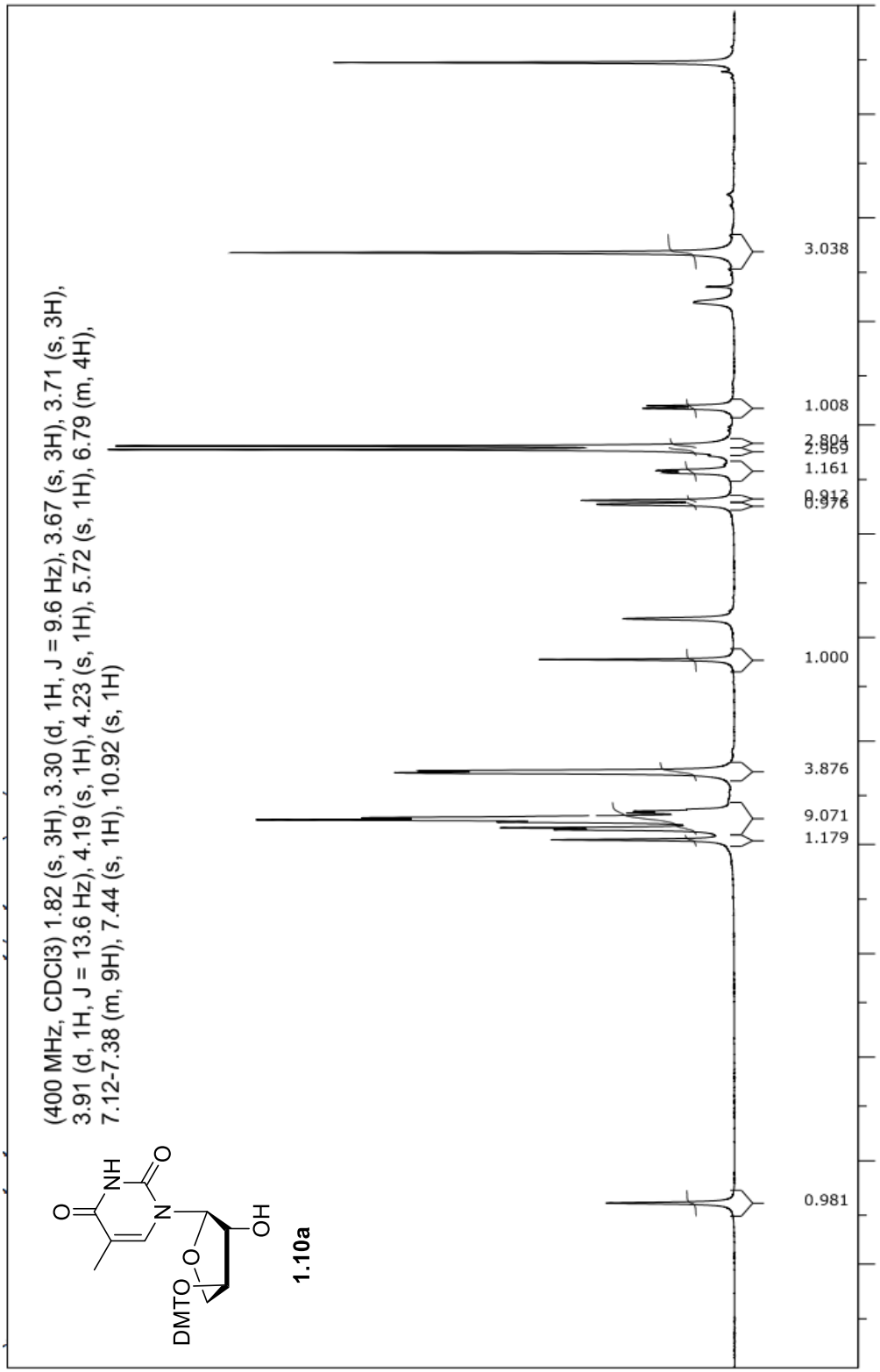
^{1}H NMR spectrum of compound **1.9c** (400 MHz, DMSO-d_6)



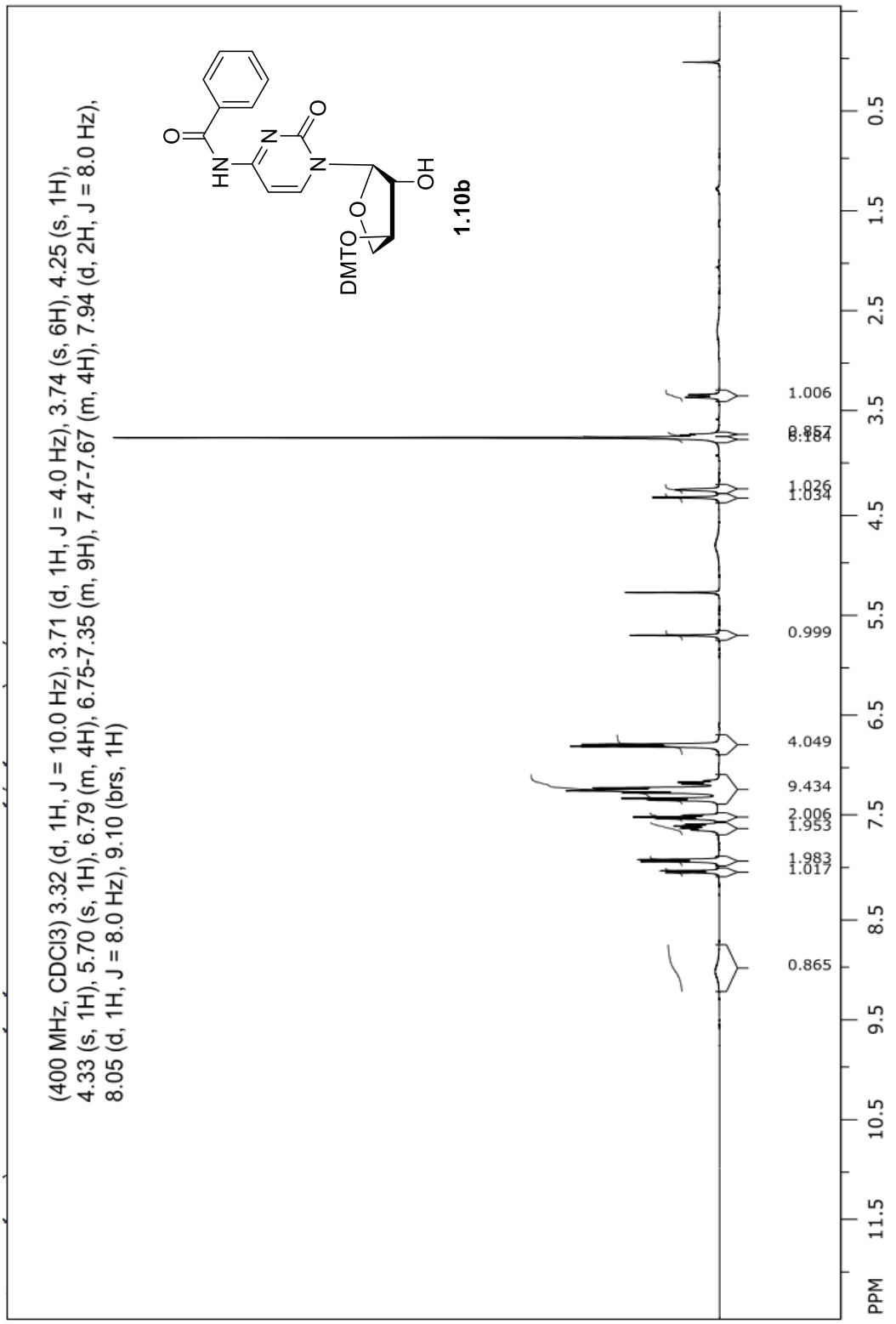
^{1}H NMR spectrum of compound **1.9d** (400 MHz, DMSO-d_6)



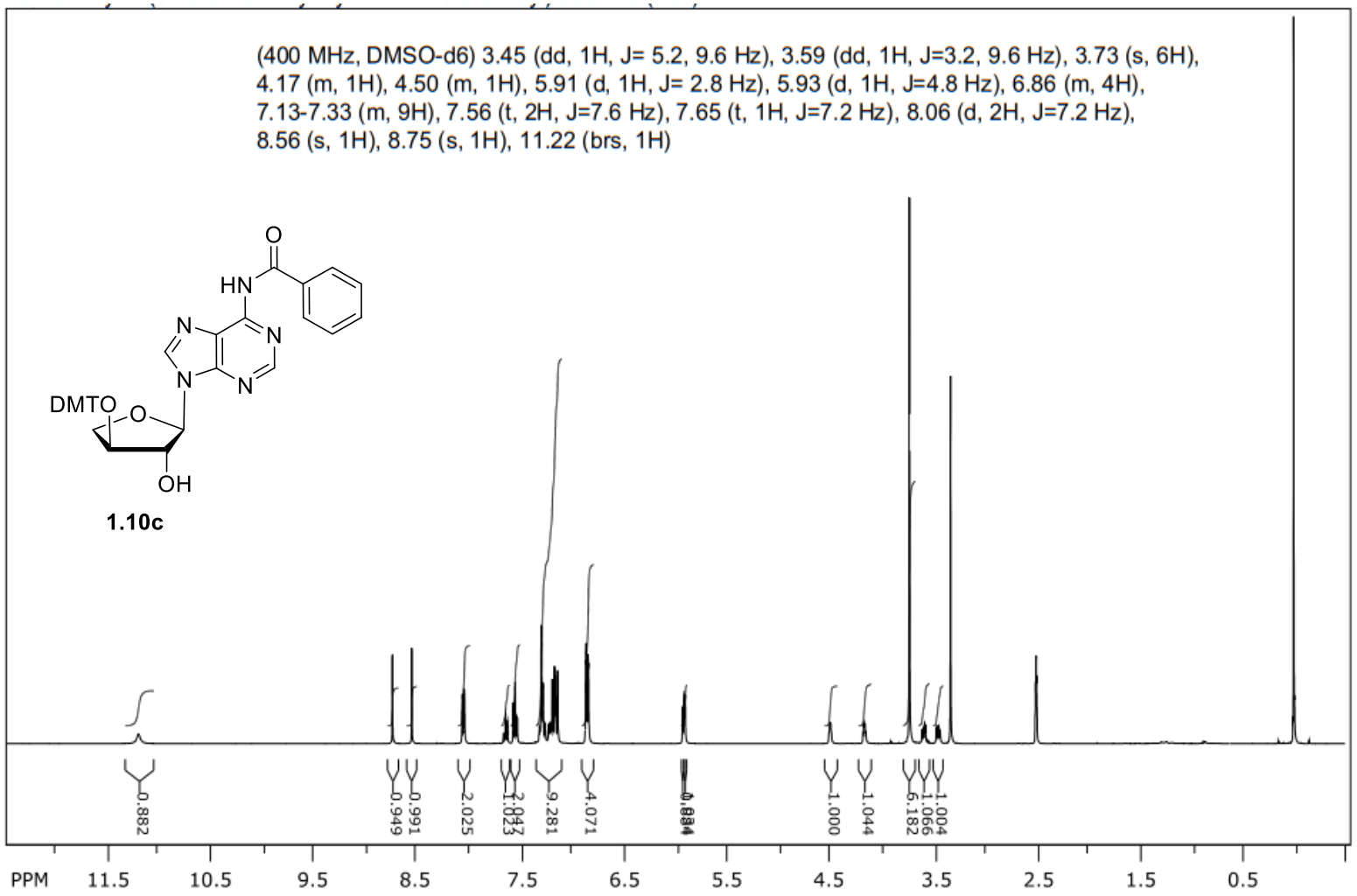
^{13}C NMR spectrum of compound **1.9d** (125 MHz, DMSO-d_6)



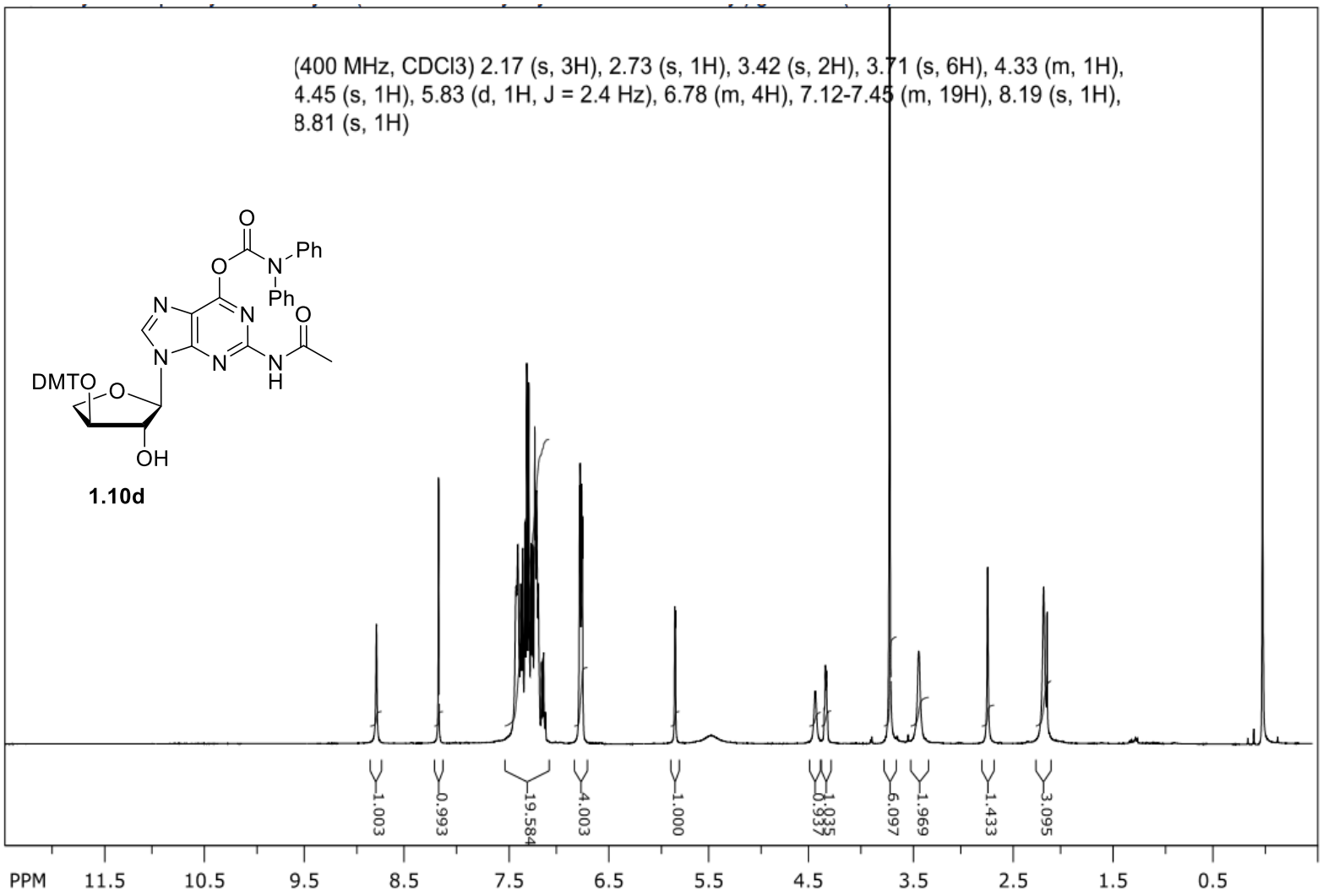
^{1}H NMR spectrum of compound **1.10a** (400 MHz, $CDCl_3$)



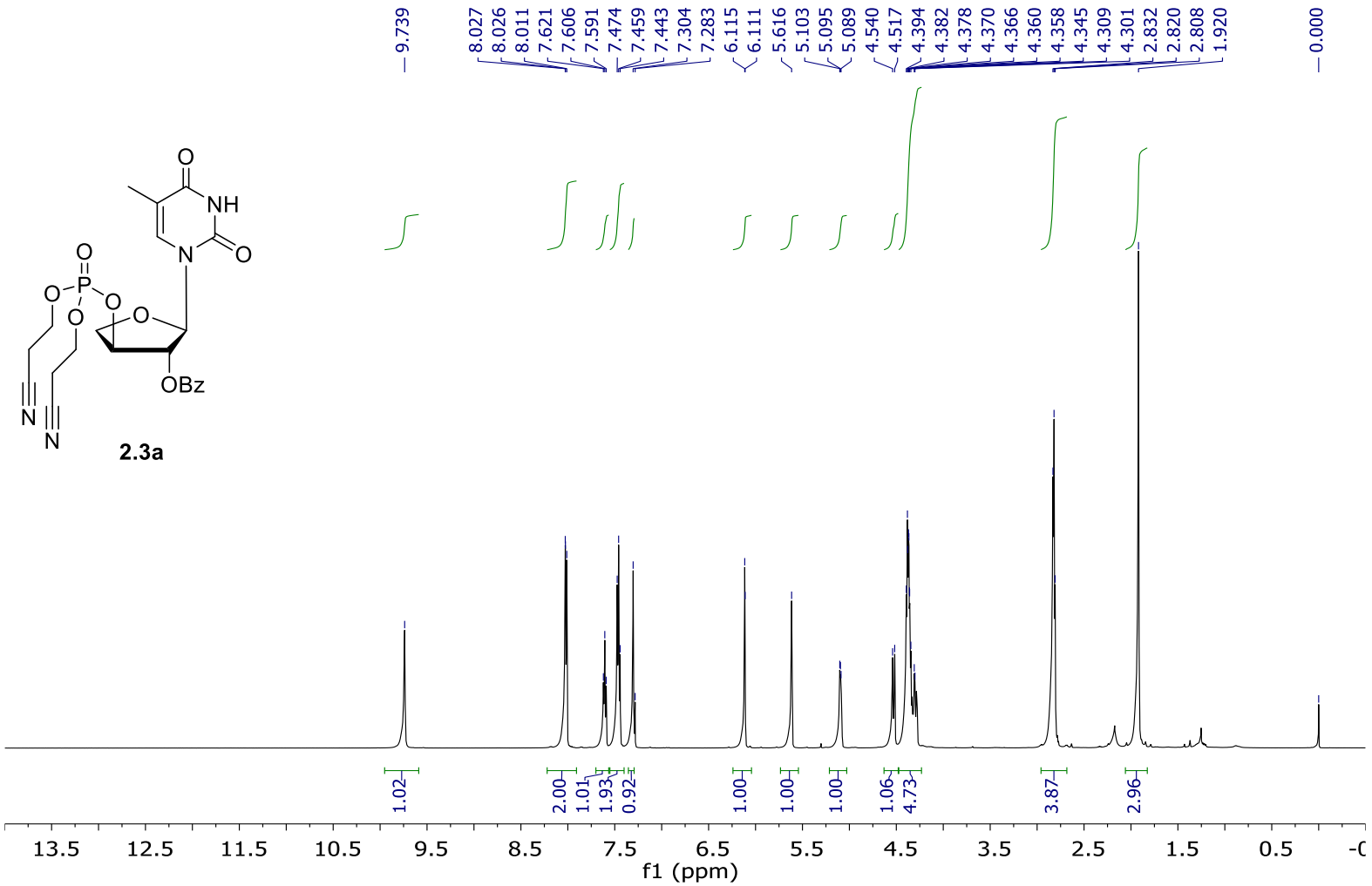
^{1}H NMR spectrum of compound **1.10b** (400 MHz, $CDCl_3$)



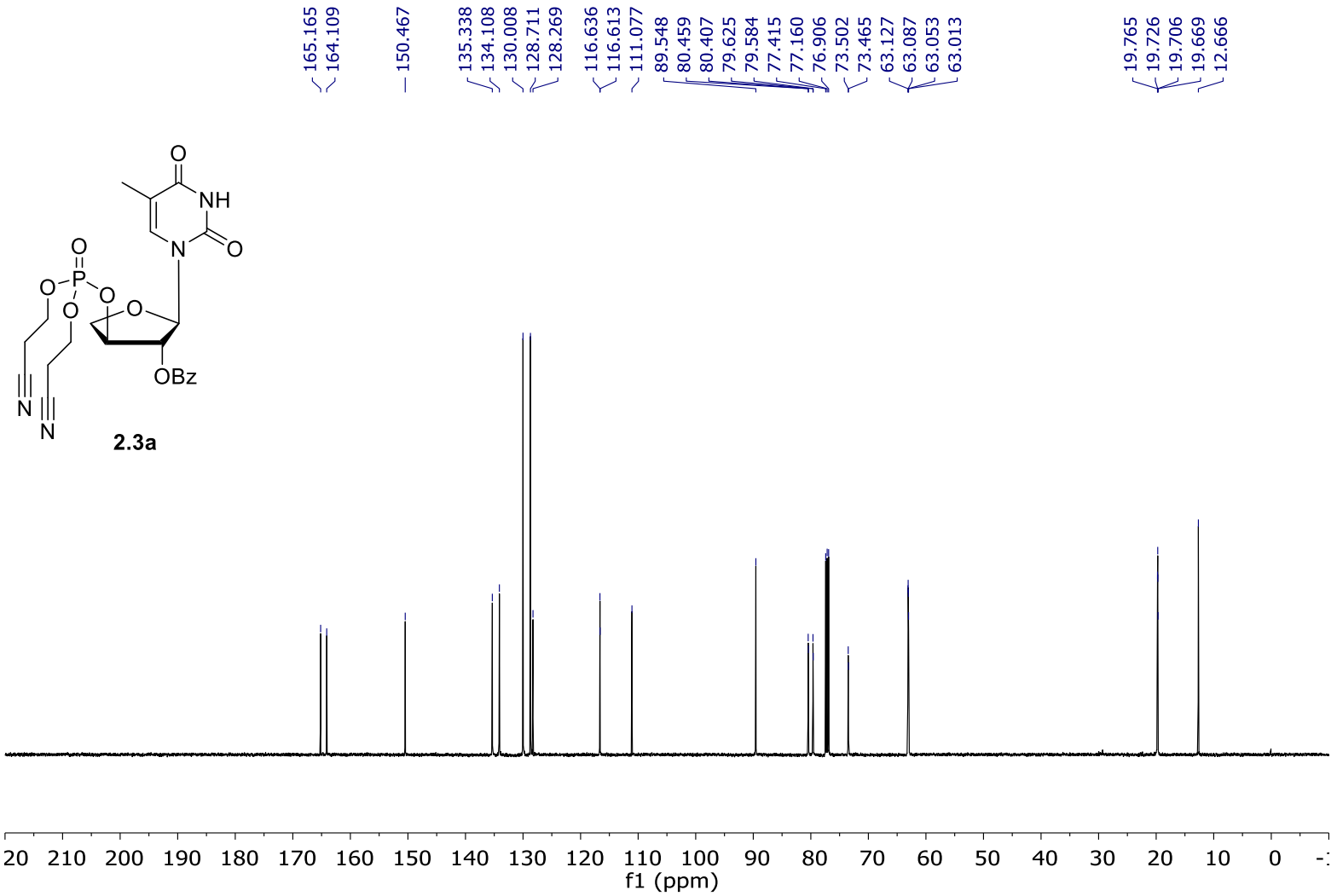
^{1}H NMR spectrum of compound **1.10c** (400 MHz, DMSO-d_6)



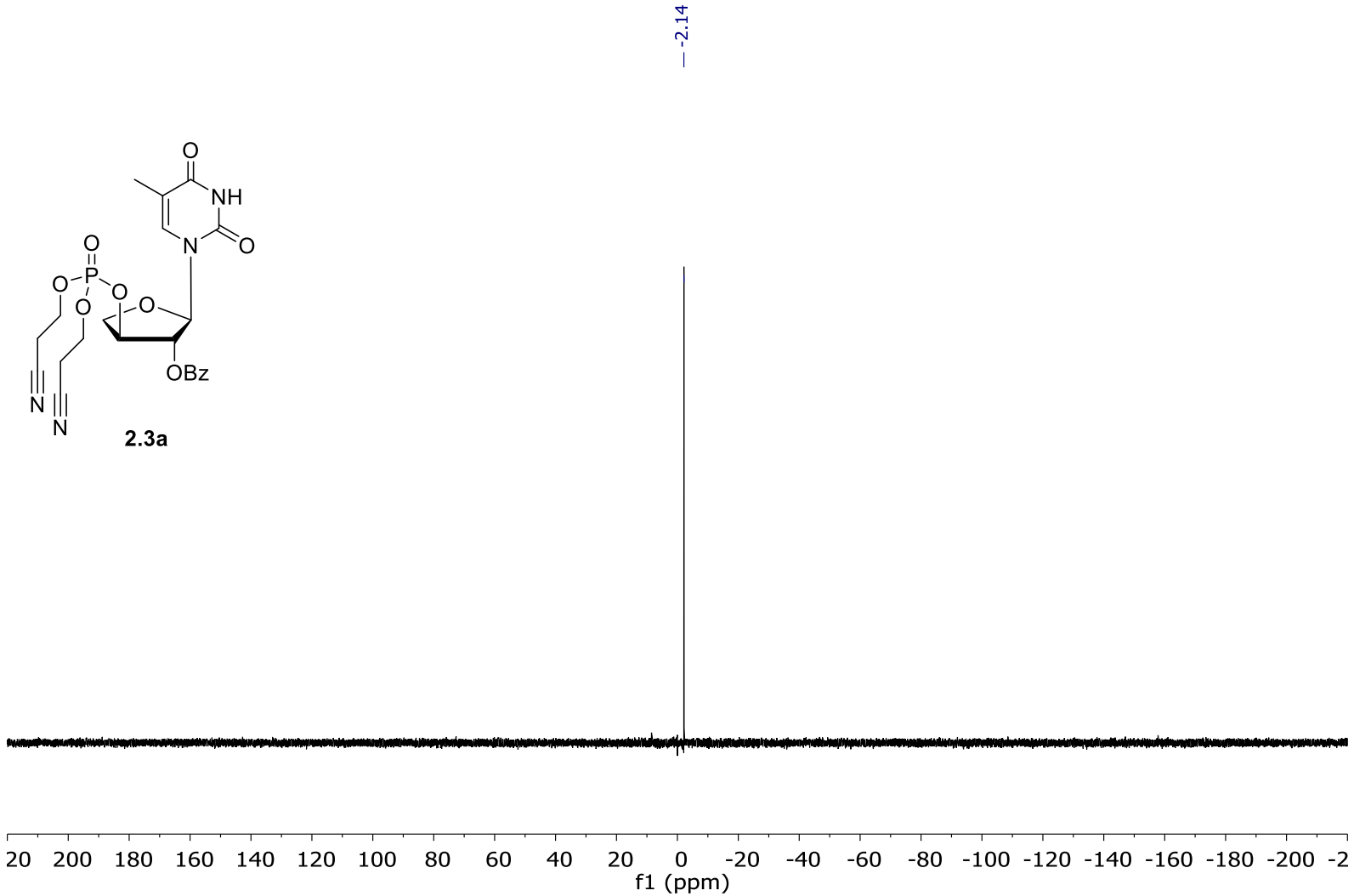
^{1}H NMR spectrum of compound **1.10d** (400 MHz, $CDCl_3$)



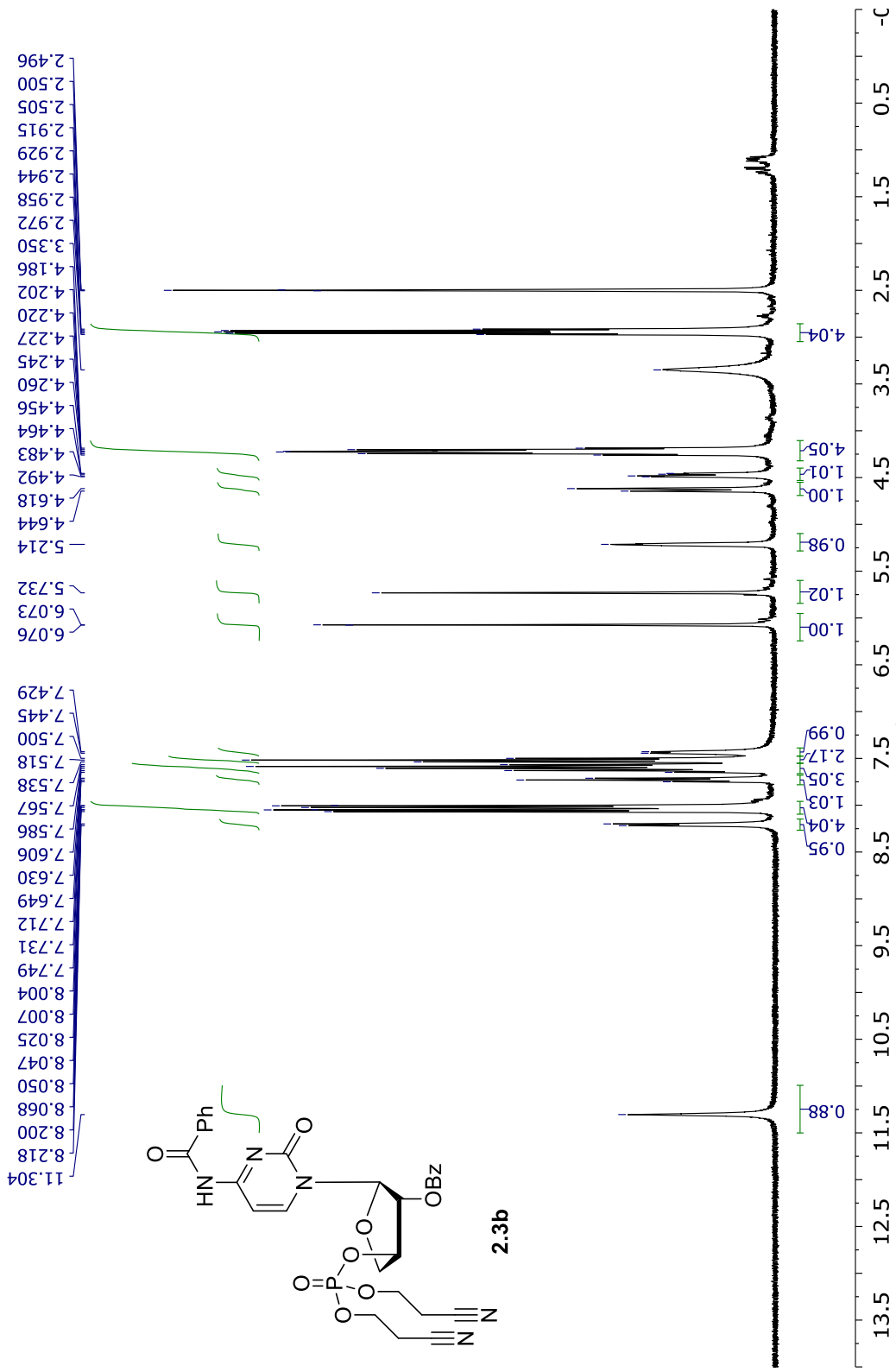
^{1}H NMR spectrum of compound **2.3a** (500 MHz, $CDCl_3$)



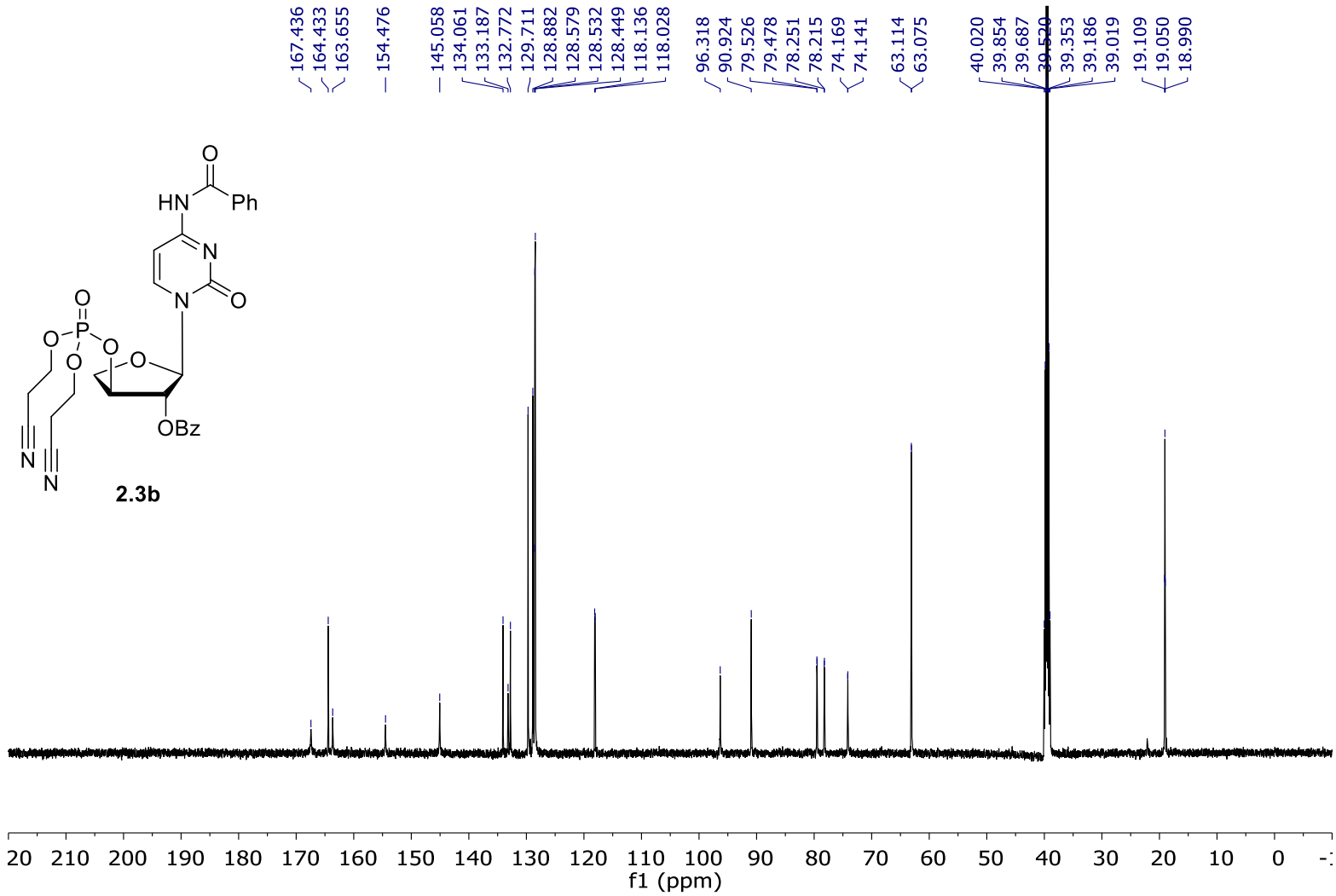
^{13}C NMR spectrum of compound **2.3a** (125.8 MHz, $CDCl_3$)



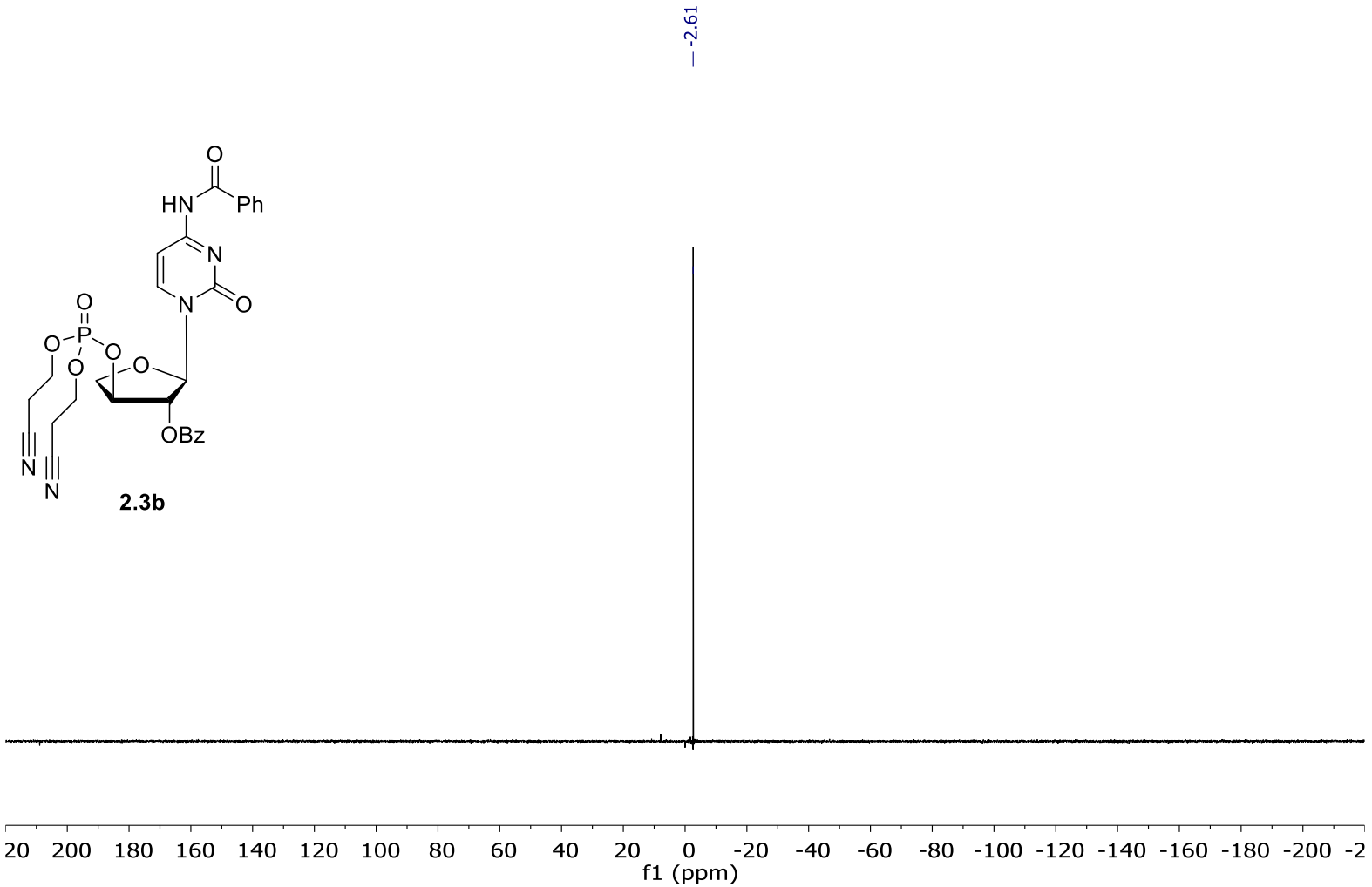
^{31}P NMR spectrum of compound **2.3a** (162 MHz, $CDCl_3$)



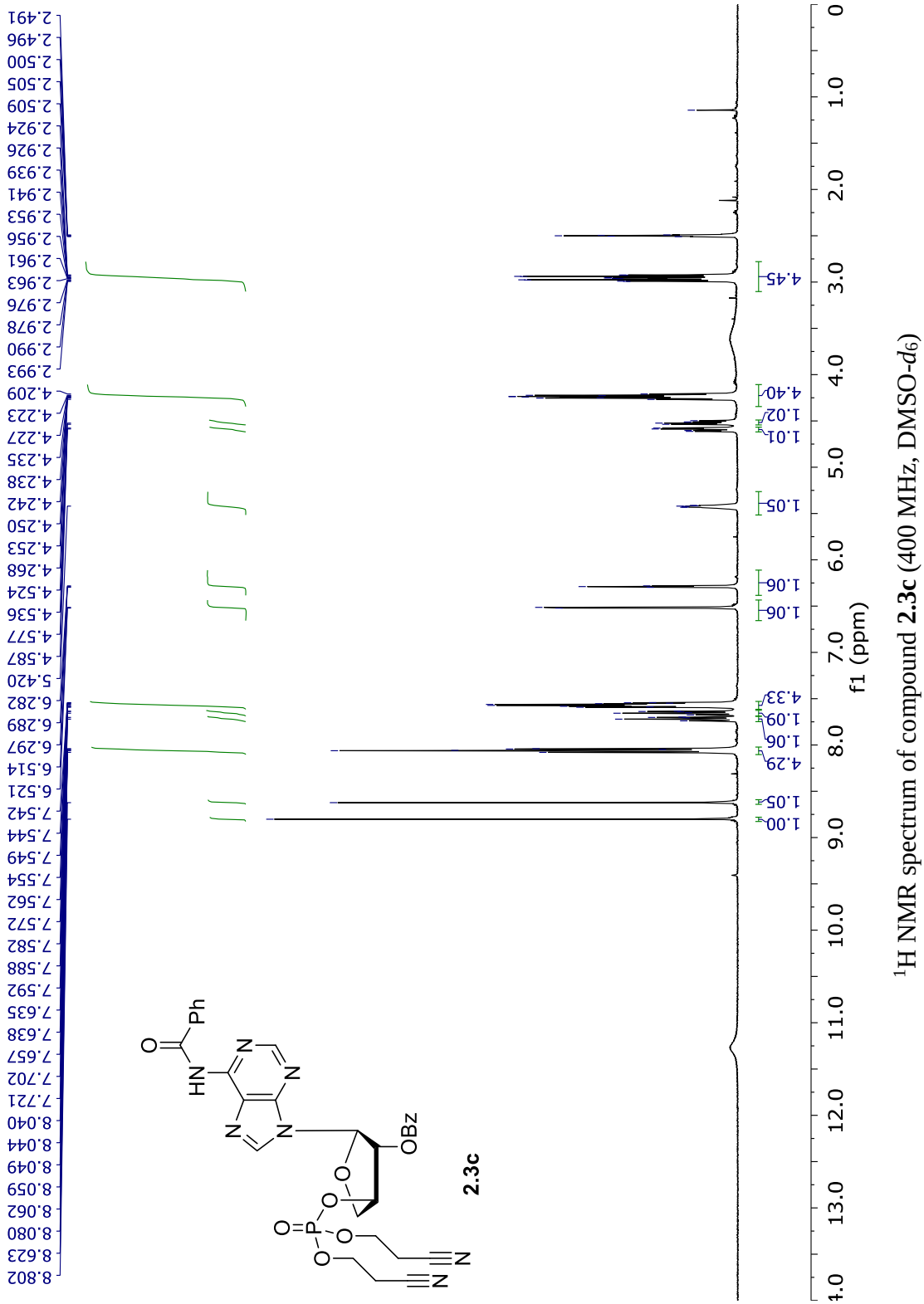
^{1}H NMR spectrum of compound **2.3b** (400 MHz, DMSO-d_6)



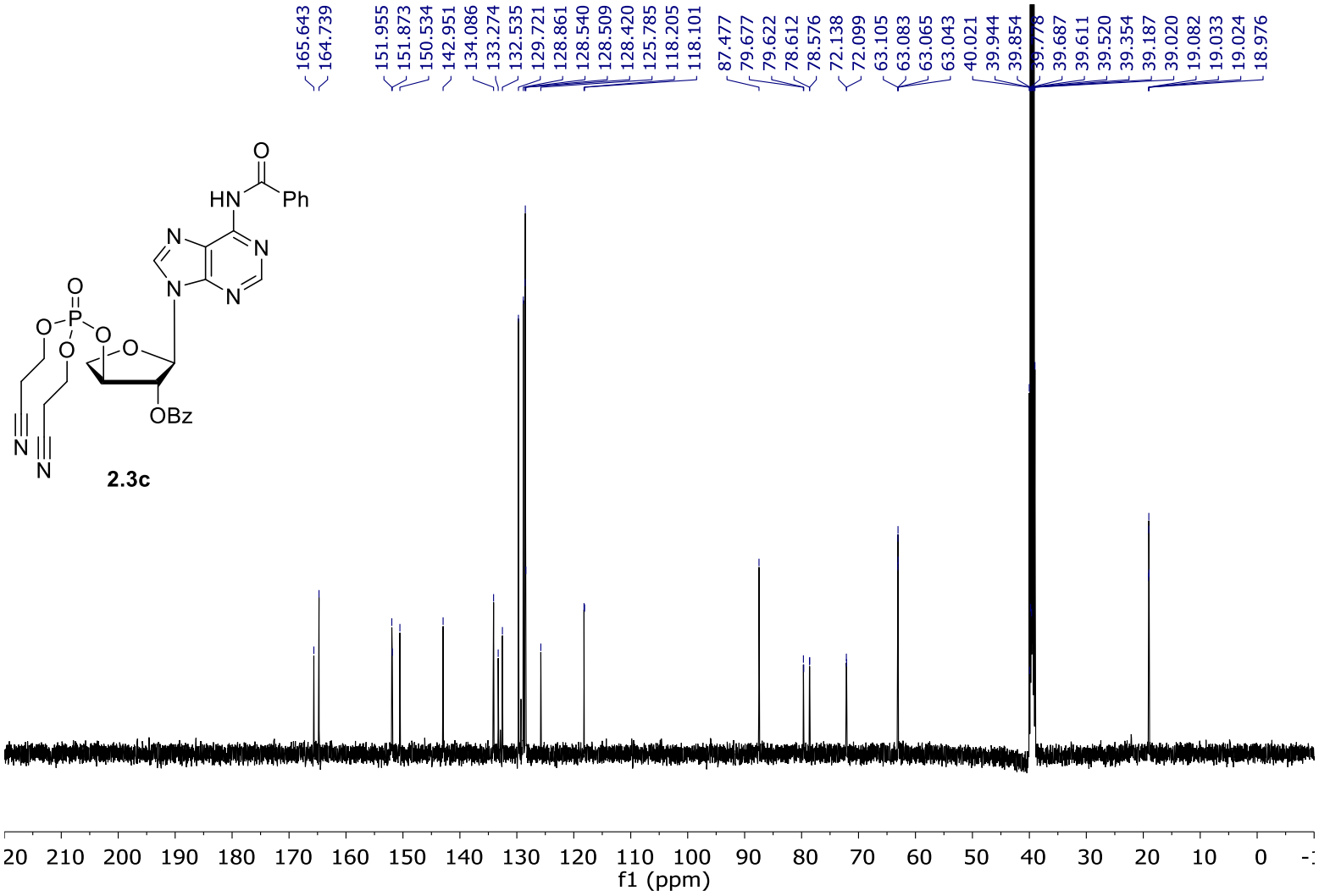
^{13}C NMR spectrum of compound **2.3b** (125.8 MHz, DMSO-d_6)



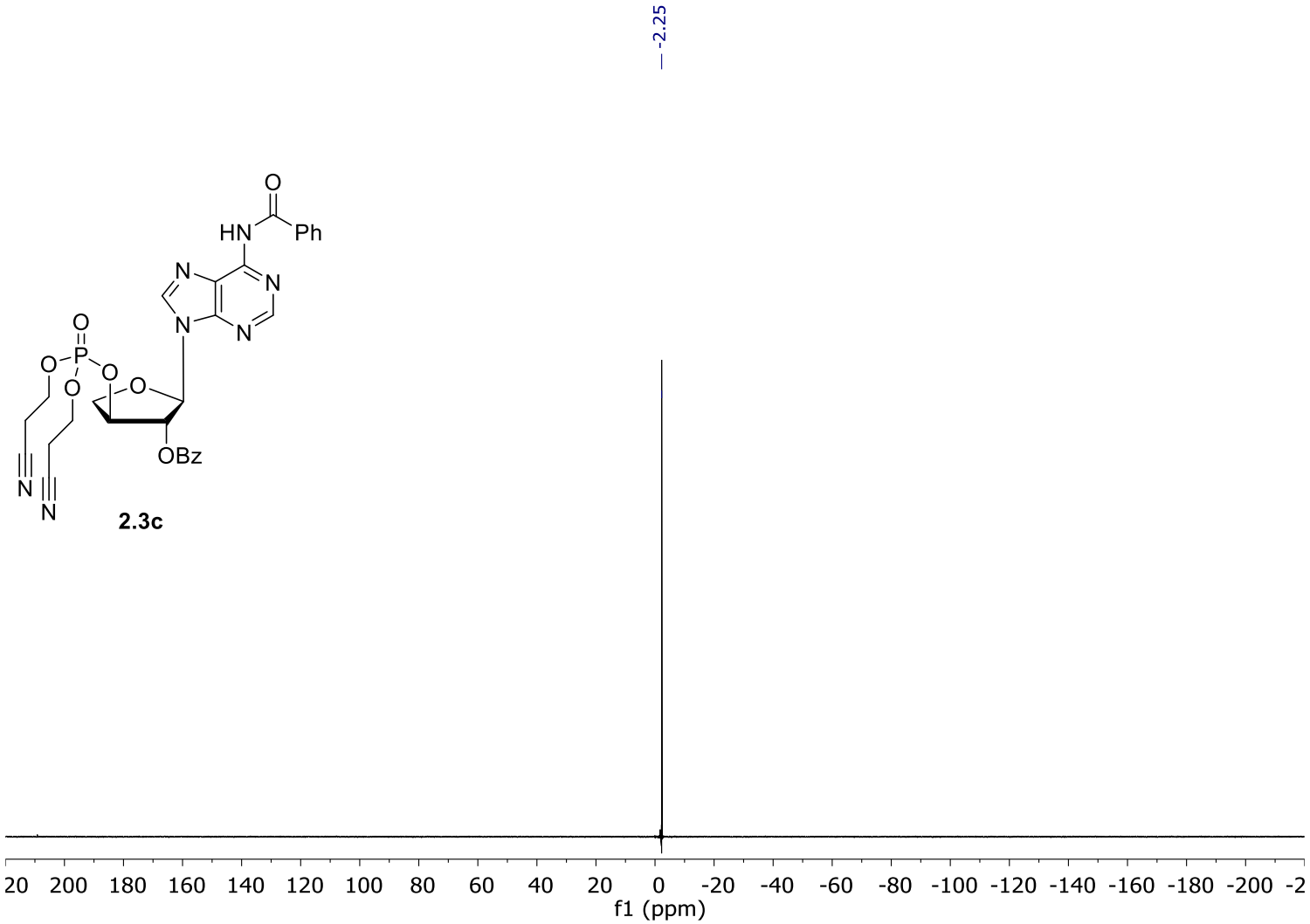
^{31}P NMR spectrum of compound **2.3b** (162 MHz, DMSO-d_6)



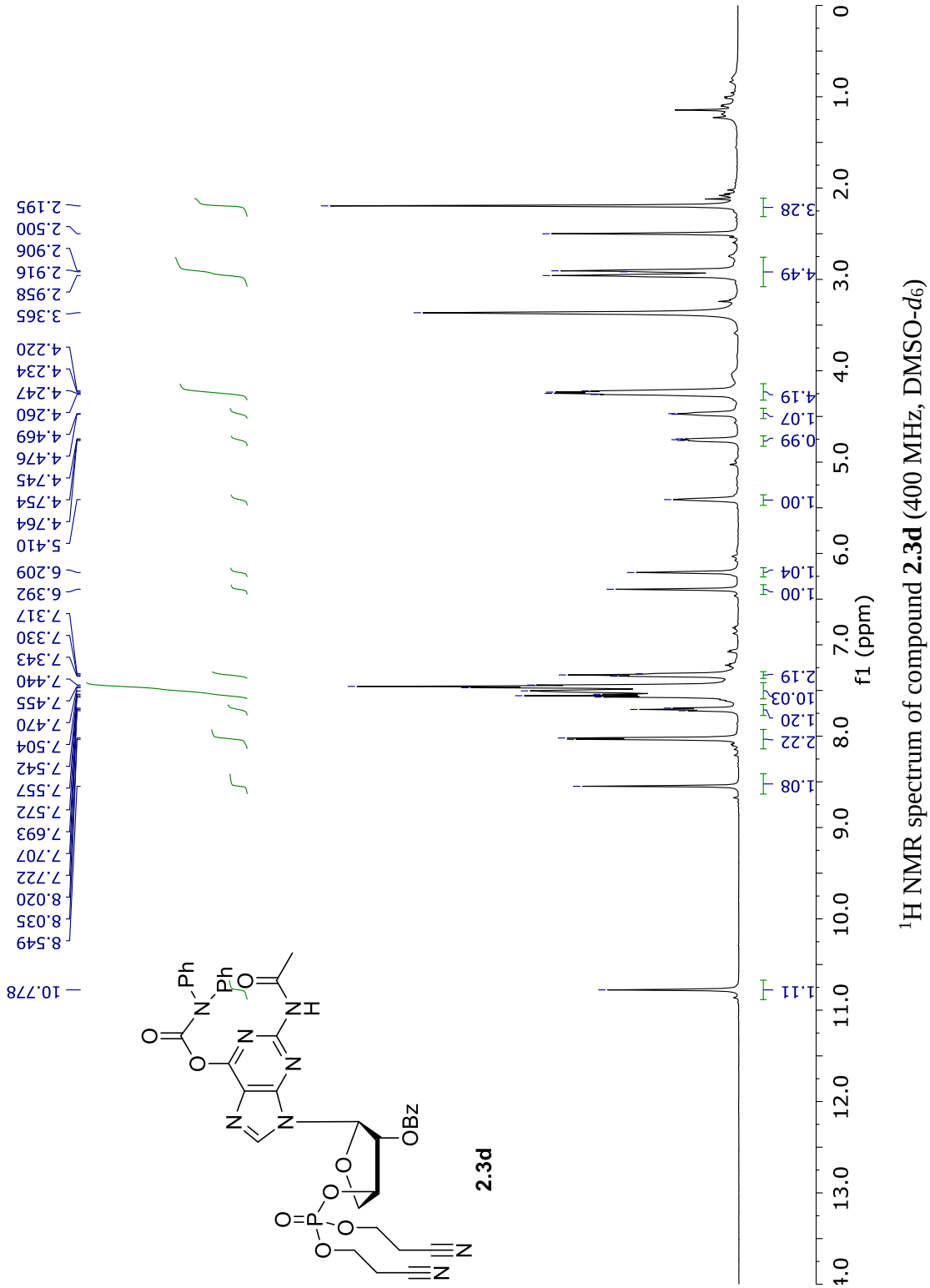
^{1}H NMR spectrum of compound **2.3c** (400 MHz, DMSO-d_6)



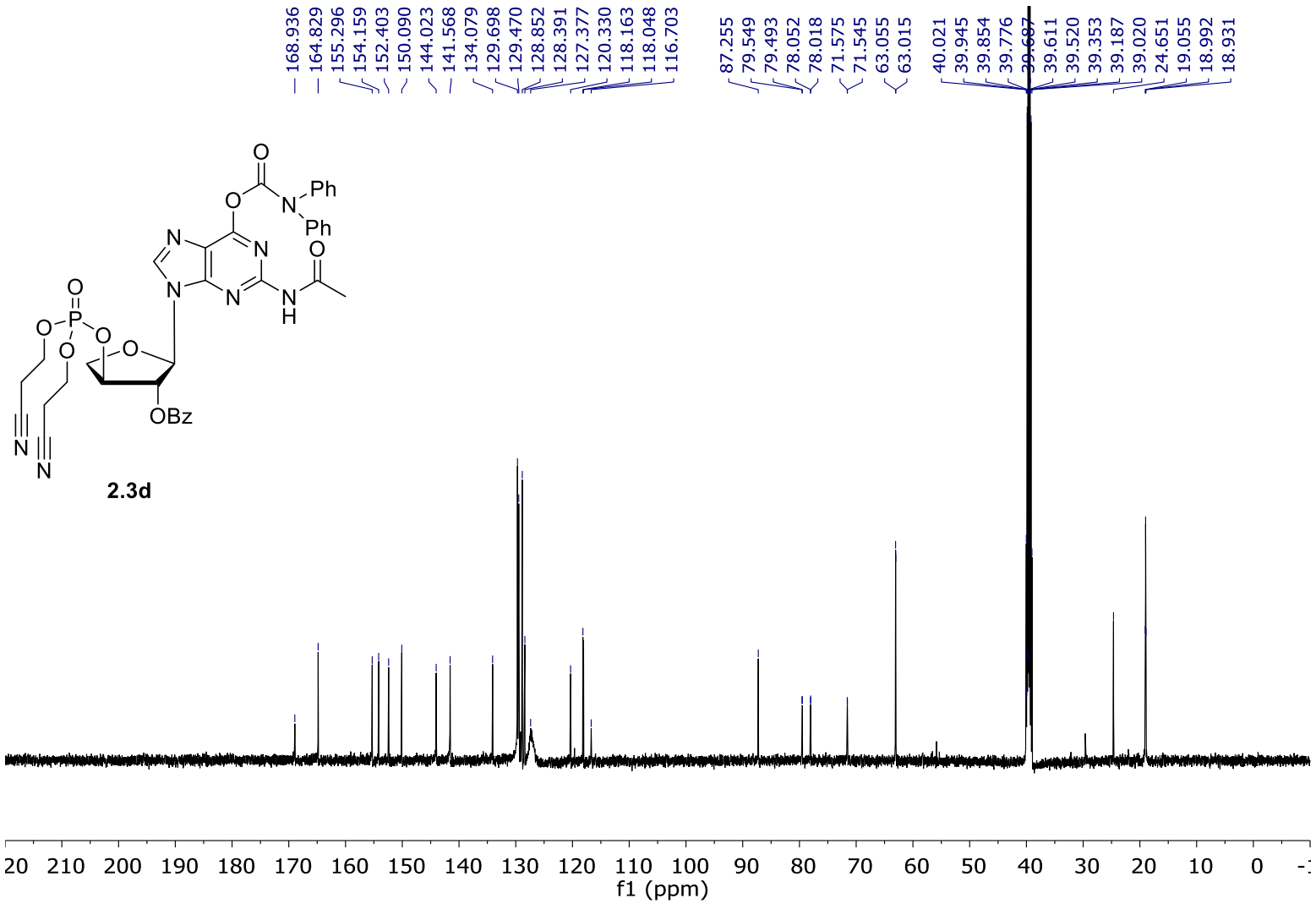
^{13}C NMR spectrum of compound **2.3c** (125.8 MHz, DMSO-d_6)



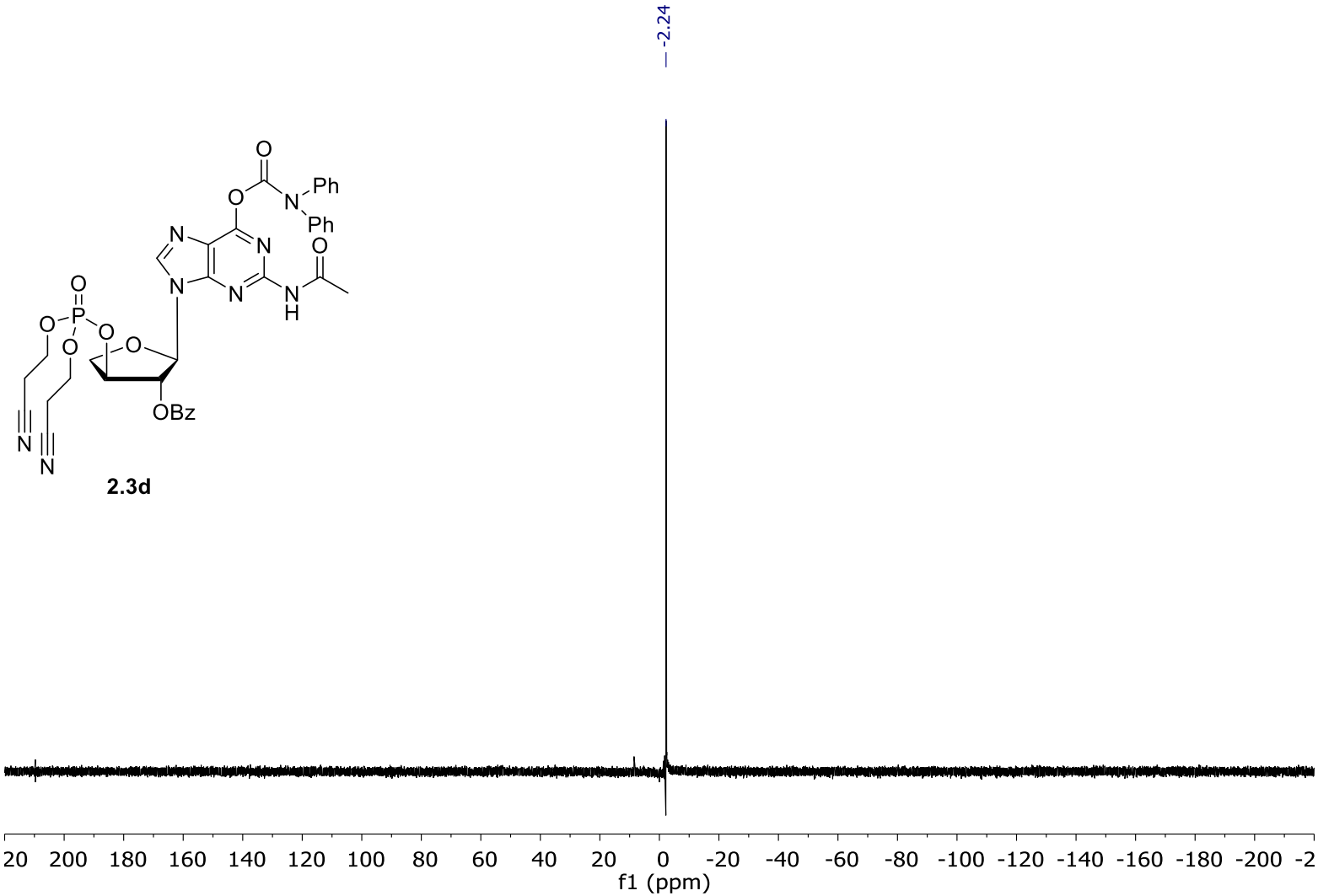
^{31}P NMR spectrum of compound **2.3c** (162 MHz, DMSO-d_6)



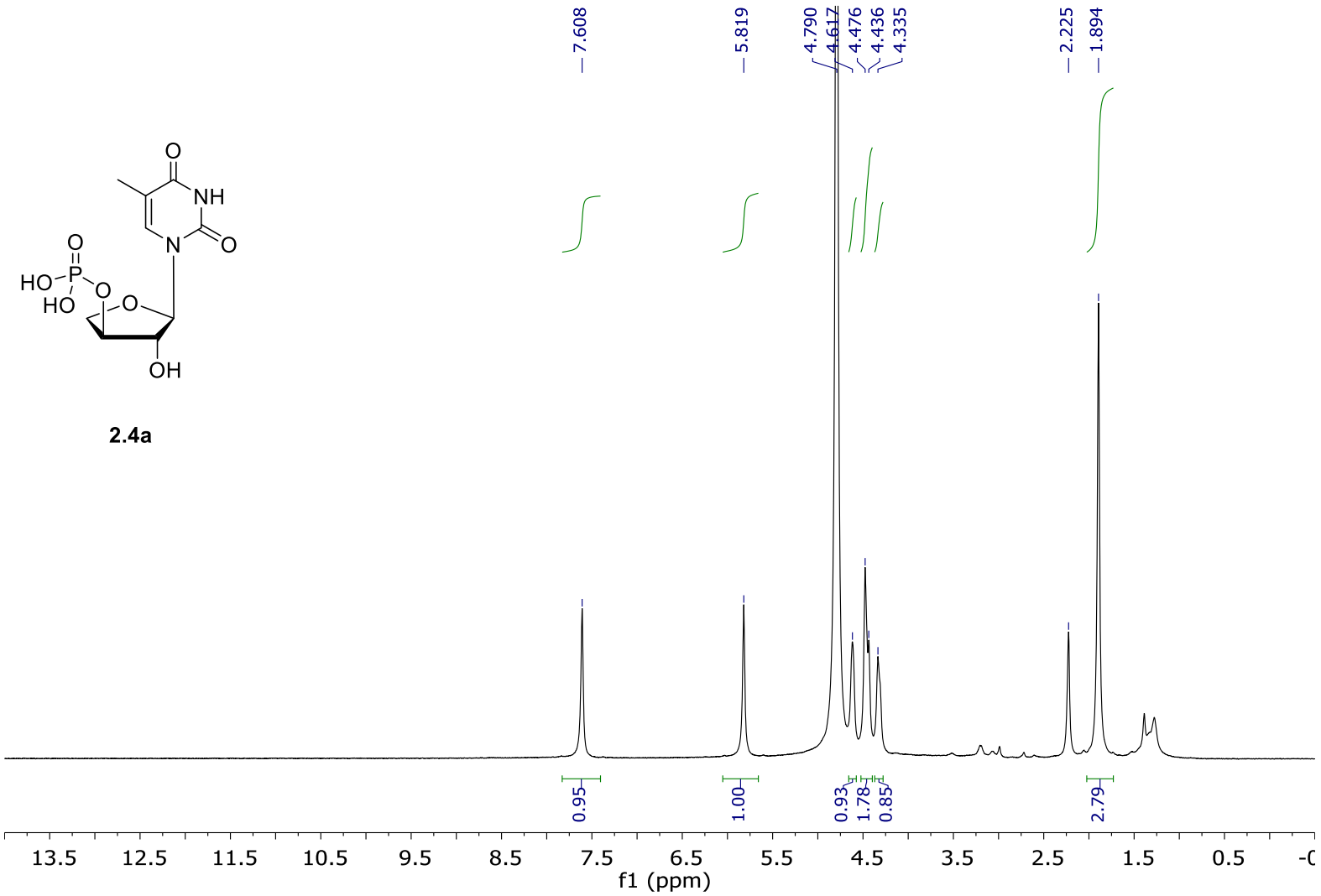
^{1}H NMR spectrum of compound **2.3d** (400 MHz, DMSO-d_6)



^{13}C NMR spectrum of compound **2.3d** (125.8 MHz, DMSO-d_6)

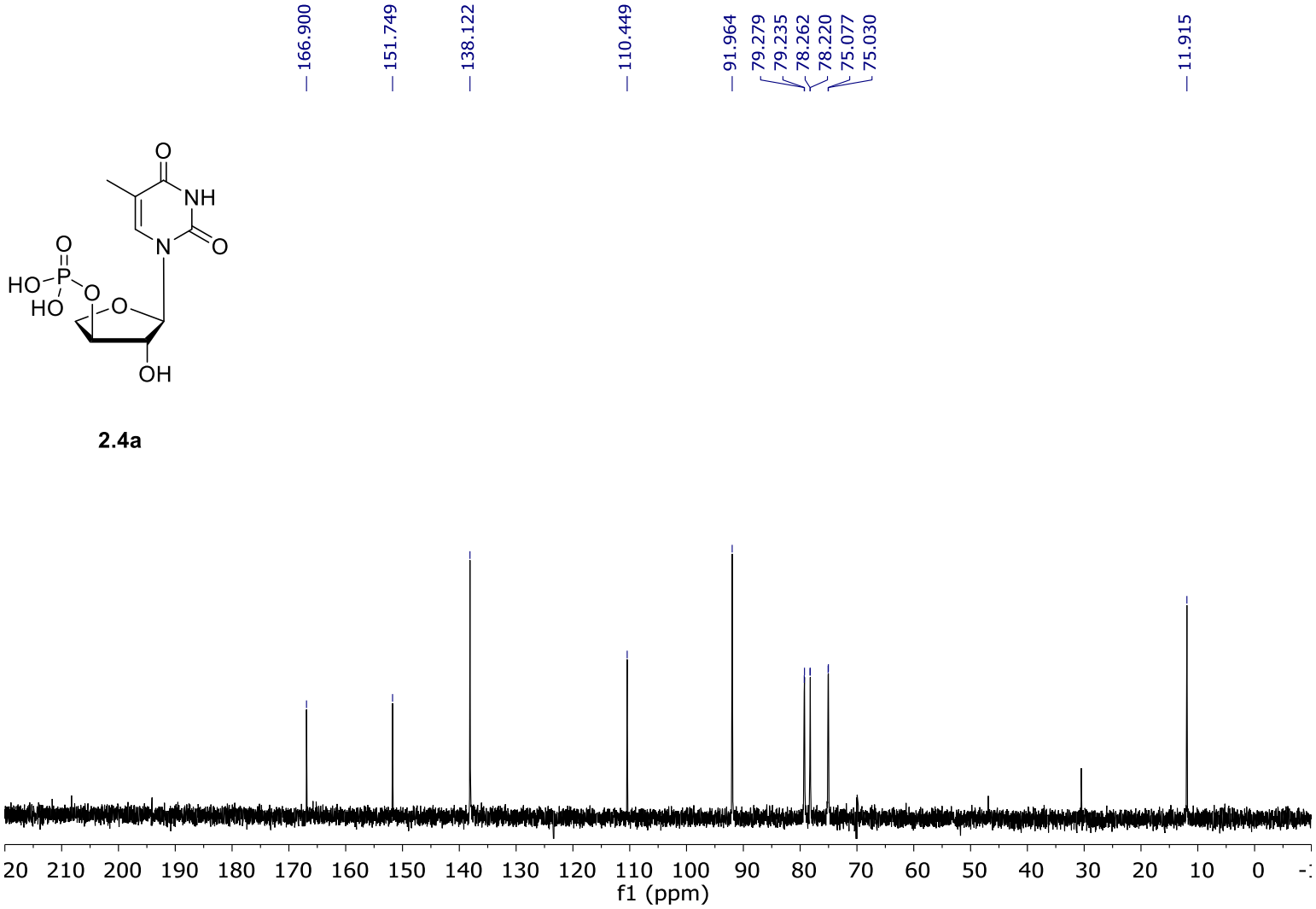


^{31}P NMR spectrum of compound **2.3d** (162 MHz, DMSO-d_6)

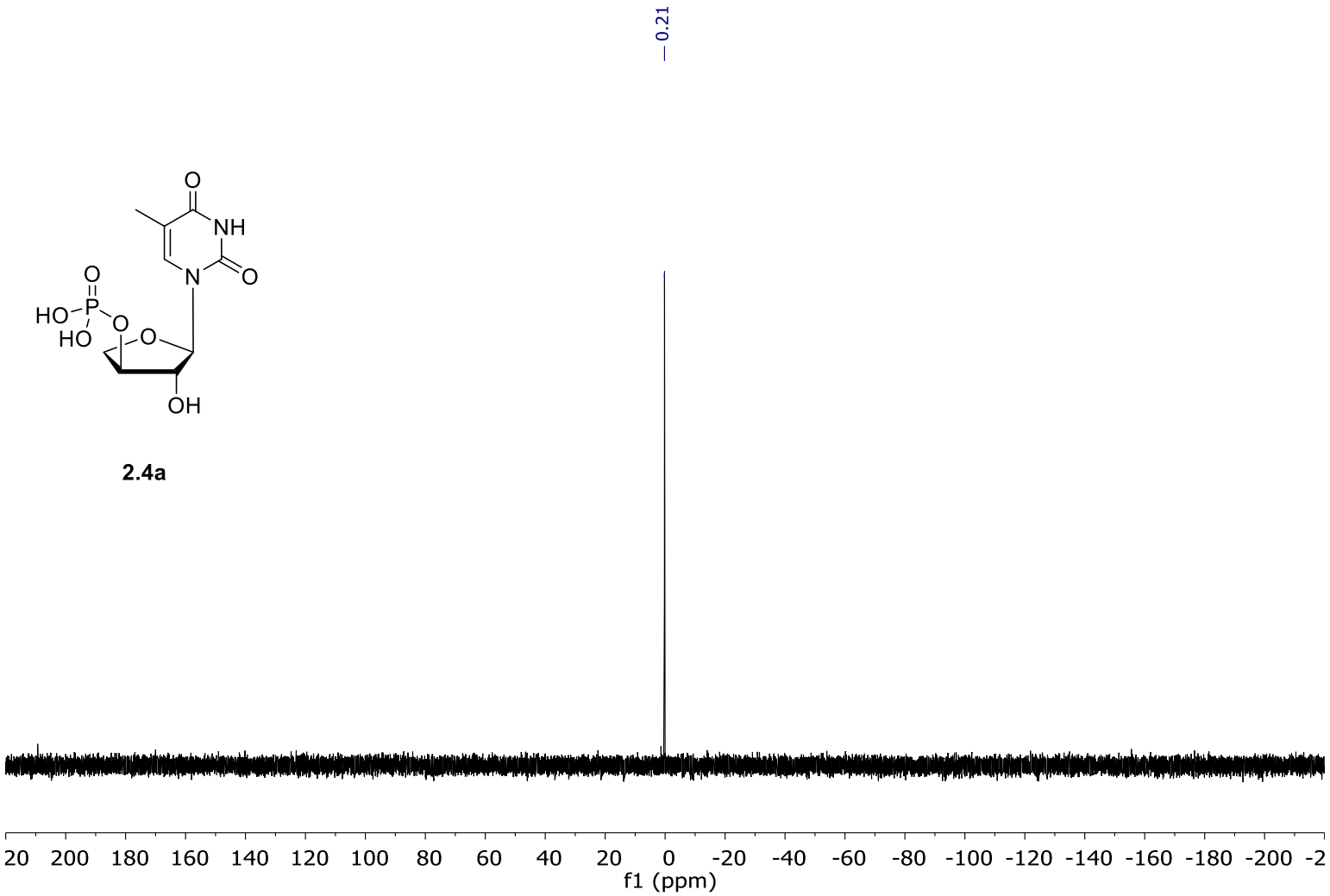


2.4a

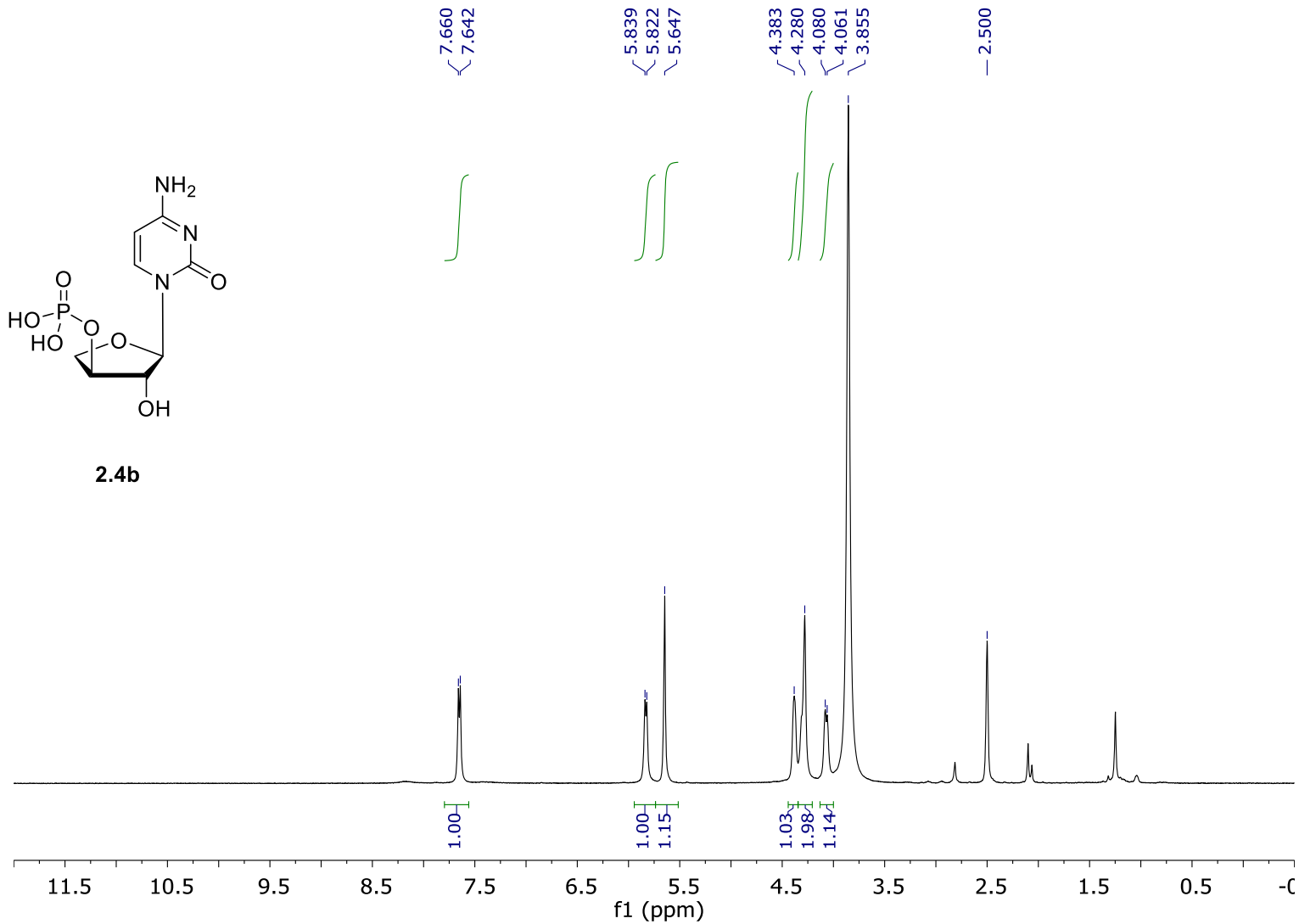
^{1}H NMR spectrum of compound **2.4a** (400 MHz, D_2O)



^{13}C NMR spectrum of compound **2.4a** (162 MHz, D_2O)

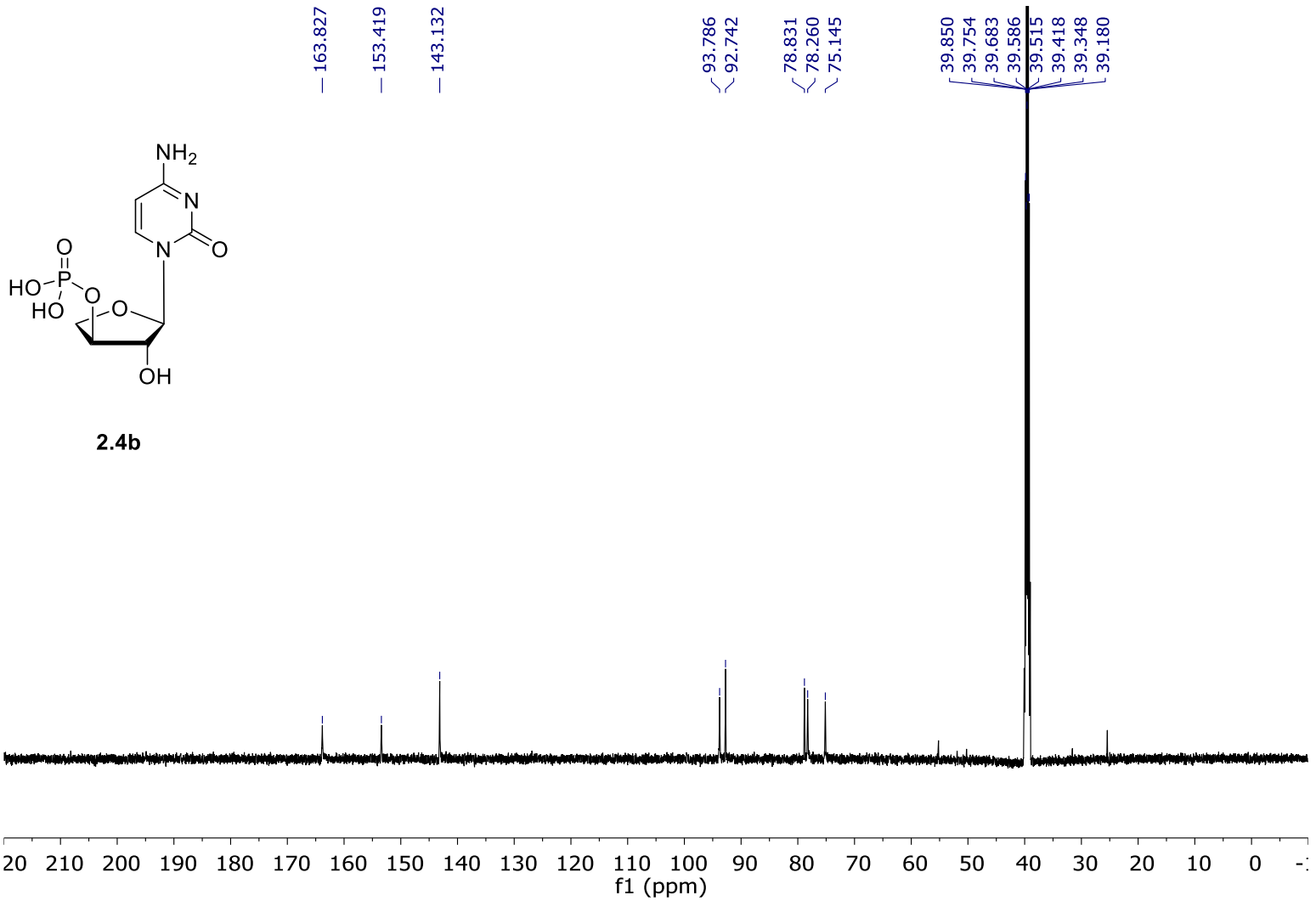


^{31}P NMR spectrum of compound **2.4a** (162 MHz, D_2O)

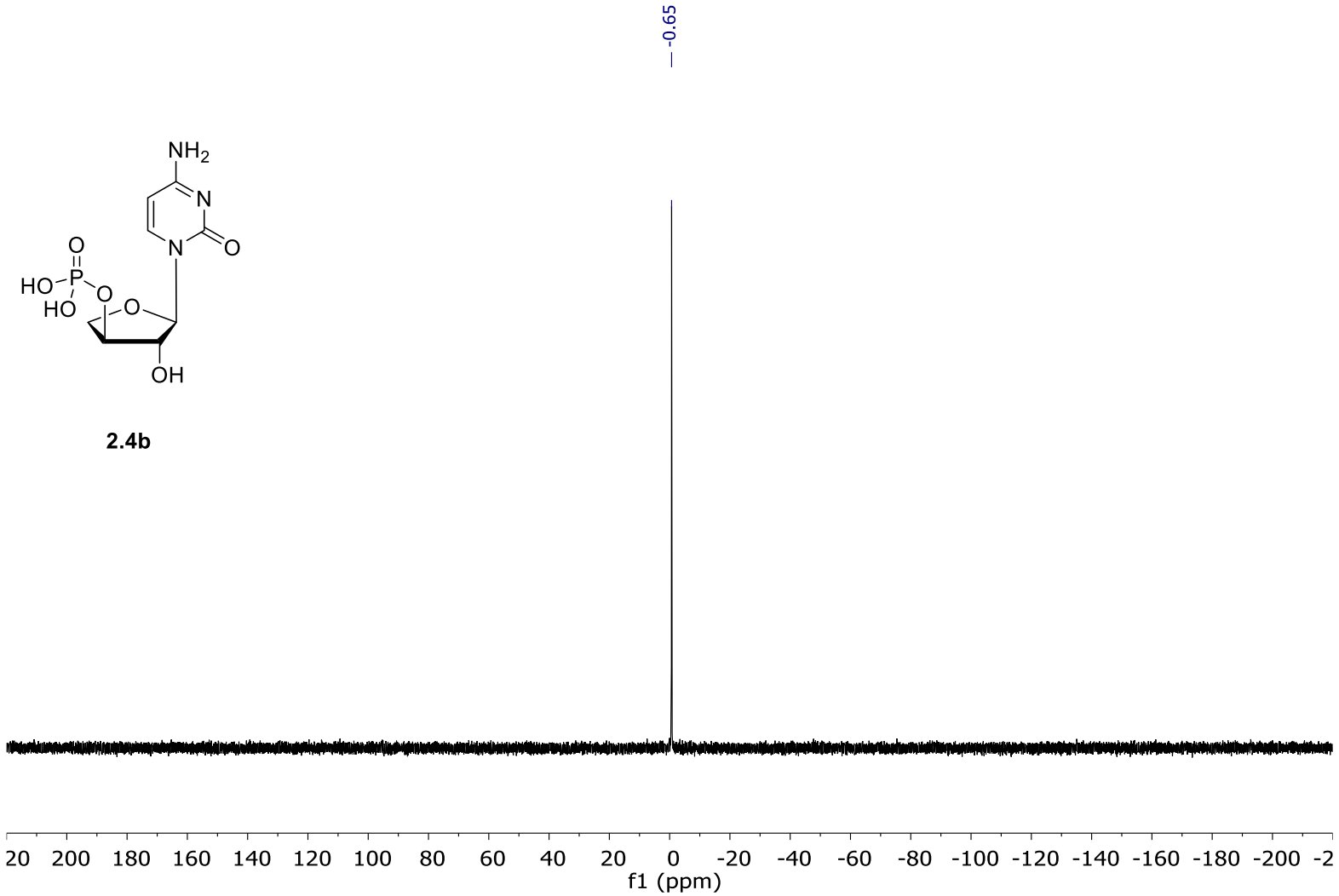


2.4b

^{1}H NMR spectrum of compound **2.4b** (400 MHz, DMSO-d_6)

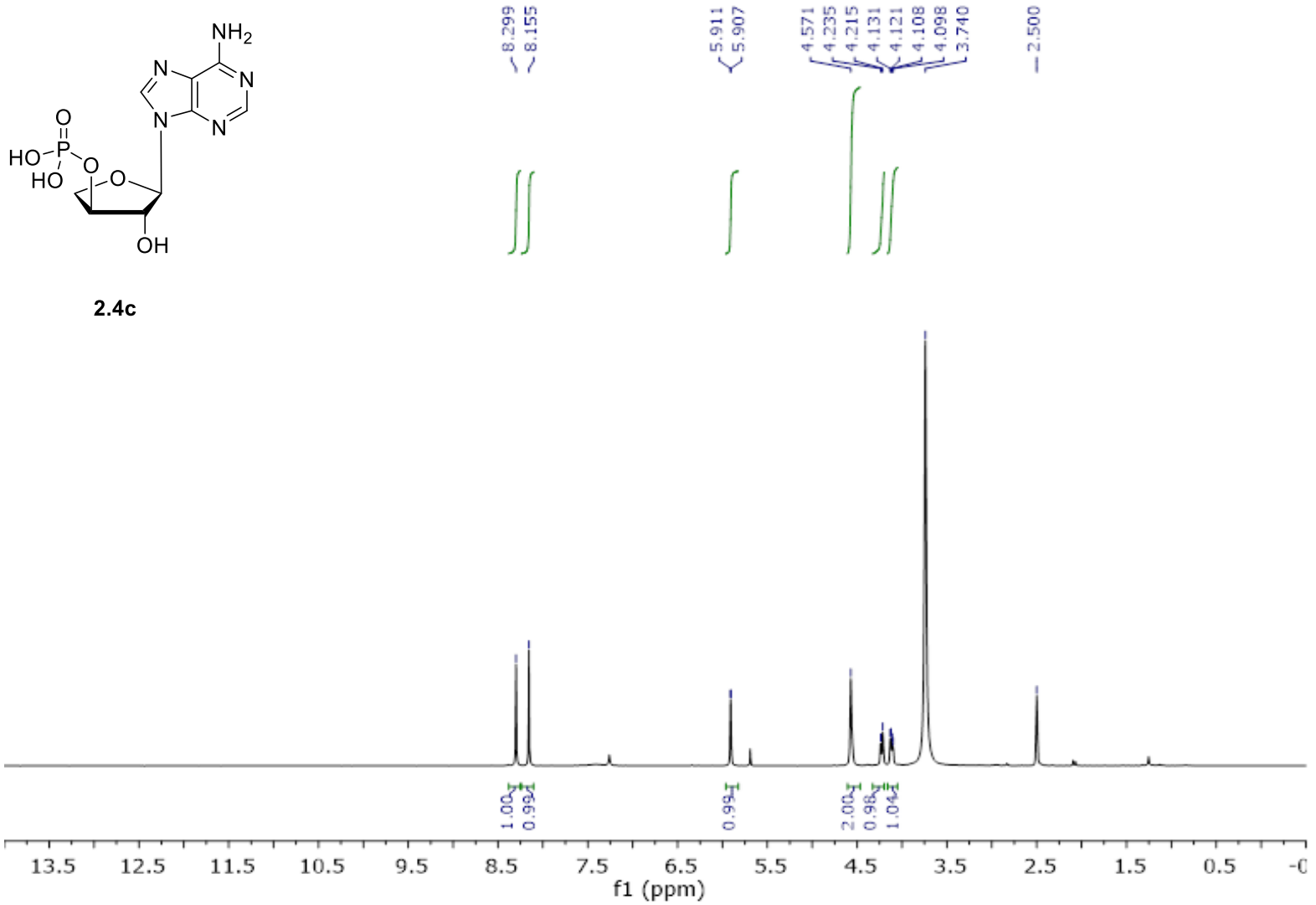


^{13}C NMR spectrum of compound **2.4b** (125.8 MHz, DMSO-d_6)



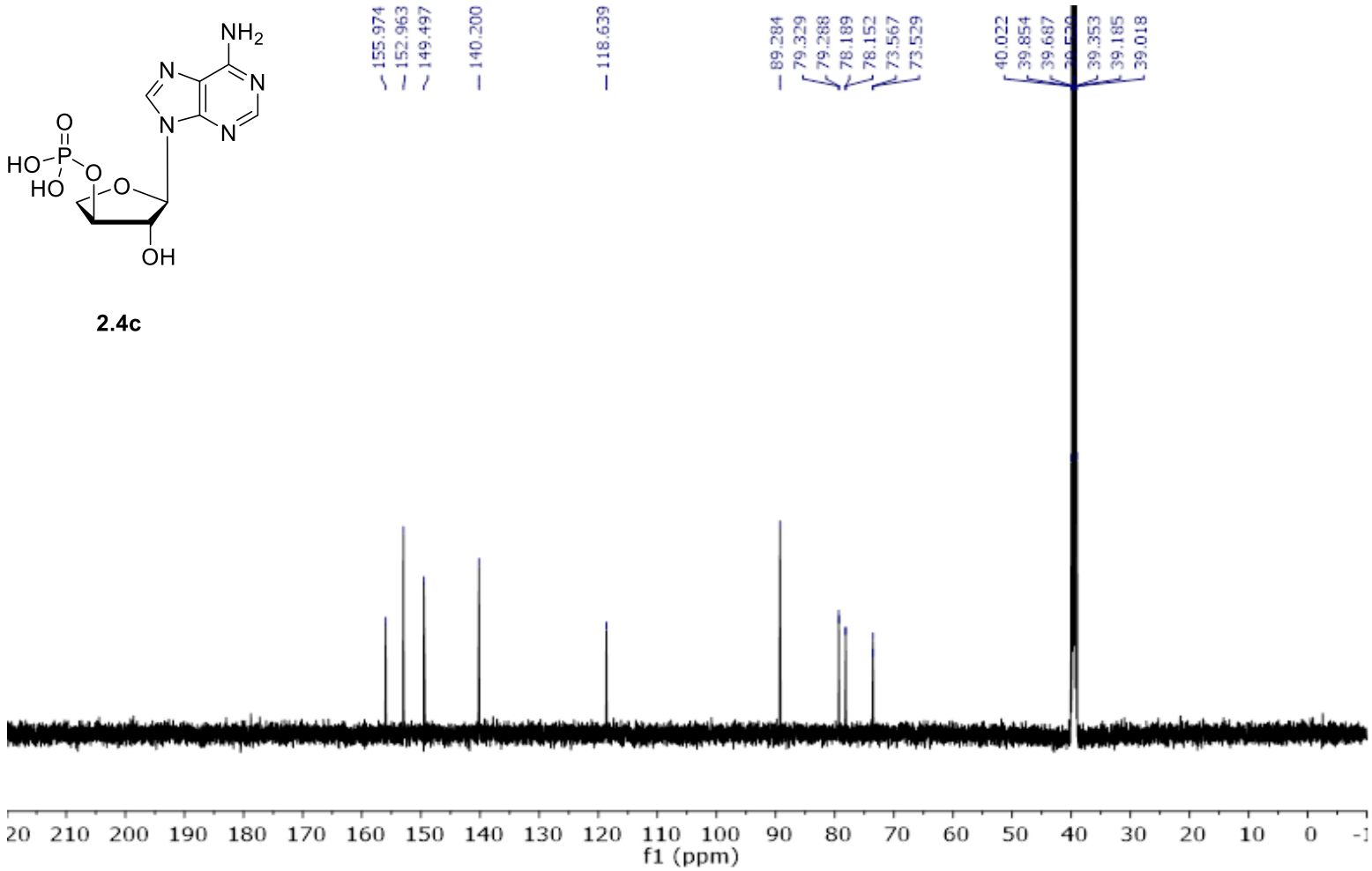
2.4b

^{31}P NMR spectrum of compound **2.4b** (162 MHz, D_2O)



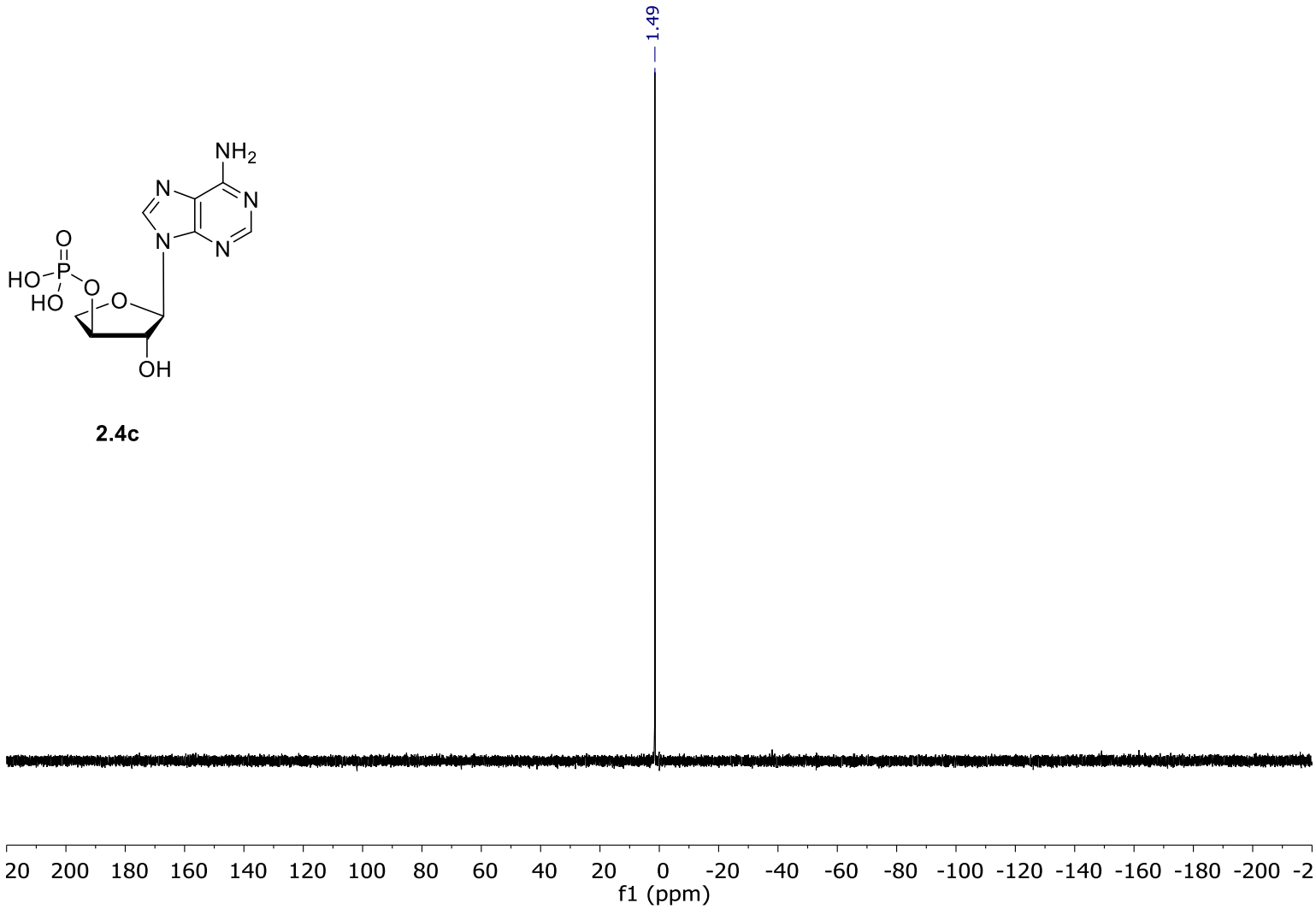
2.4c

^{1}H NMR spectrum of compound **2.4c** (400 MHz, DMSO-d_6)



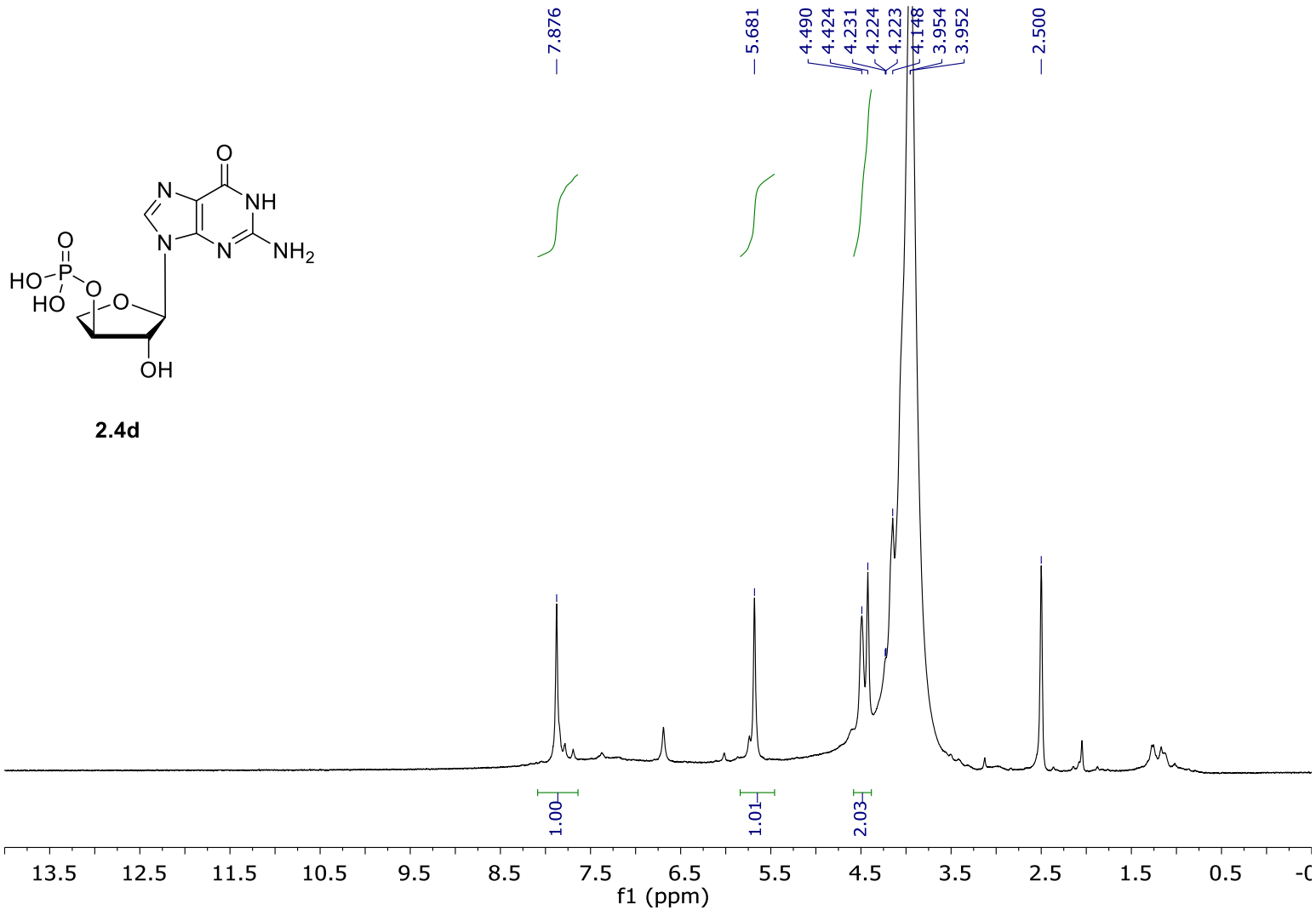
2.4c

^{13}C NMR spectrum of compound **2.4c** (125.8 MHz, DMSO-d_6)



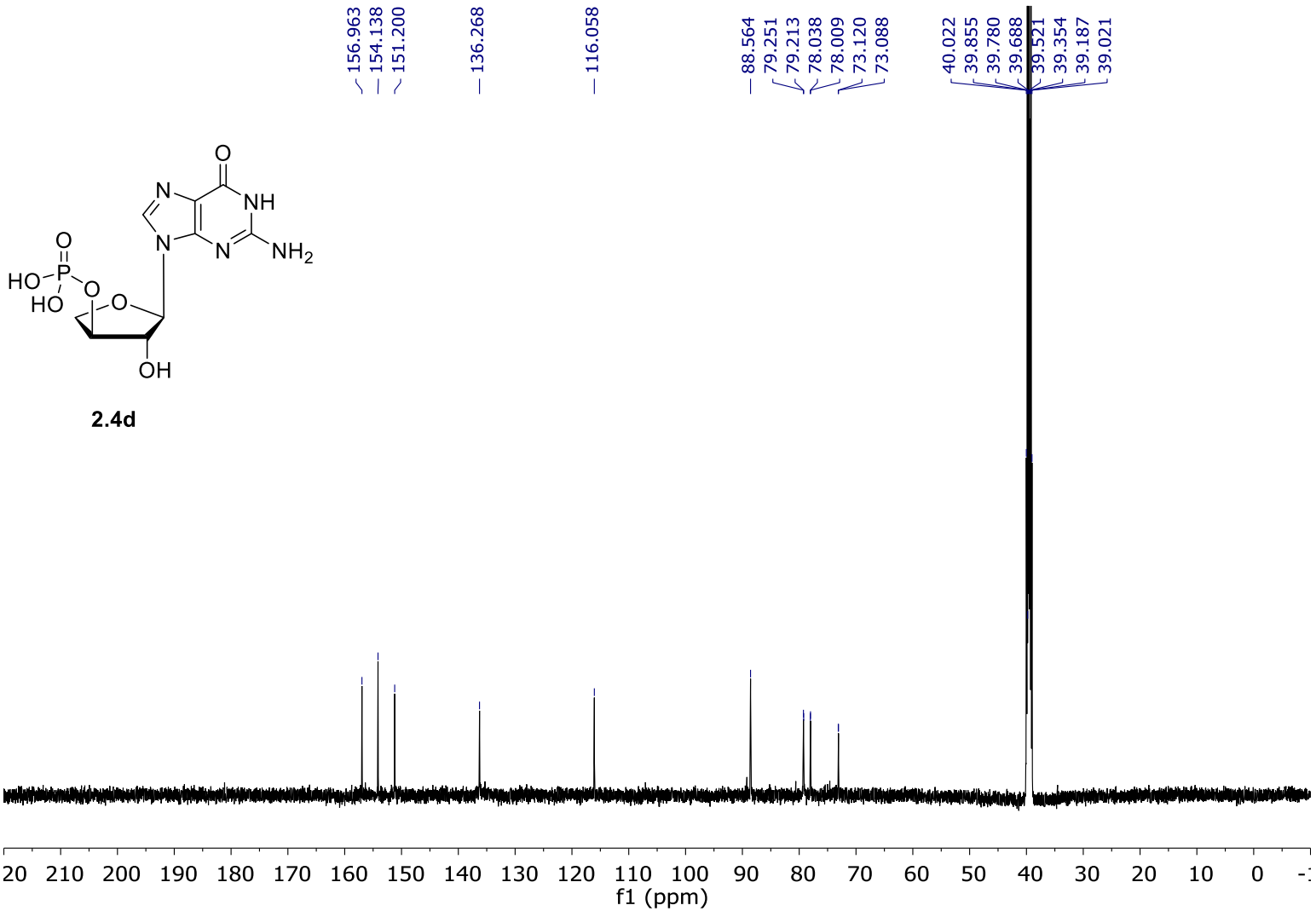
2.4c

^{31}P NMR spectrum of compound **2.4c** (162 MHz, D_2O)

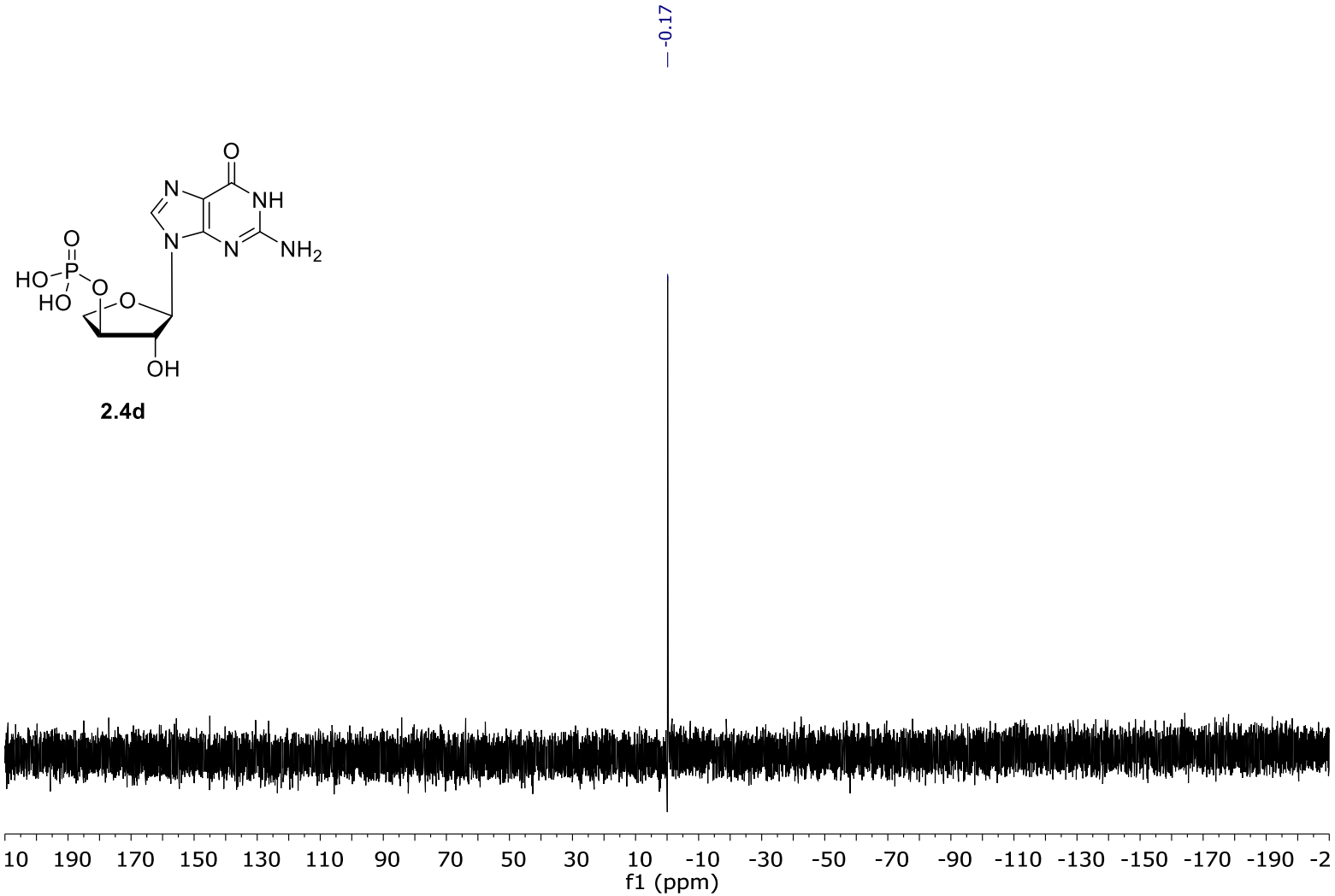


2.4d

^{1}H NMR spectrum of compound **2.4d** (400 MHz, DMSO-d_6)

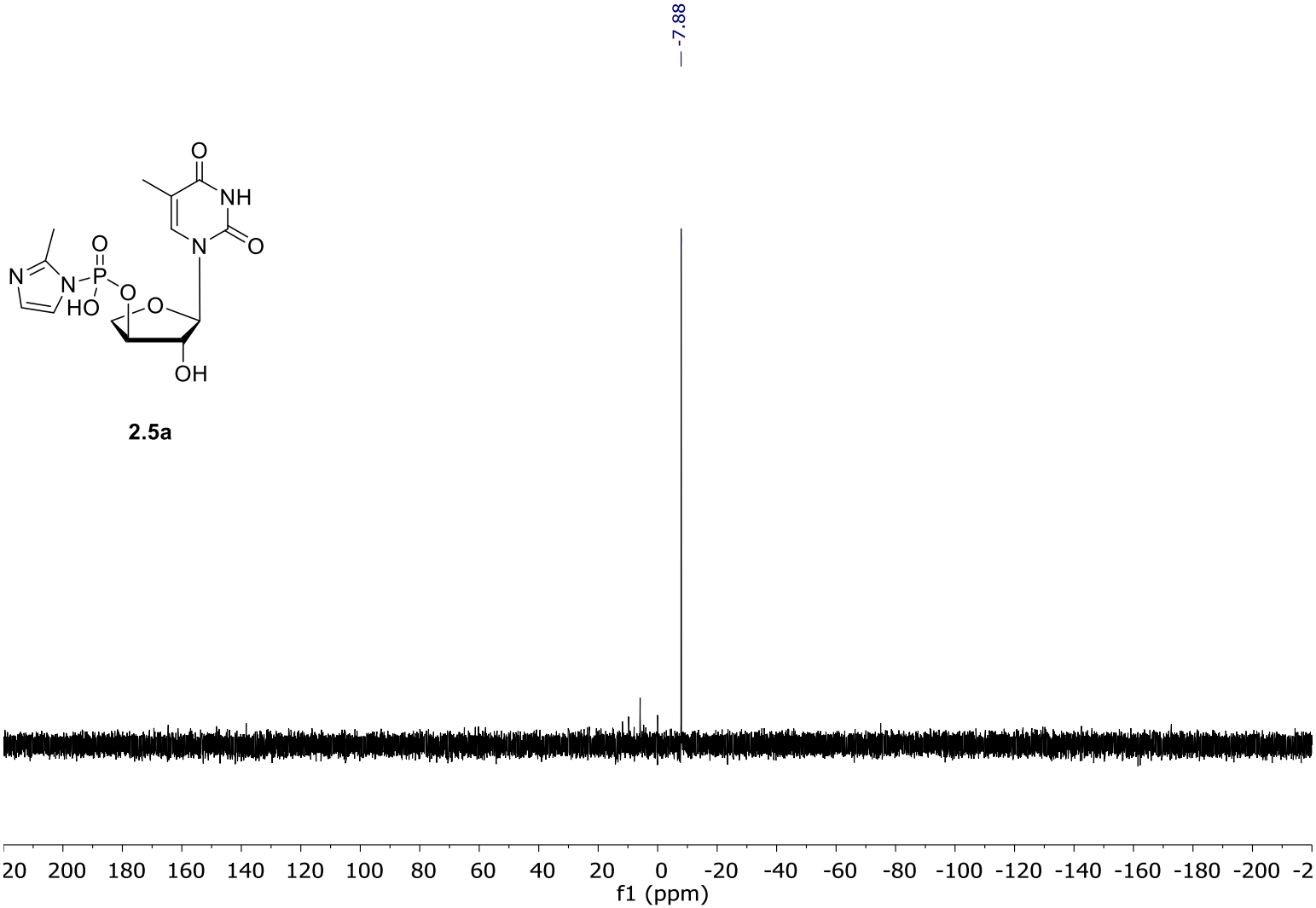


^{13}C NMR spectrum of compound **2.4d** (100.6 MHz, DMSO-d_6)



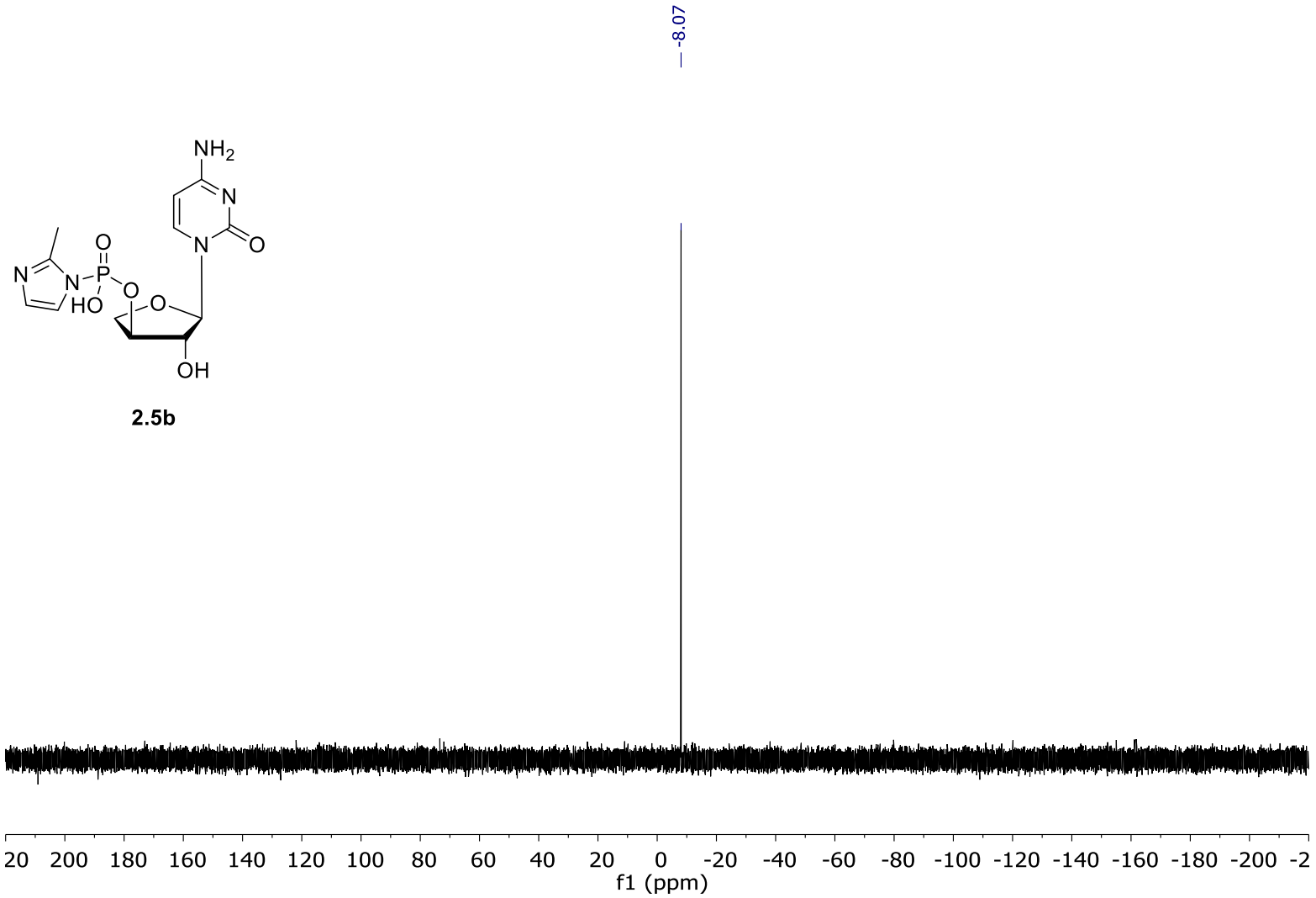
2.4d

^{31}P NMR spectrum of compound **2.4d** (162 MHz, D_2O)

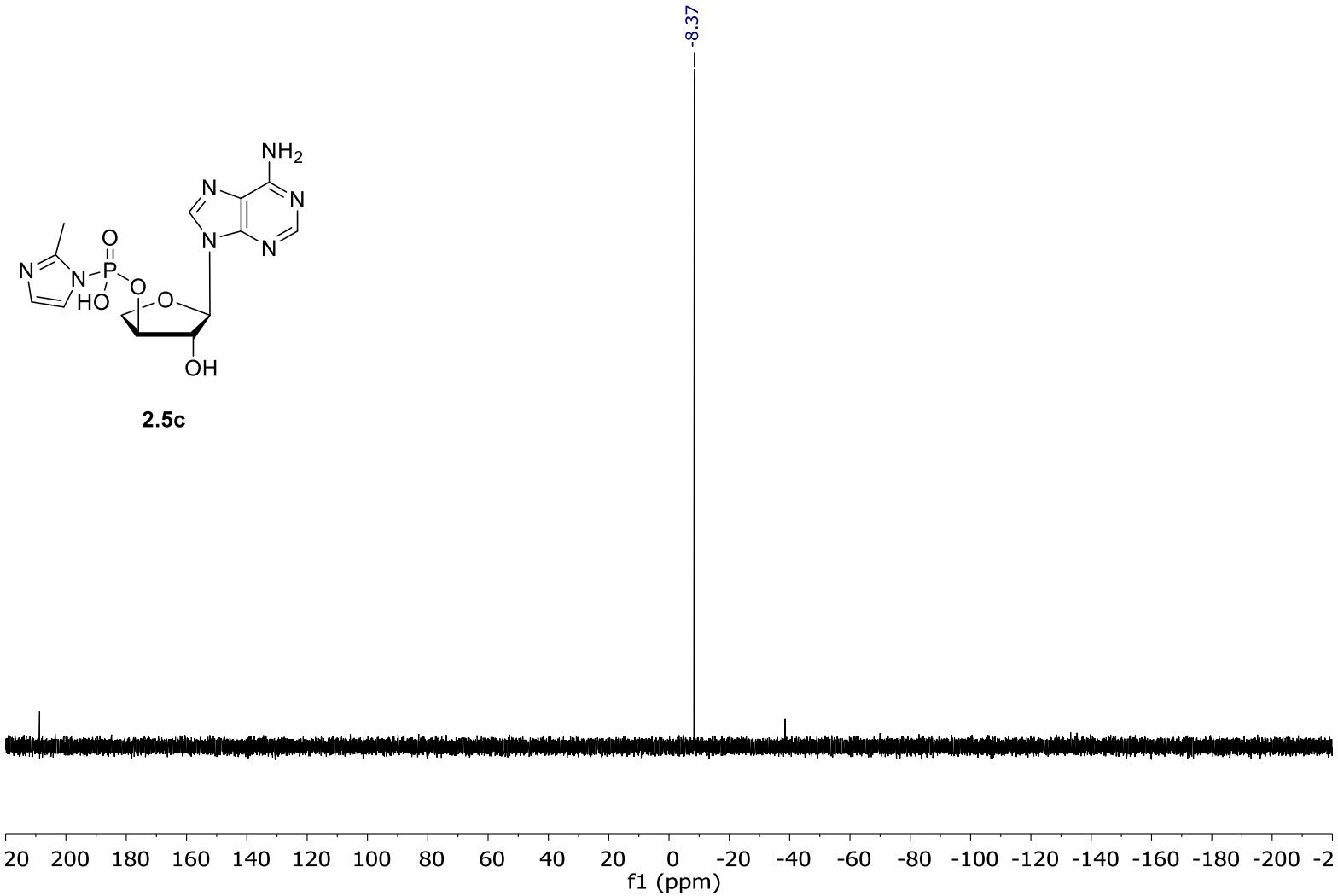


2.5a

^{31}P NMR spectrum of compound **2.5a** (162 MHz, D_2O)

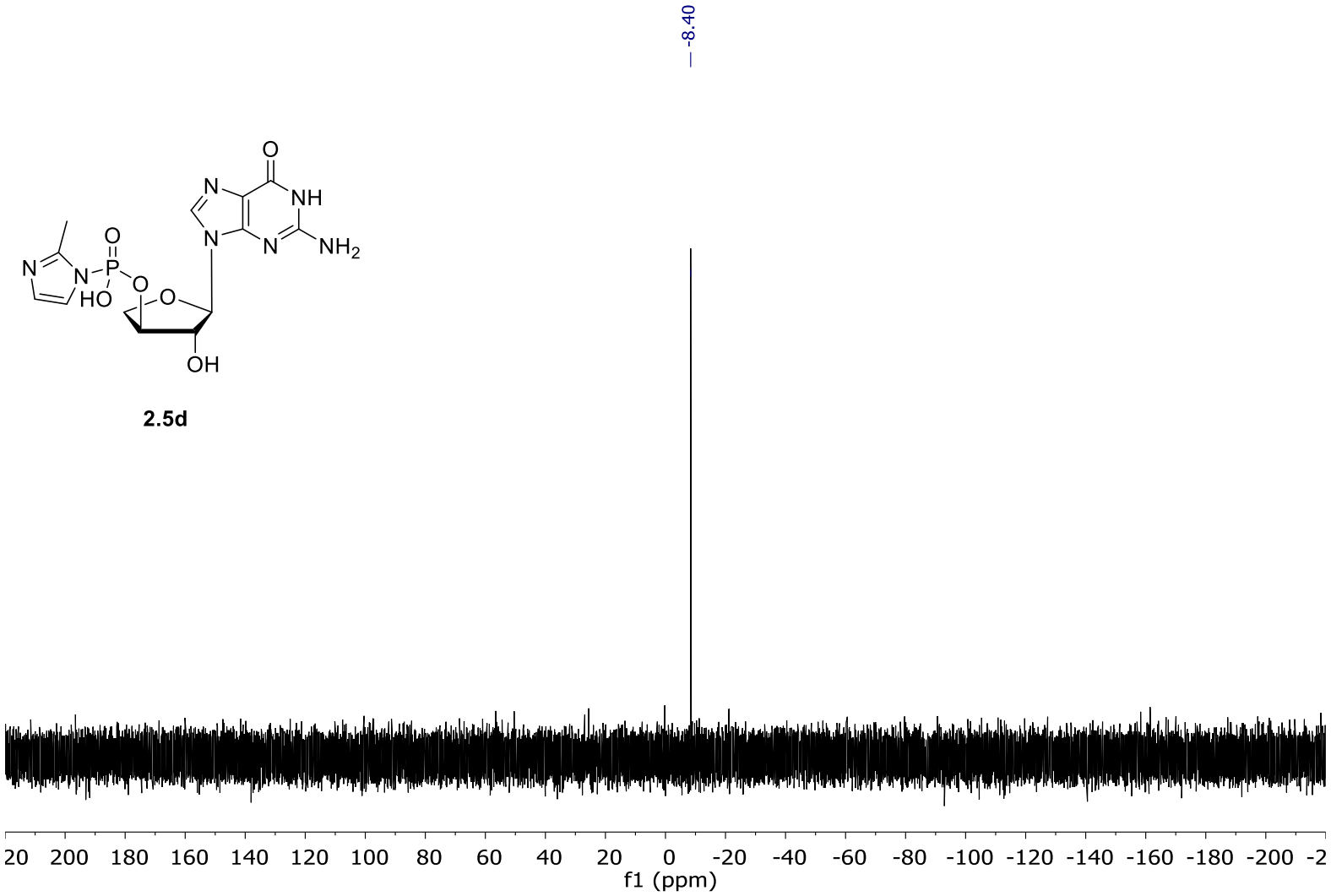


^{31}P NMR spectrum of compound **2.5b** (162 MHz, D_2O)

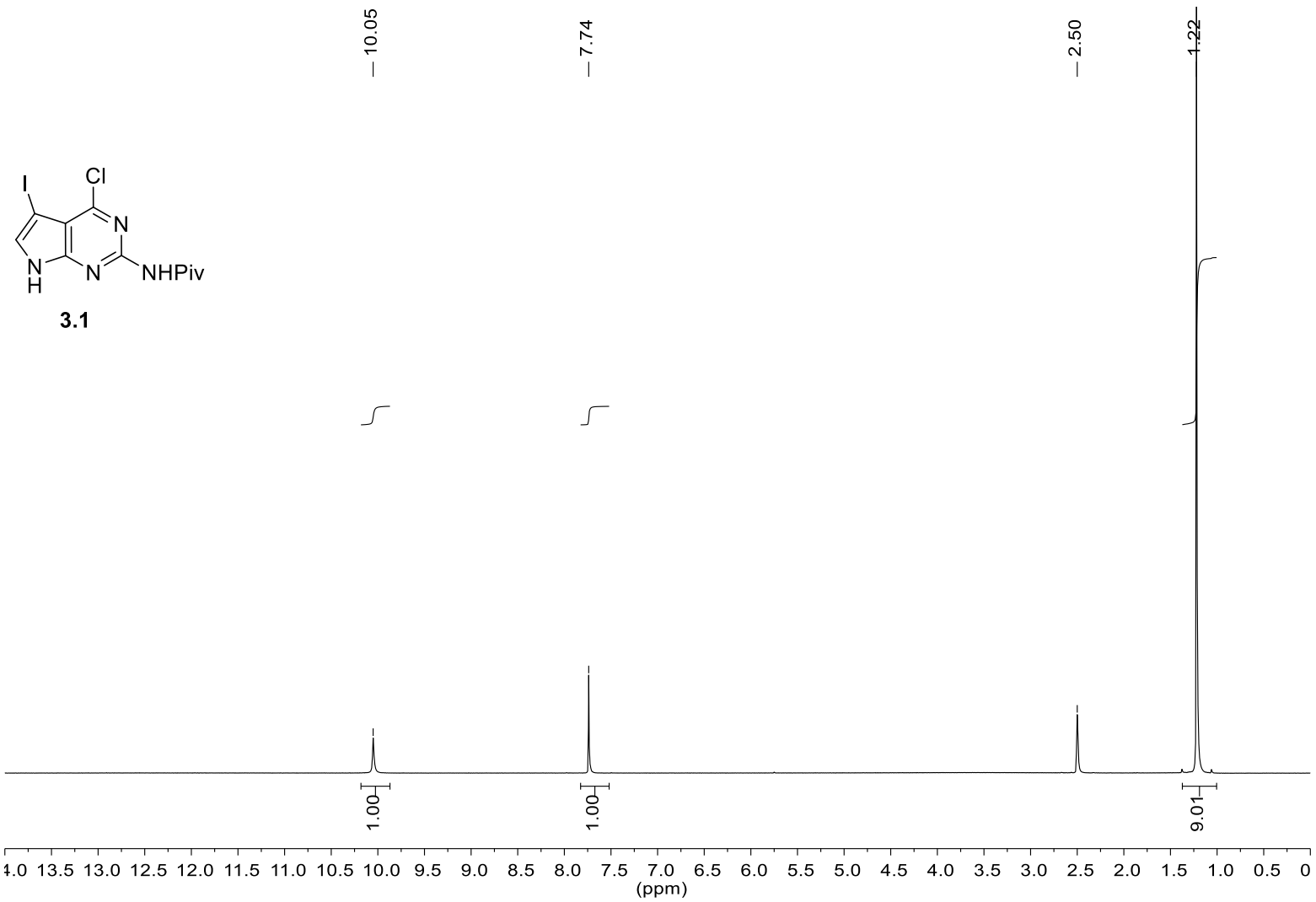


2.5c

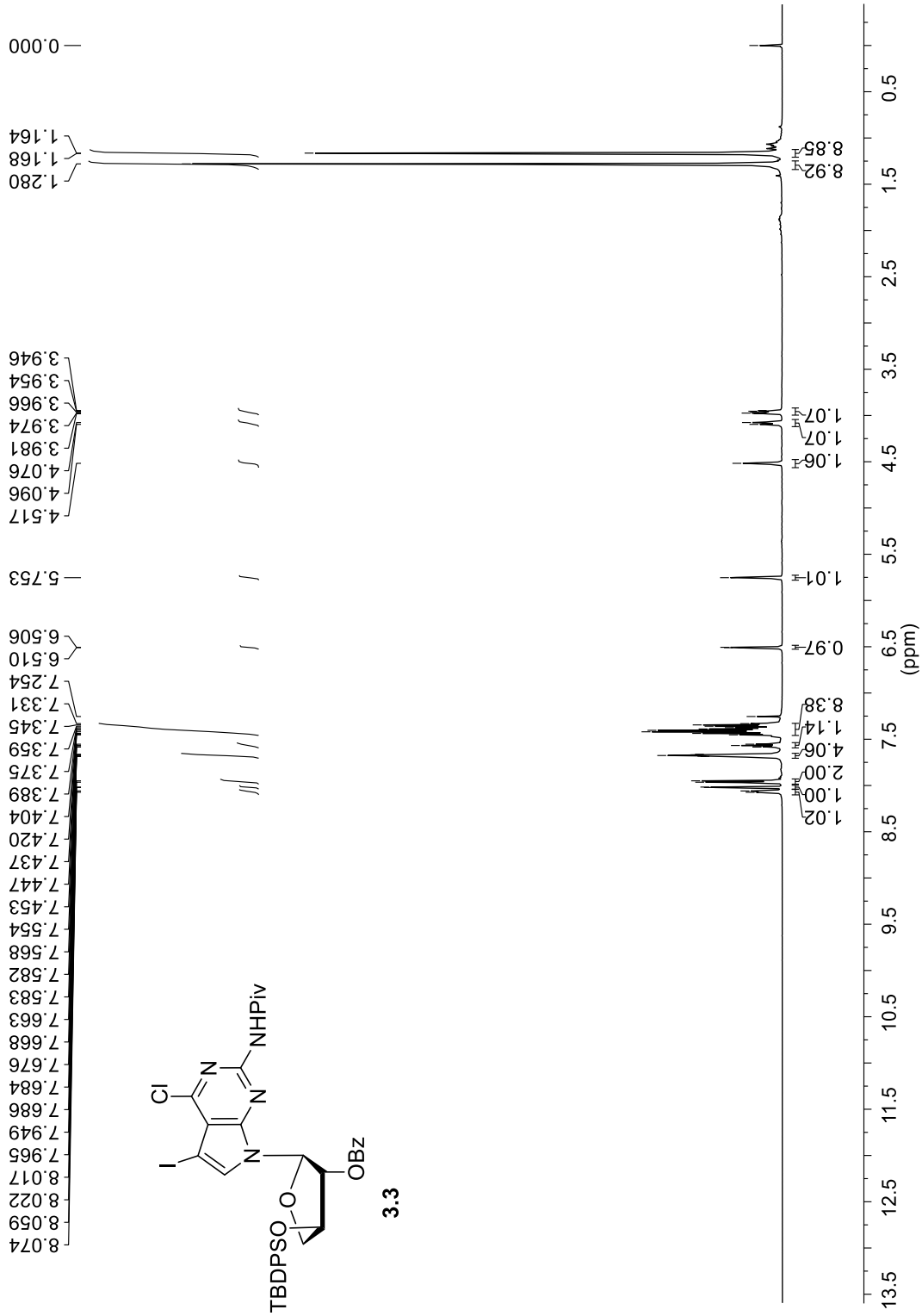
^{31}P NMR spectrum of compound **2.5c** (162 MHz, D_2O)



^{31}P NMR spectrum of compound **2.5d** (162 MHz, D_2O)

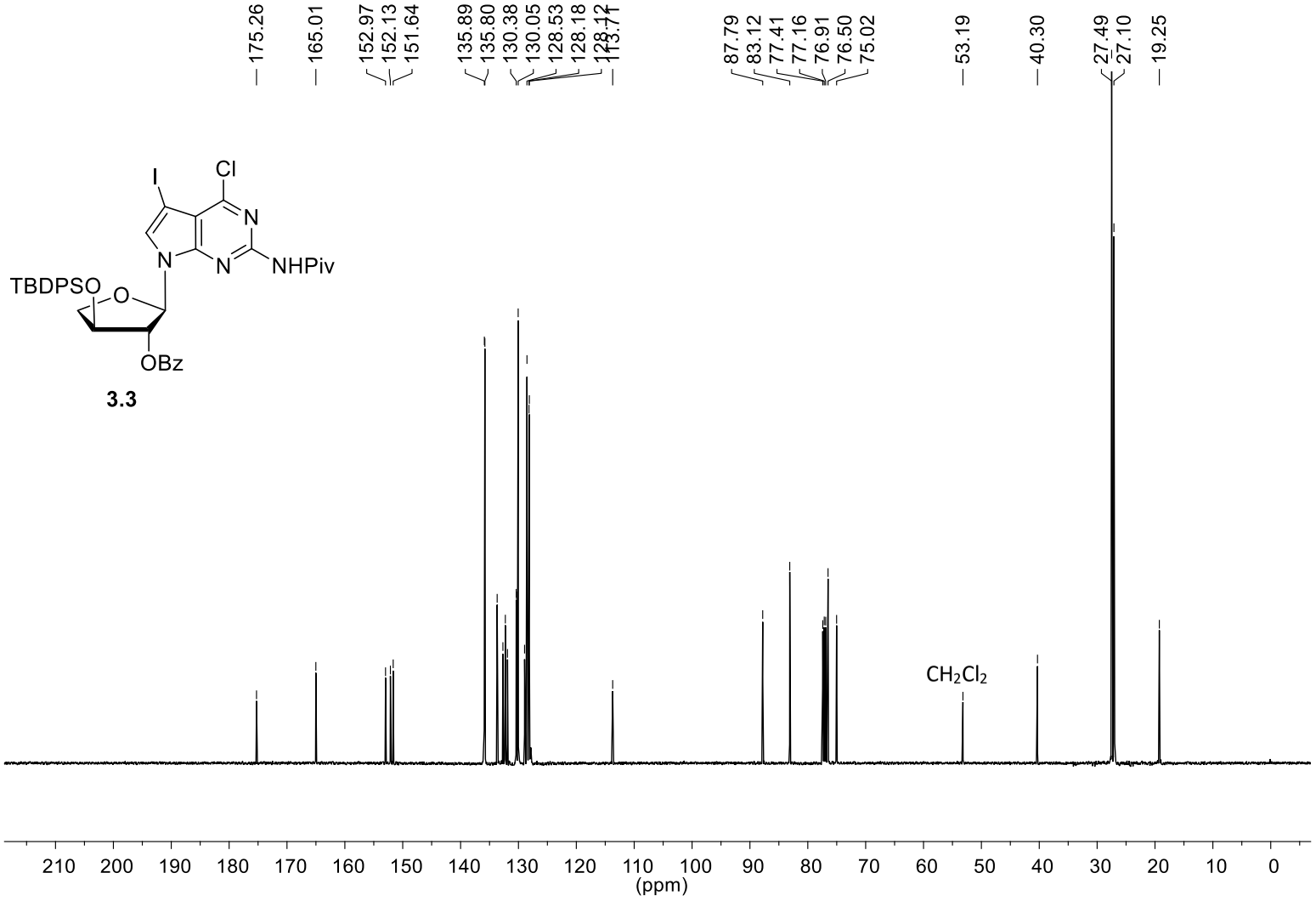


^{1}H NMR spectrum of compound **3.1** (400 MHz, DMSO-d_6)

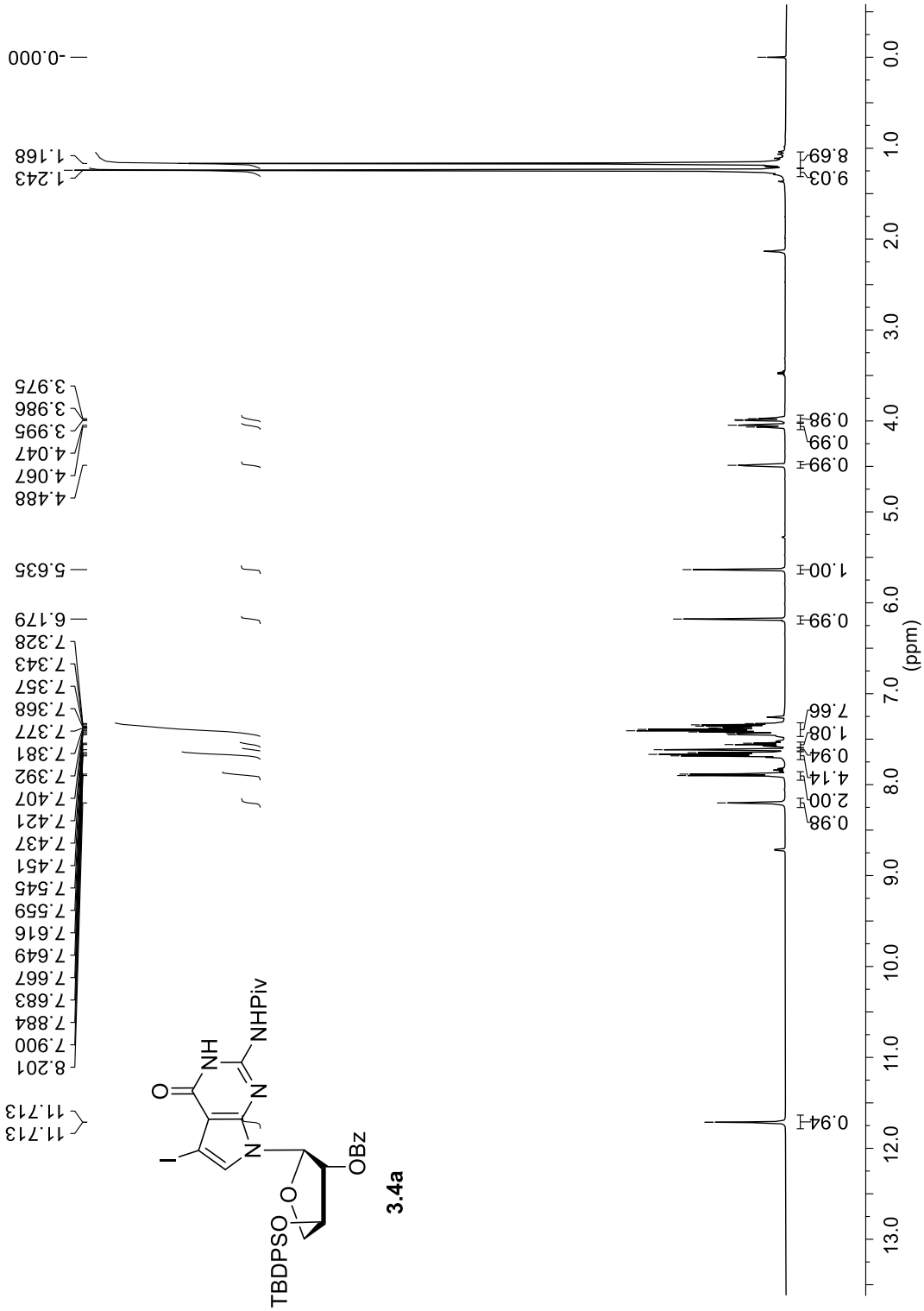


3.3

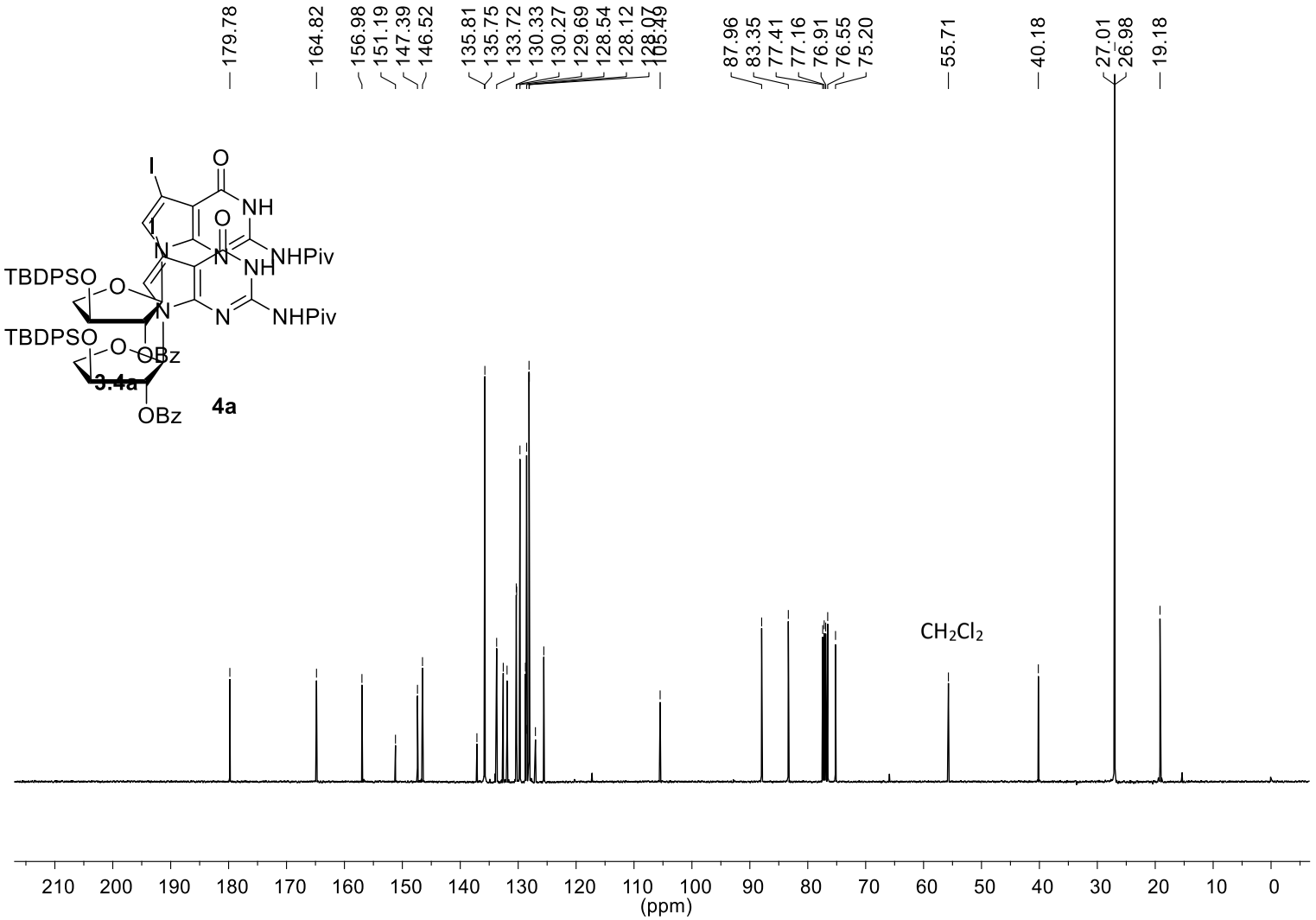
^{1}H NMR spectrum of compound **3.3** (500 MHz, $CDCl_3$).



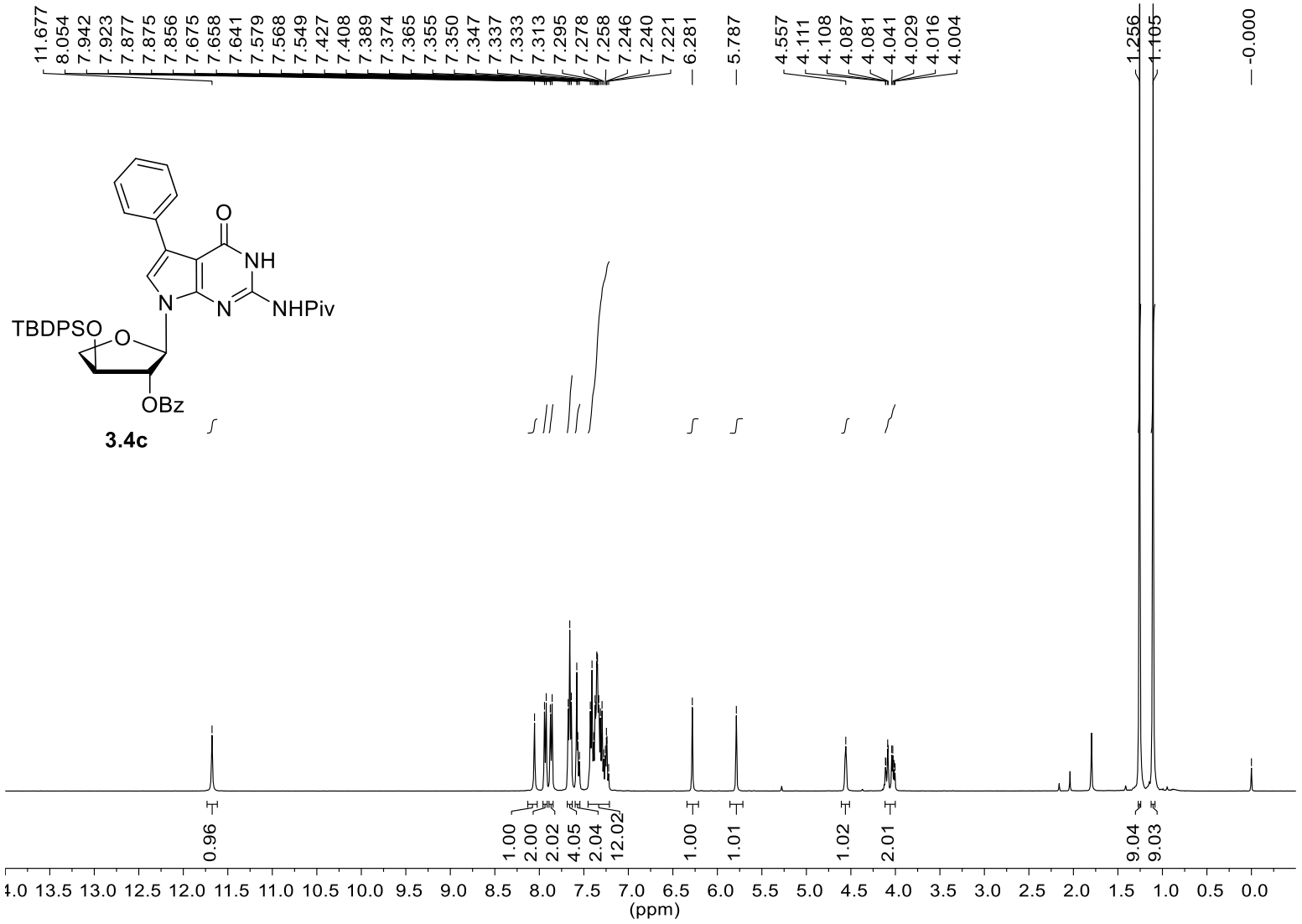
^{13}C NMR spectrum of compound **3.3** (125.8 MHz, $CDCl_3$).



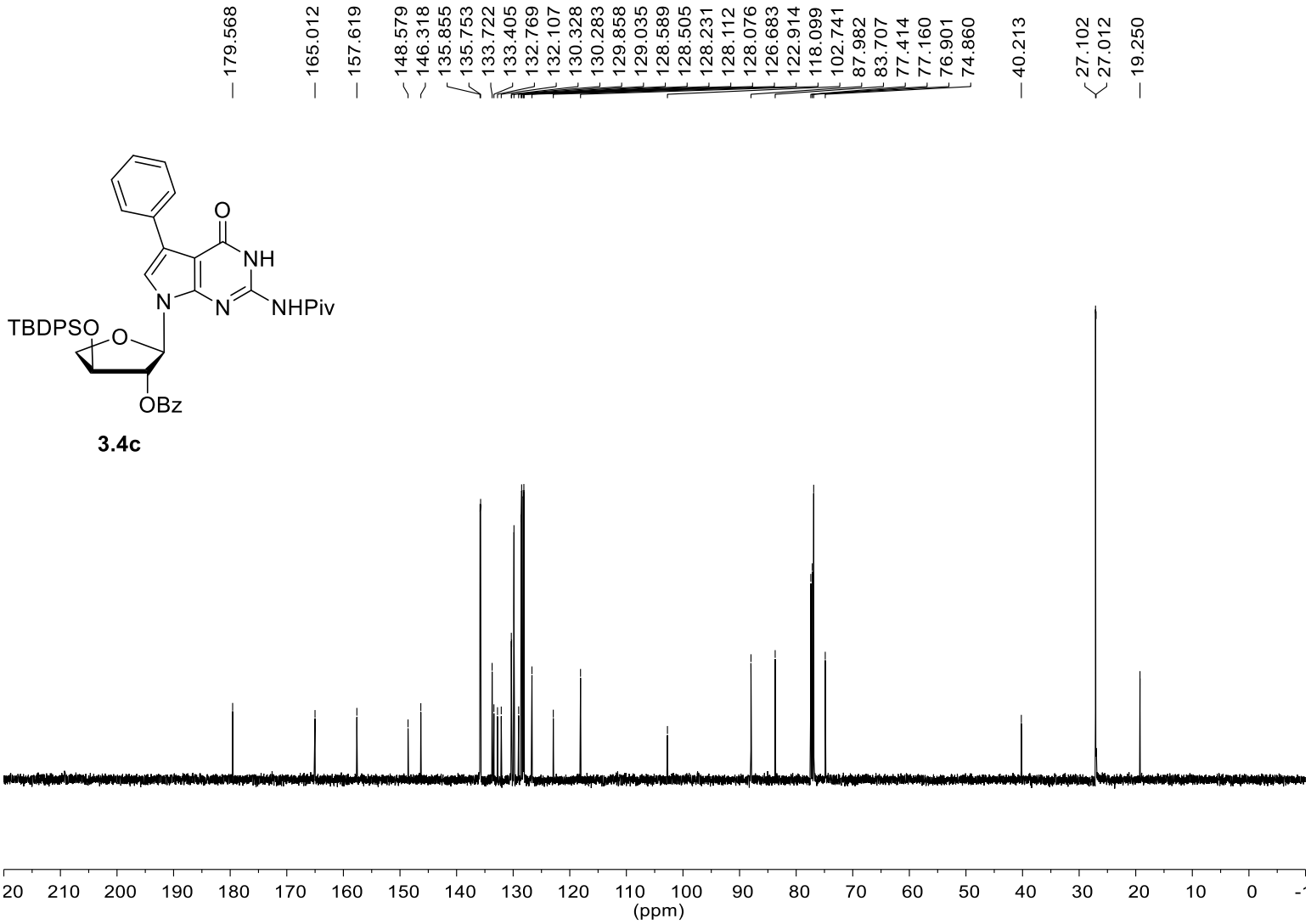
^{1}H NMR spectrum of compound **3.4a** (500 MHz, $CDCl_3$).



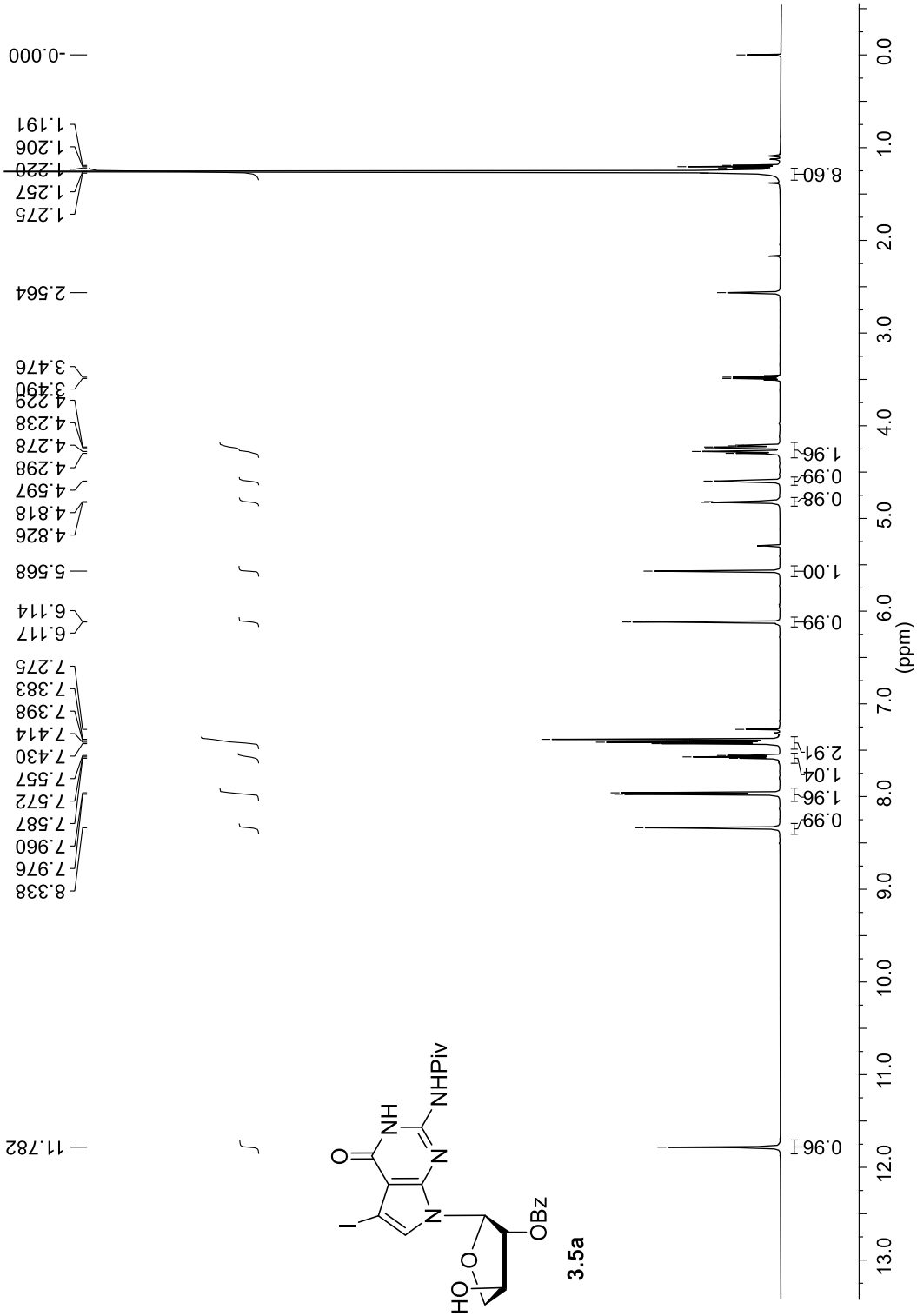
^{13}C NMR spectrum of compound **3.4a** (125.8 MHz, $CDCl_3$).



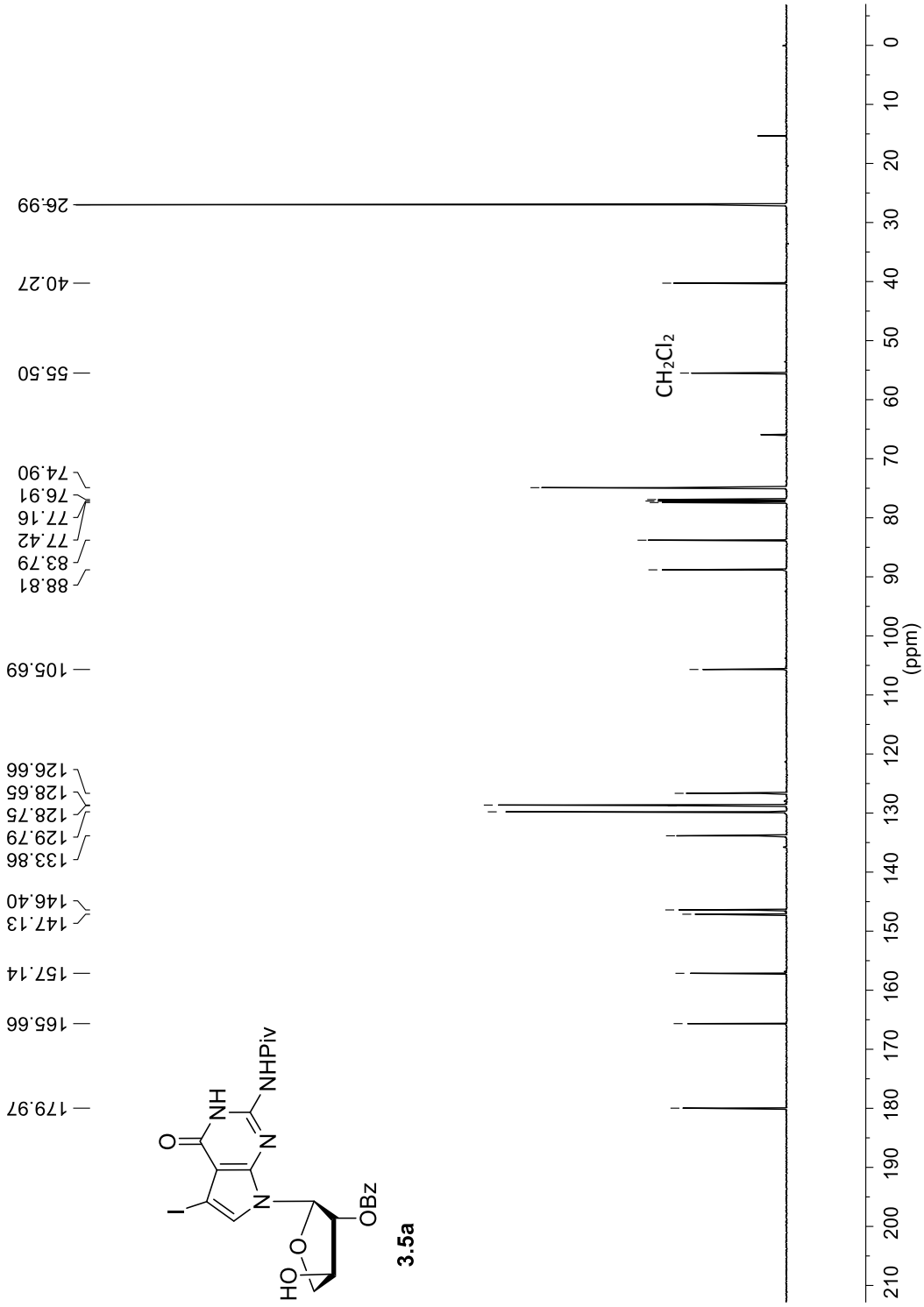
^{1}H NMR spectrum of compound **3.4c** (400 MHz, $CDCl_3$).



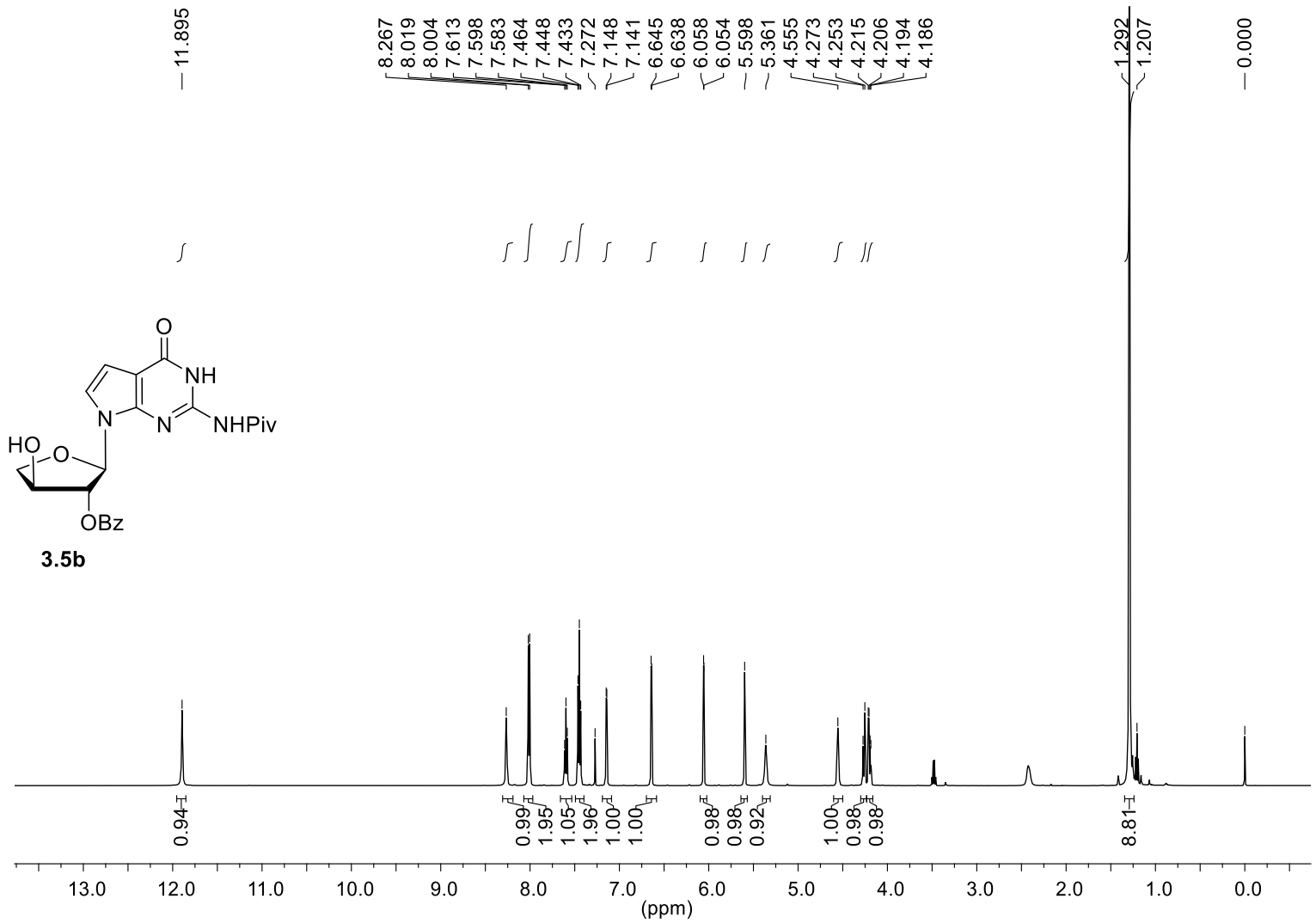
^{13}C NMR spectrum of compound **3.4c** (125.8 MHz, $CDCl_3$)



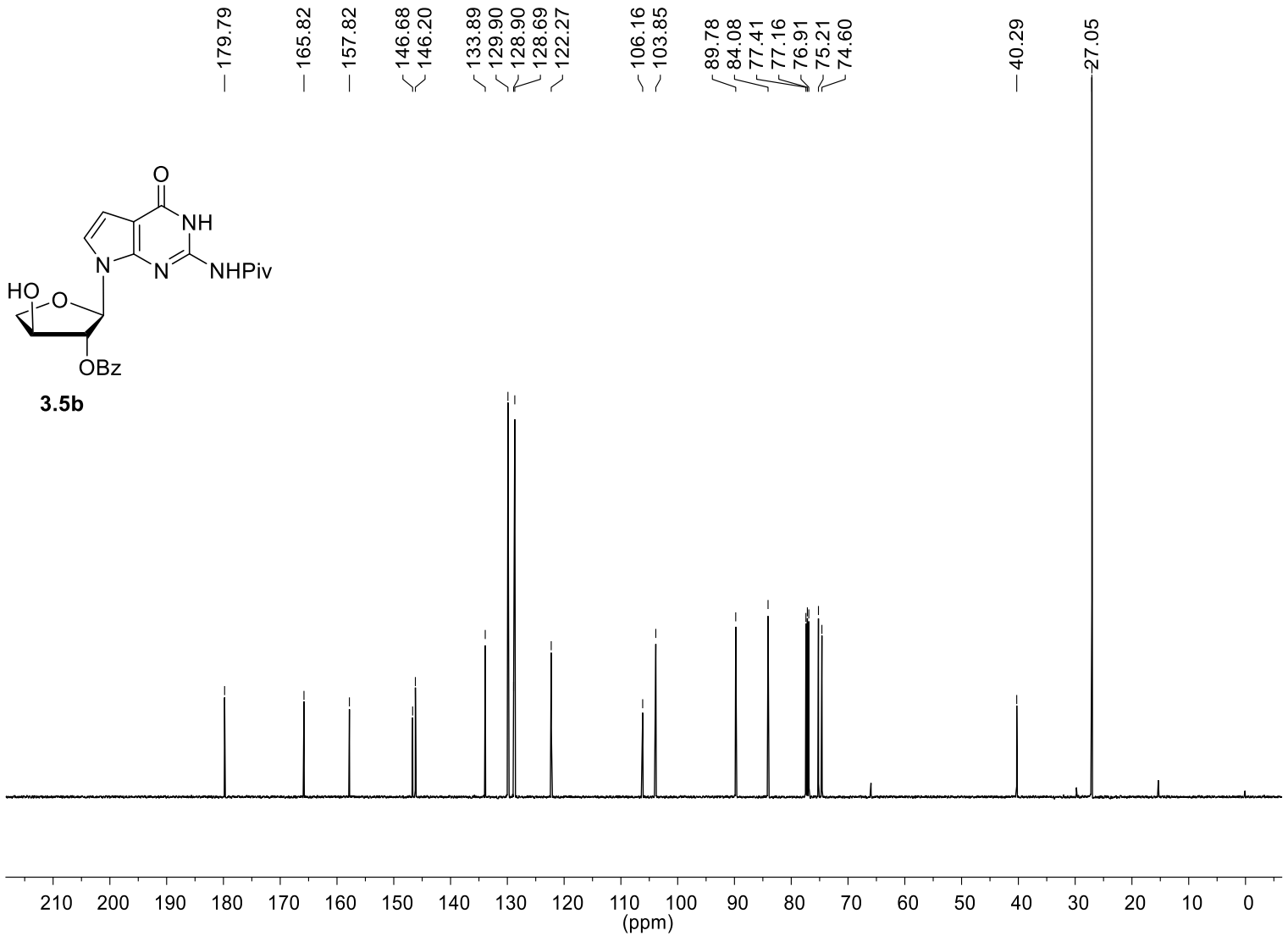
^{1}H NMR spectrum of compound **3.5a** (500 MHz, $CDCl_3$).



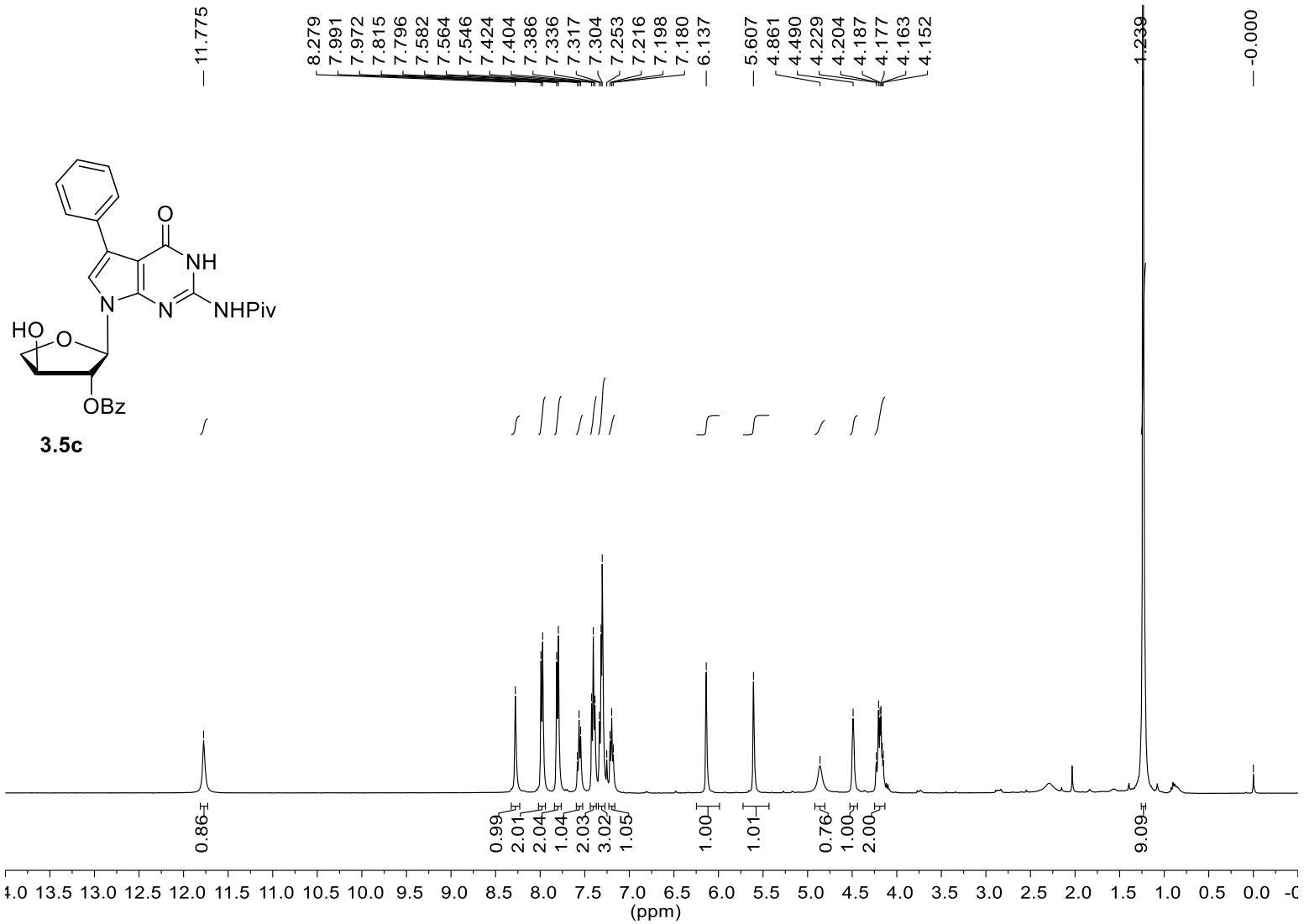
^{13}C NMR spectrum of compound **3.5a** (125.8 MHz, $CDCl_3$).



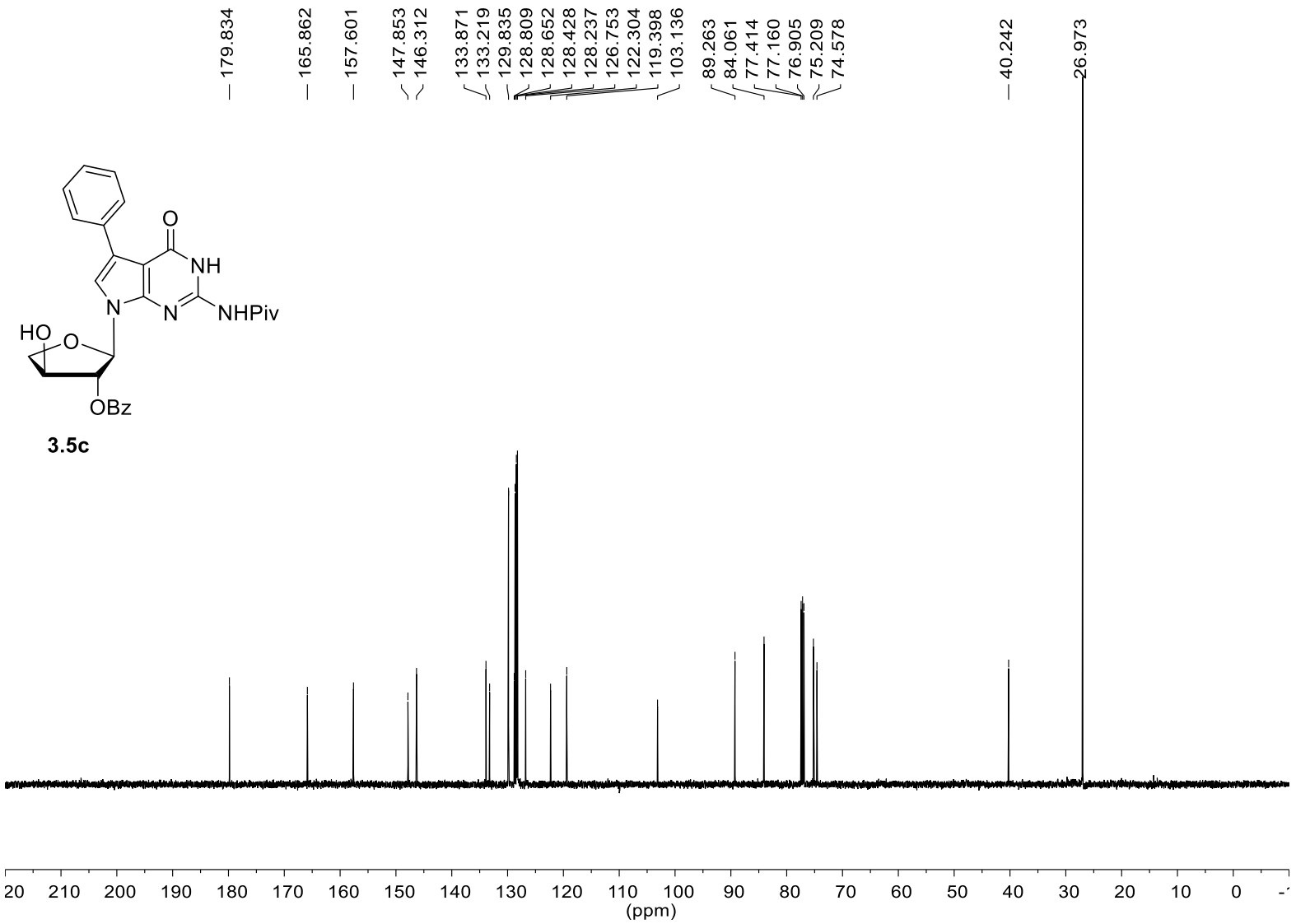
^{1}H NMR spectrum of compound **3.5b** (500 MHz, $CDCl_3$).



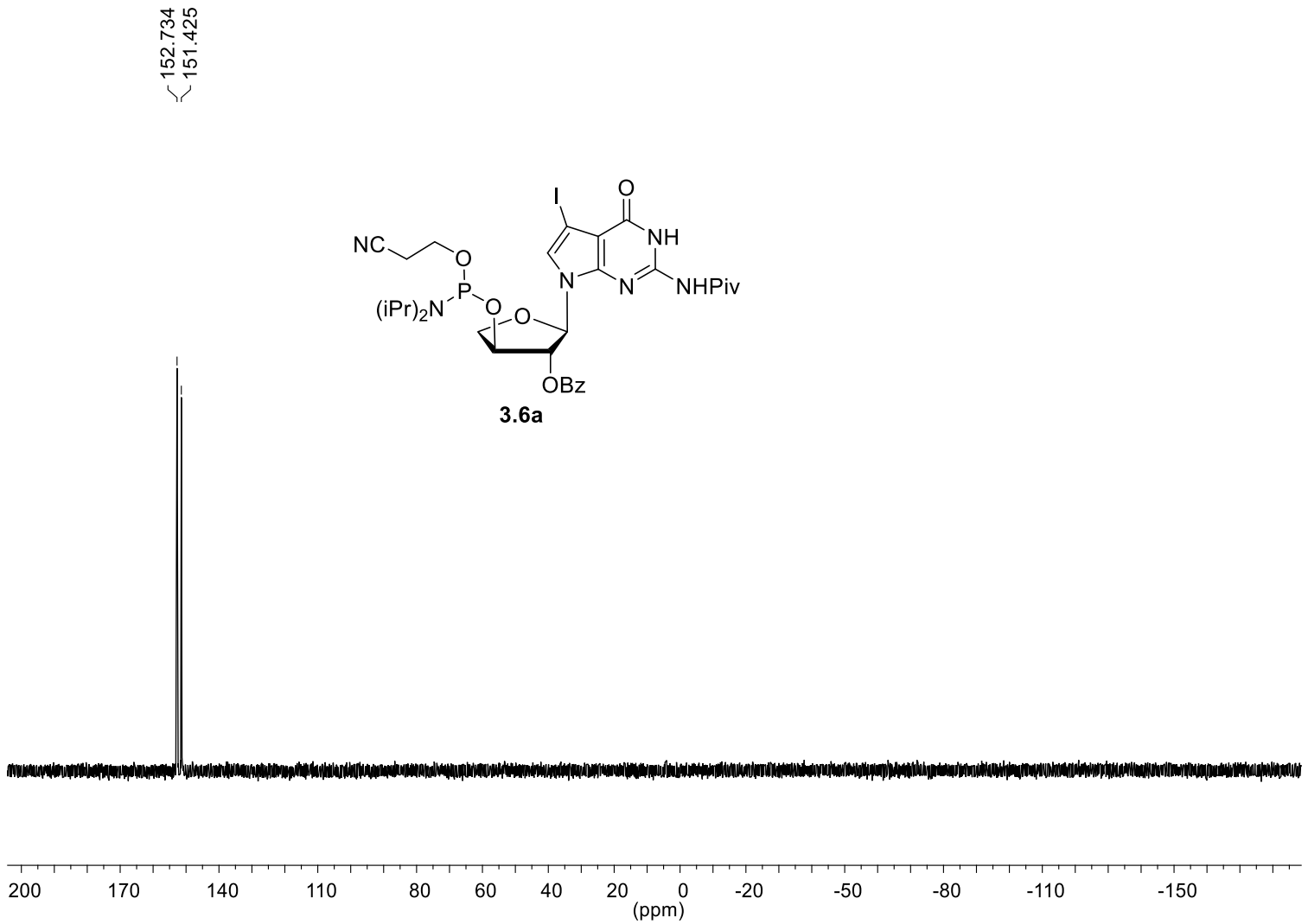
^{13}C NMR spectrum of compound **3.5b** (125.8 MHz, $CDCl_3$).



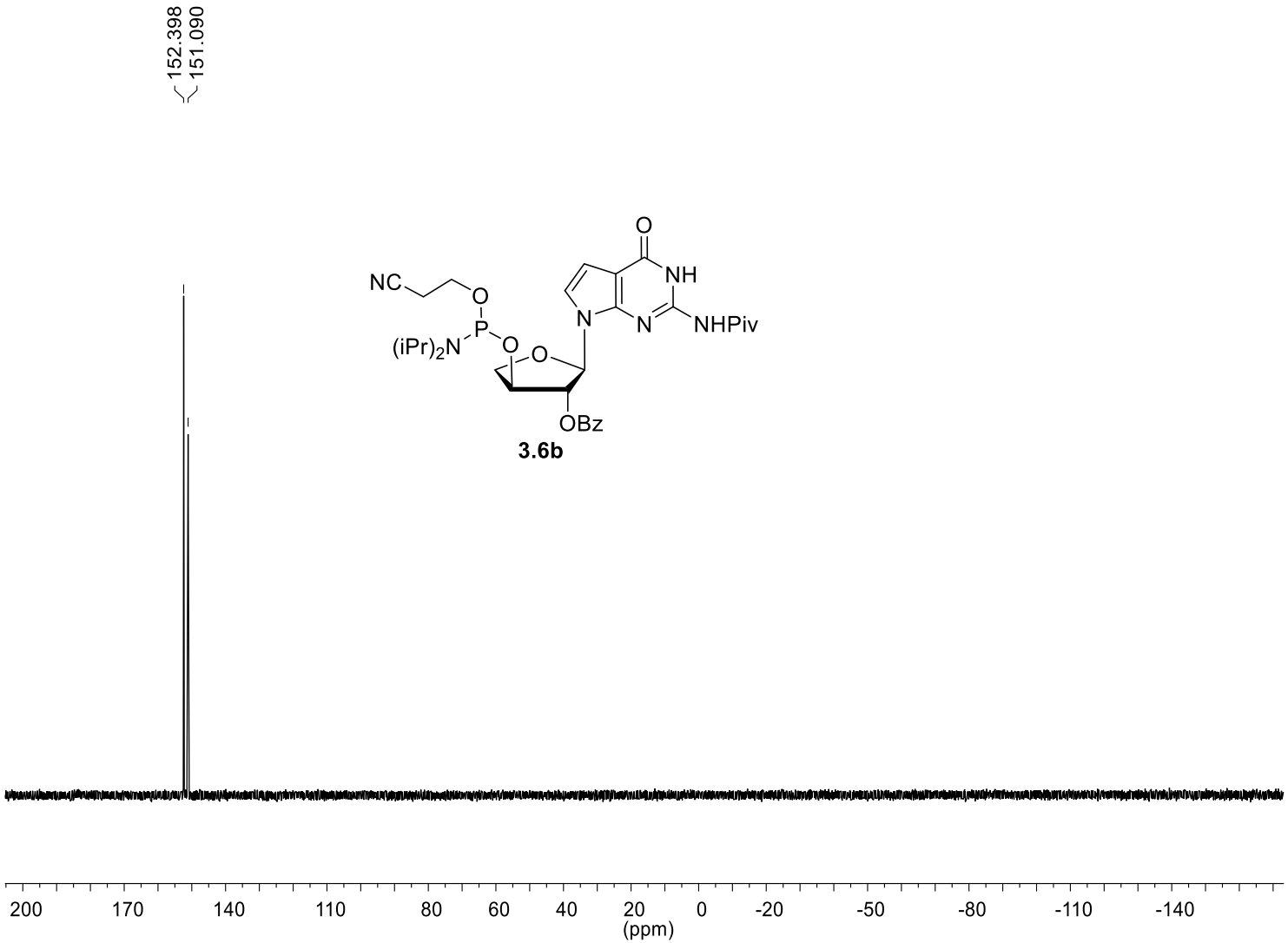
^{1}H NMR spectrum of compound **3.5c** (400 MHz, $CDCl_3$)



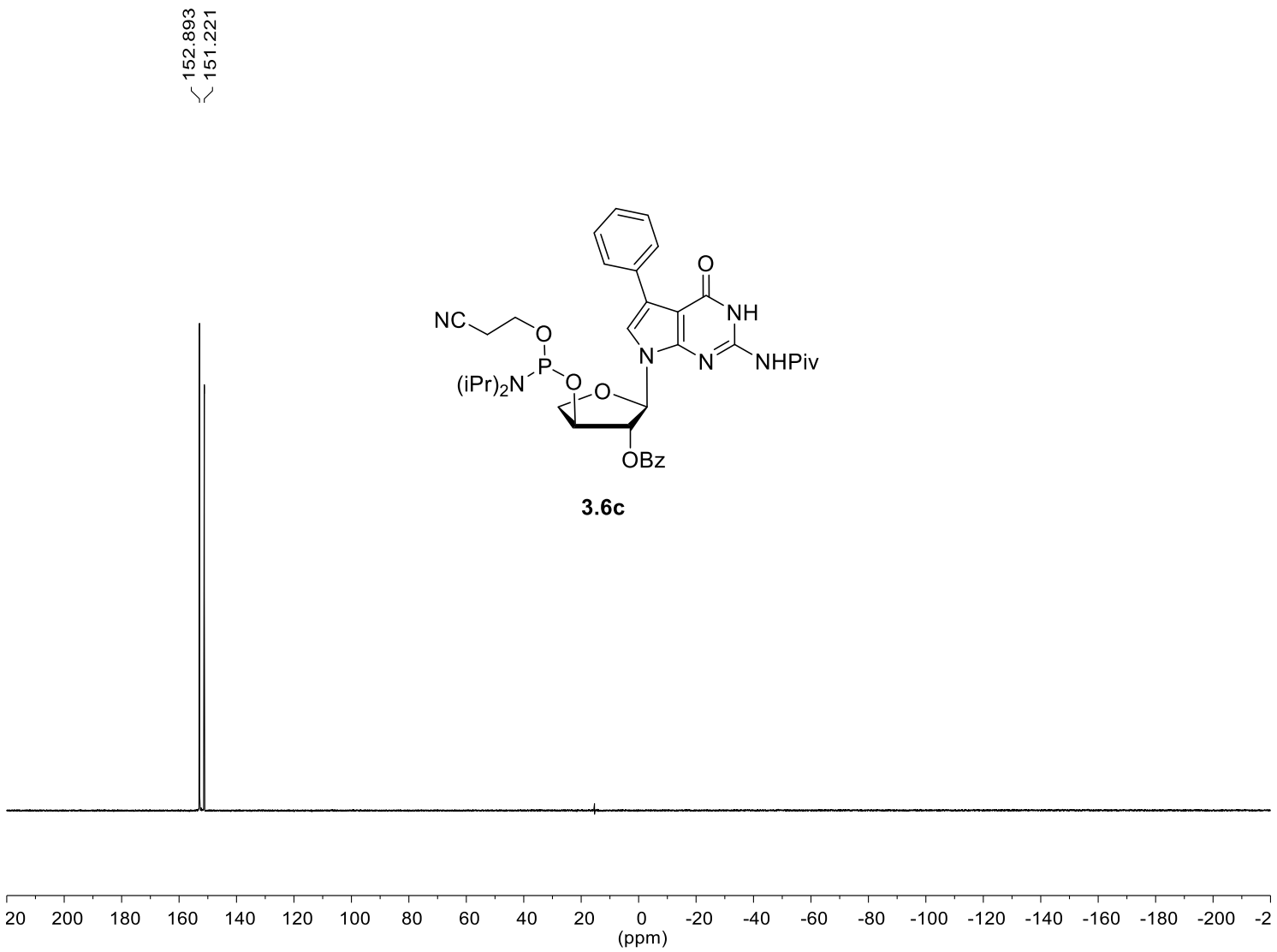
^{13}C NMR spectrum of compound **3.5c** (125.8 MHz, $CDCl_3$).



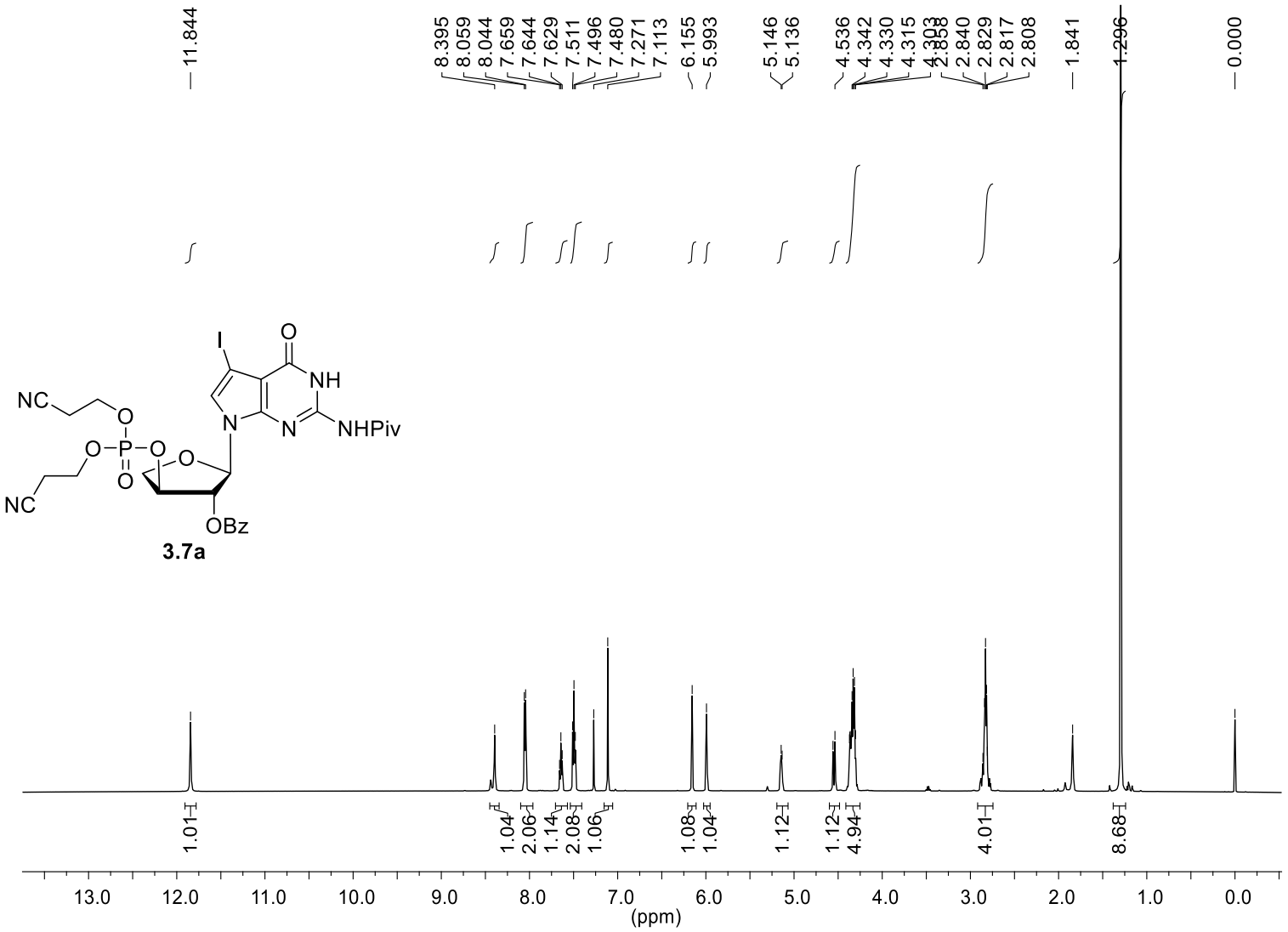
^{31}P NMR spectrum of compound **3.6a** (162 MHz, $CDCl_3$).



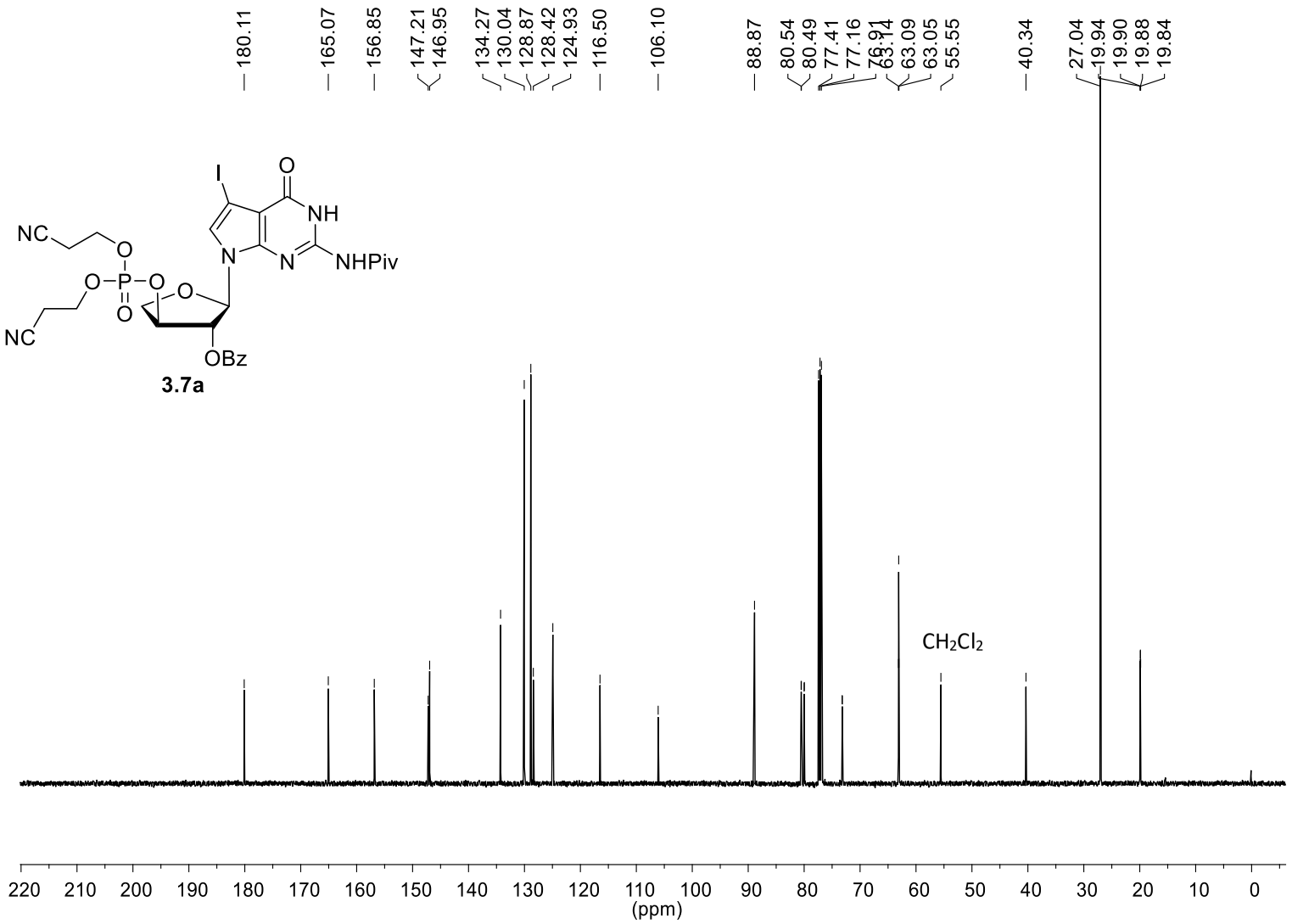
^{31}P NMR spectrum of compound **3.6b** (162 MHz, $CDCl_3$).



^{31}P NMR spectrum of compound **3.6c** (162 MHz, $CDCl_3$).

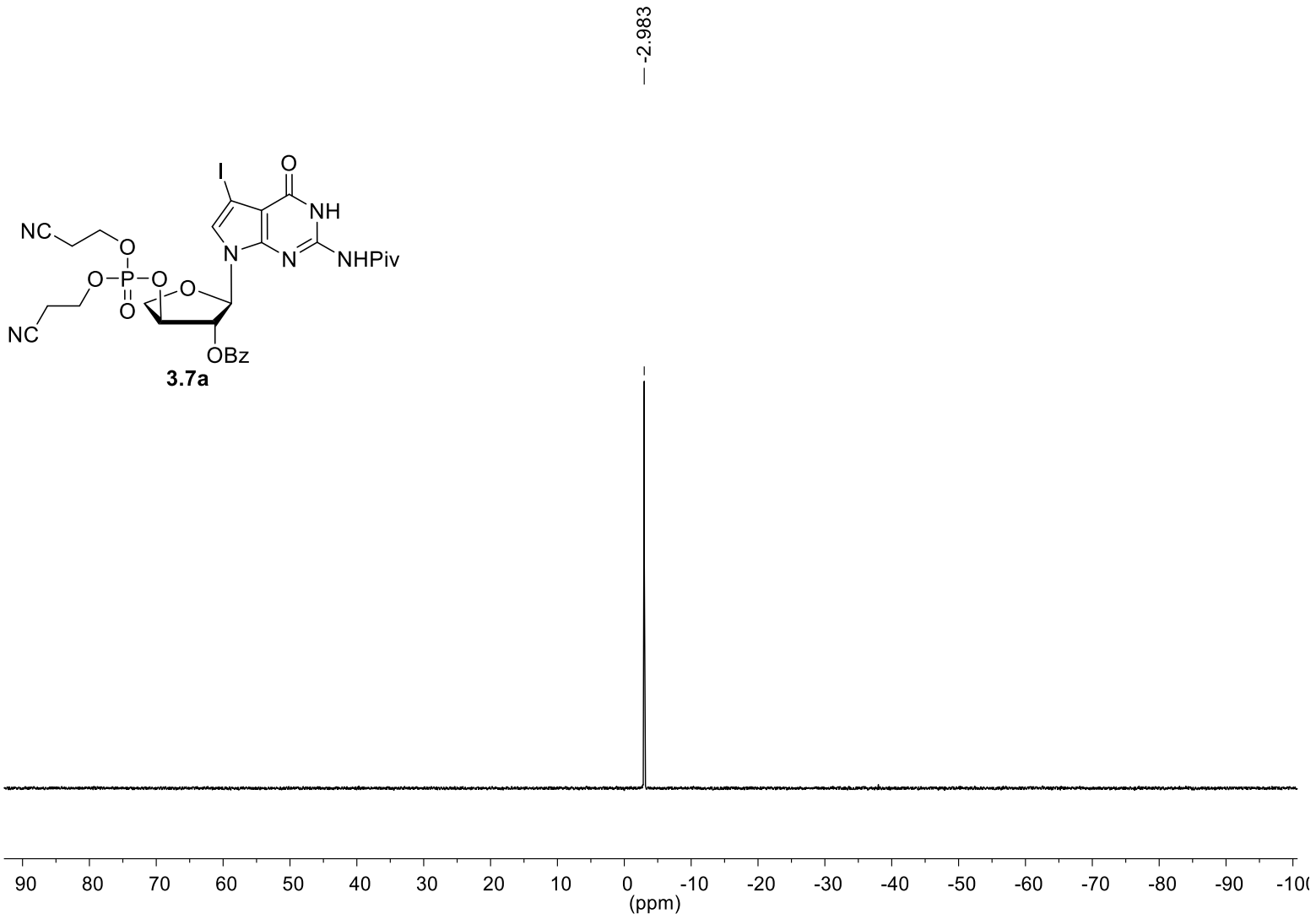


^{1}H NMR spectrum of compound **3.7a** (500 MHz, $CDCl_3$).

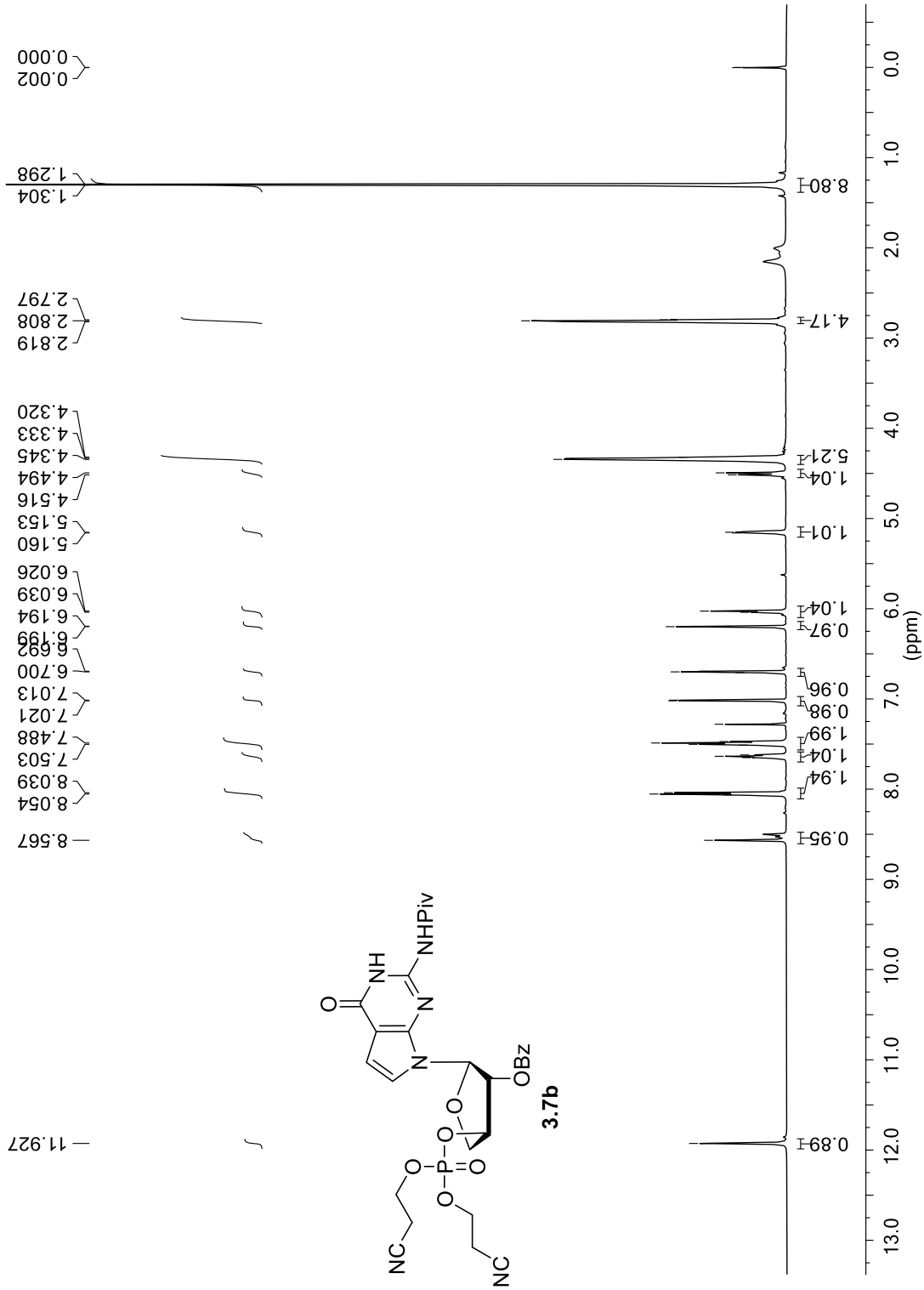


3.7a

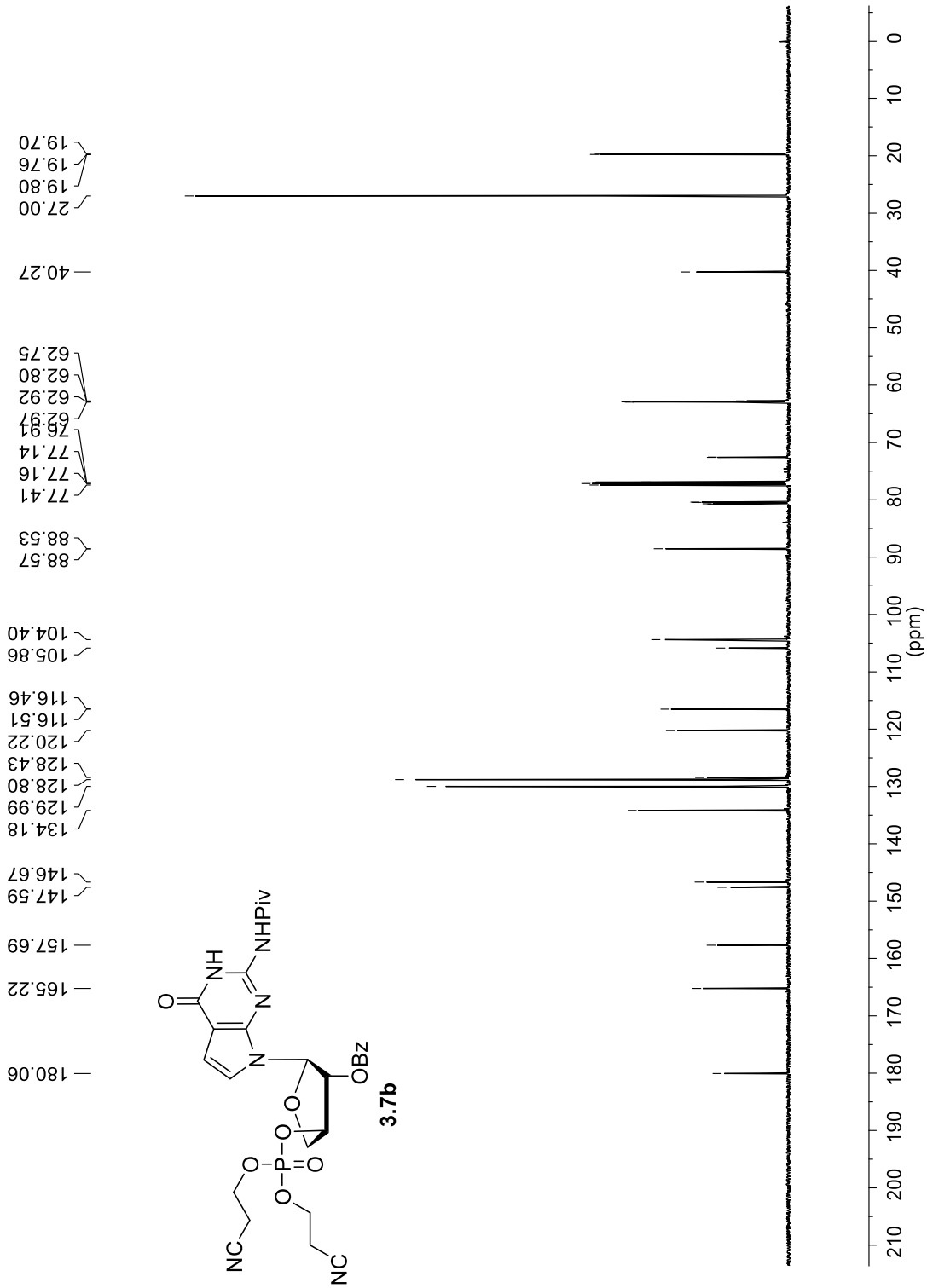
^{13}C NMR spectrum of compound **3.7a** (125.8 MHz, $CDCl_3$).



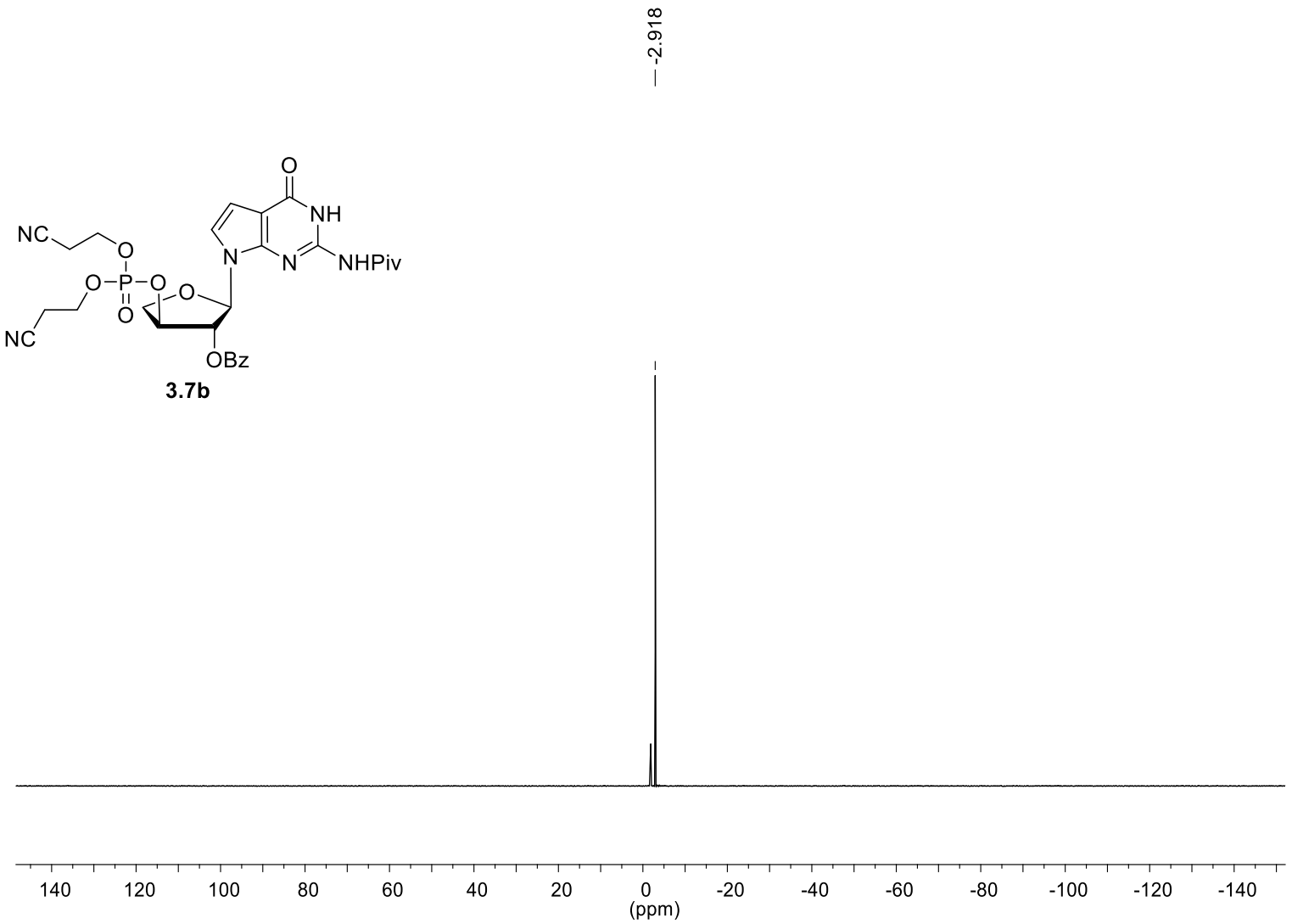
^{31}P NMR spectrum of compound **3.7a** (162 MHz, $CDCl_3$).



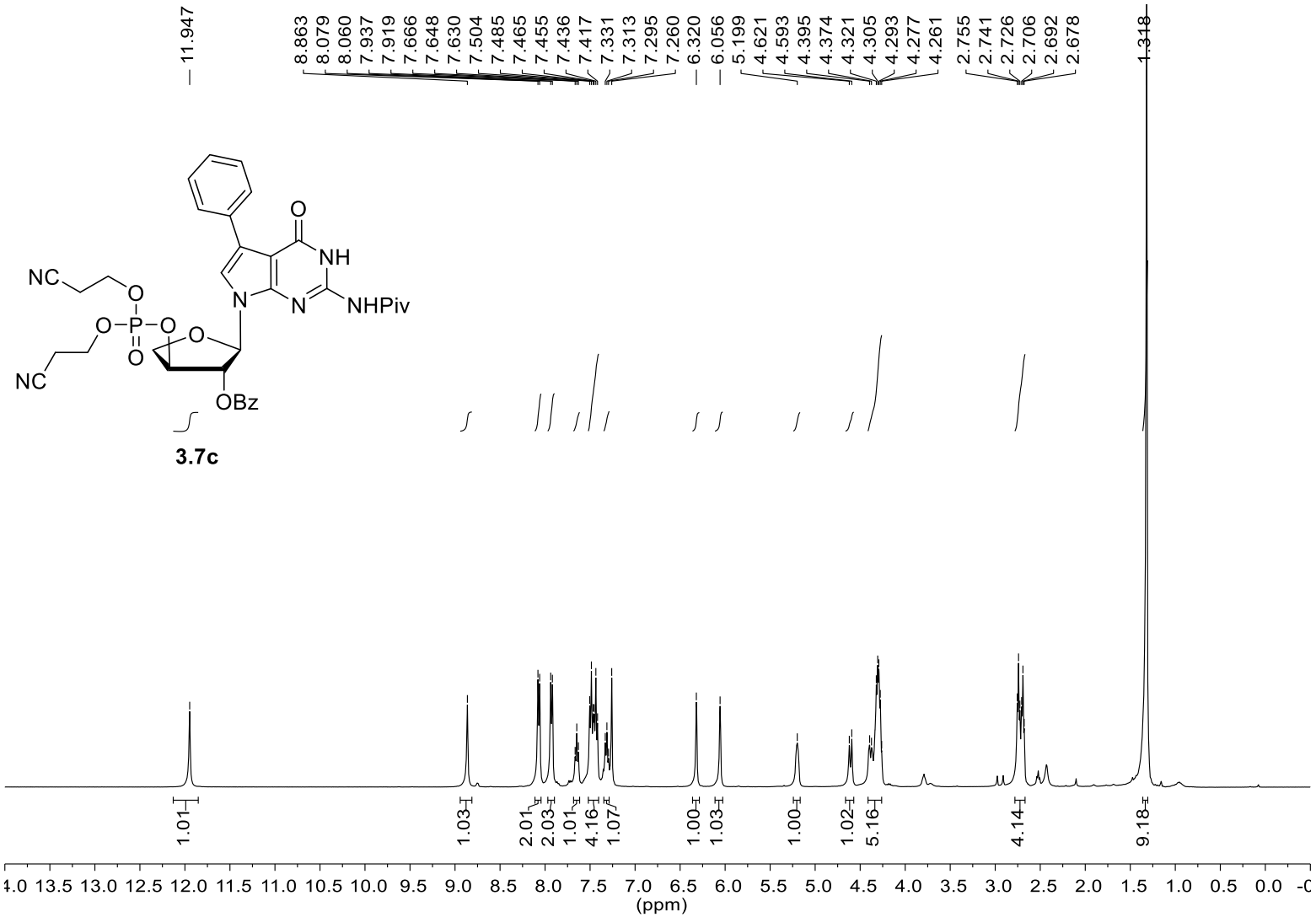
^{1}H NMR spectrum of compound **3.7b** (500 MHz, $CDCl_3$).



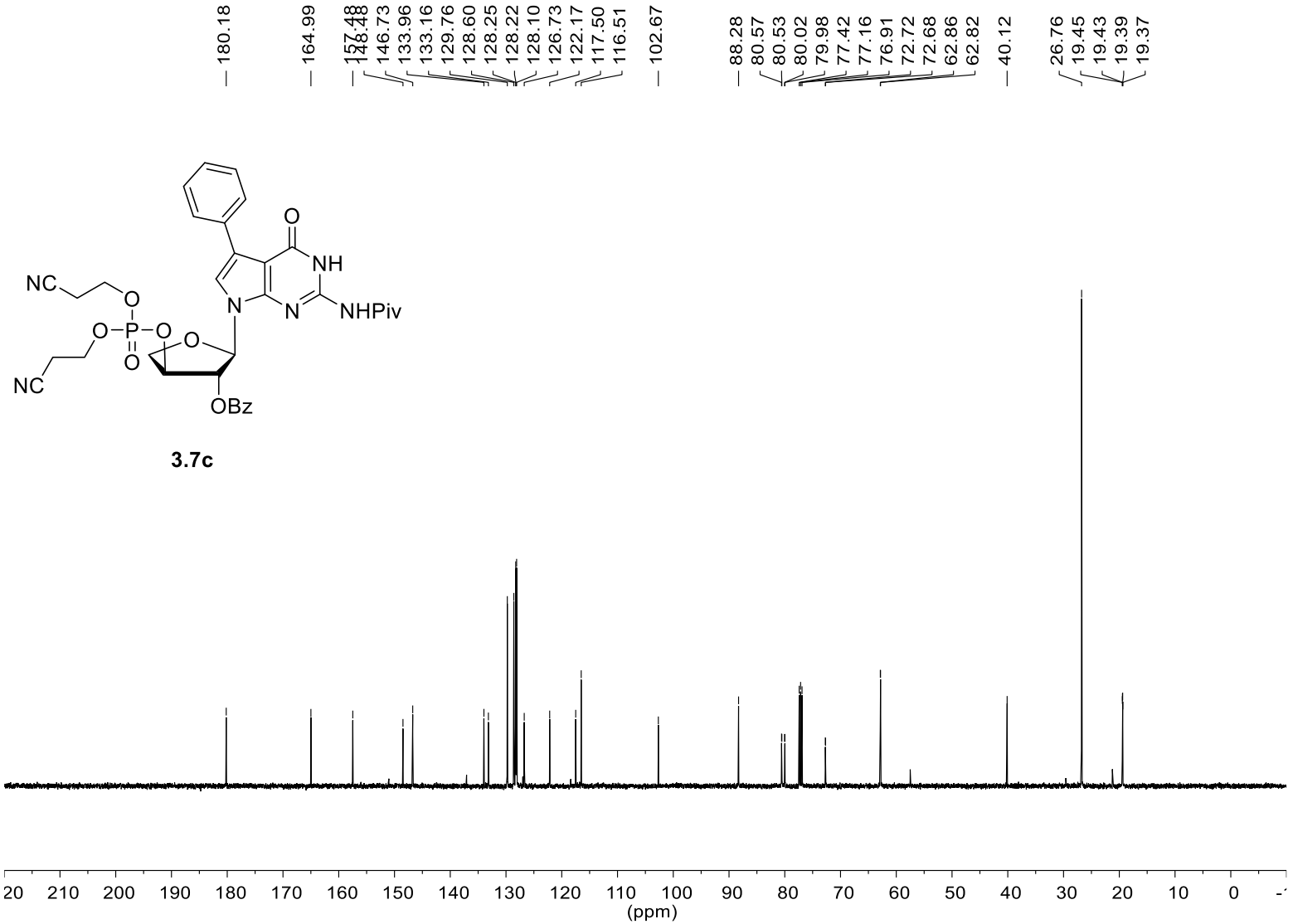
^{13}C NMR spectrum of compound **3.7b** (125.8 MHz, $CDCl_3$).



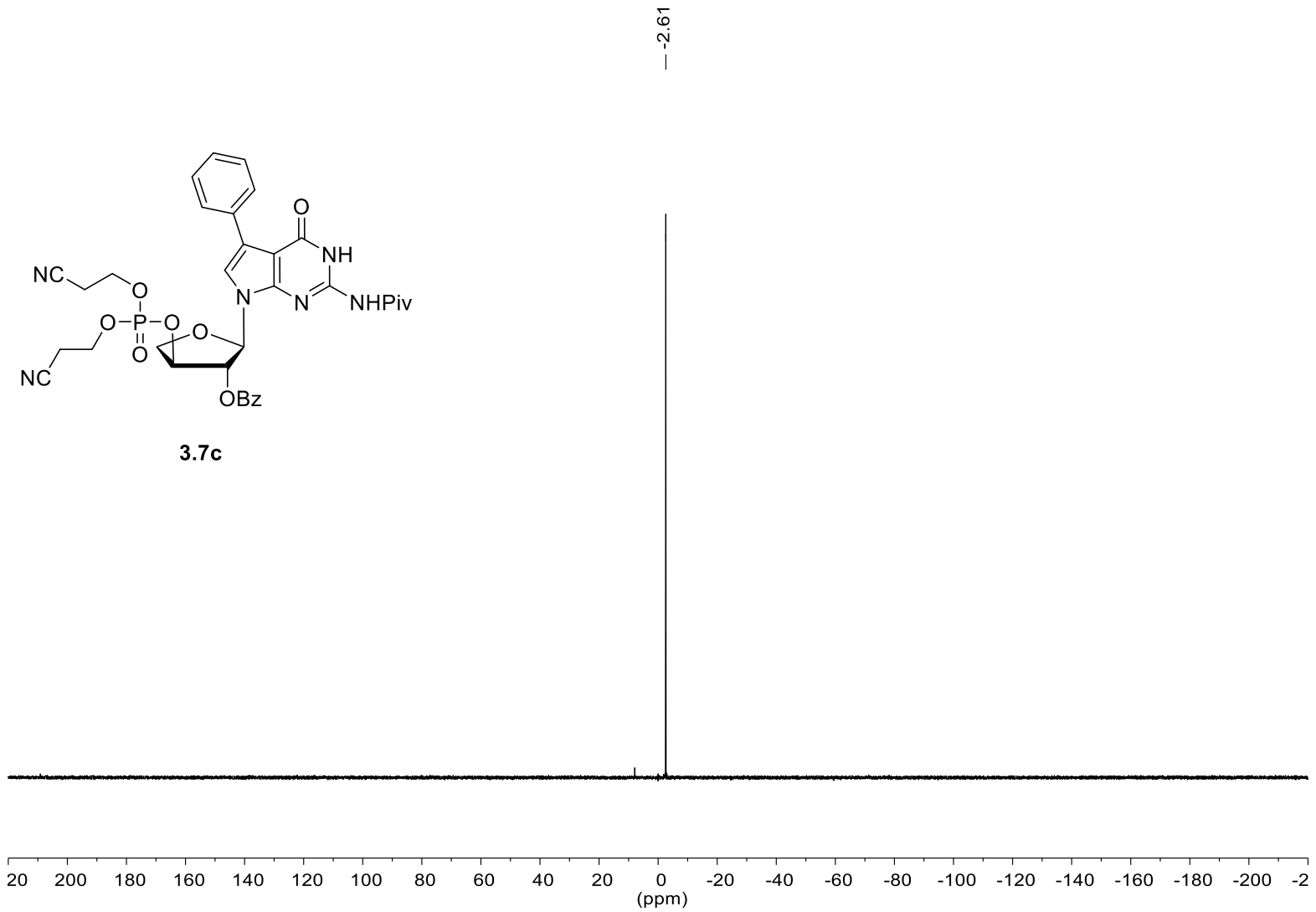
^{31}P NMR spectrum of compound **3.7b** (162 MHz, $CDCl_3$).



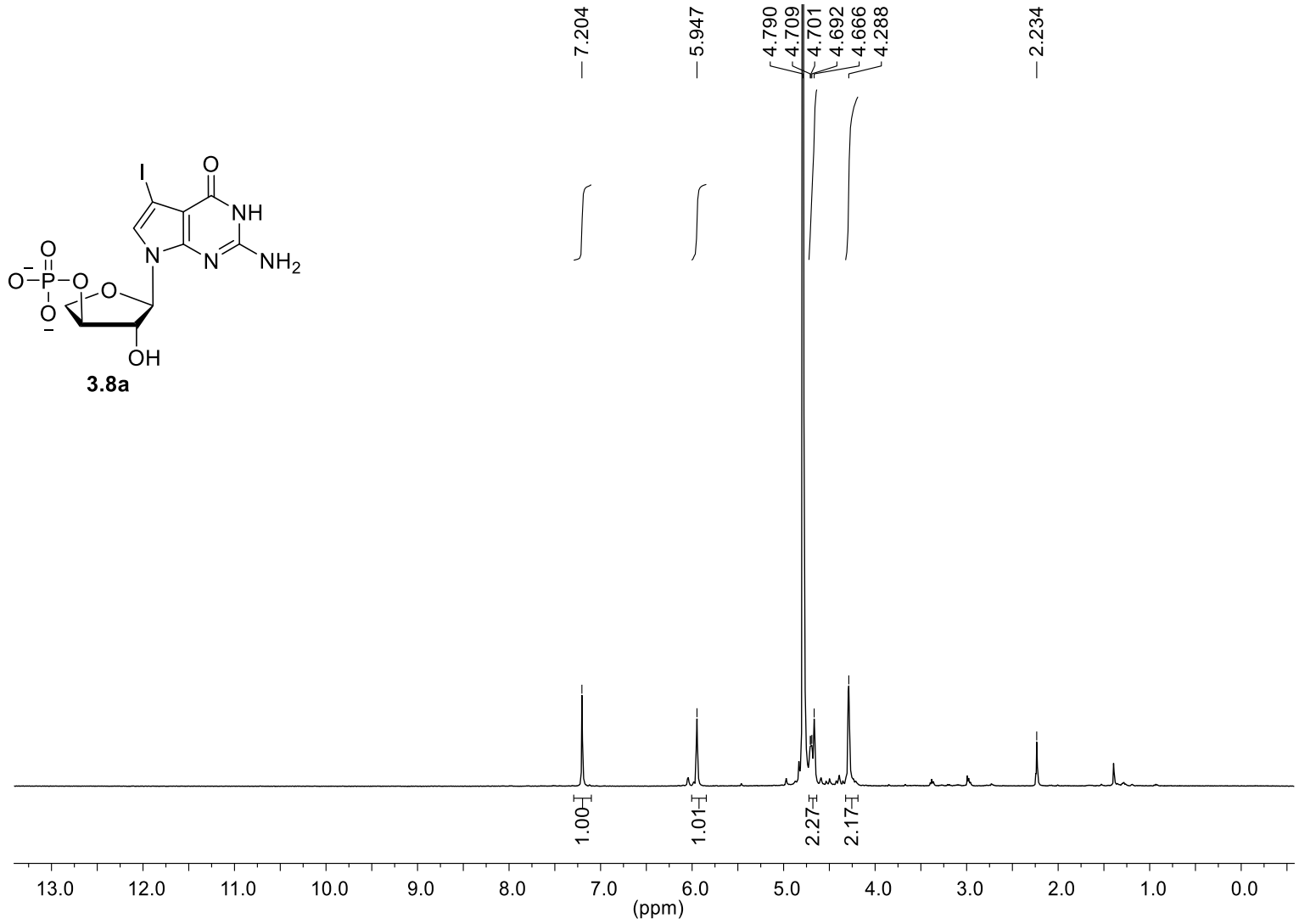
^{1}H NMR spectrum of compound **3.7c** (400 MHz, $CDCl_3$).



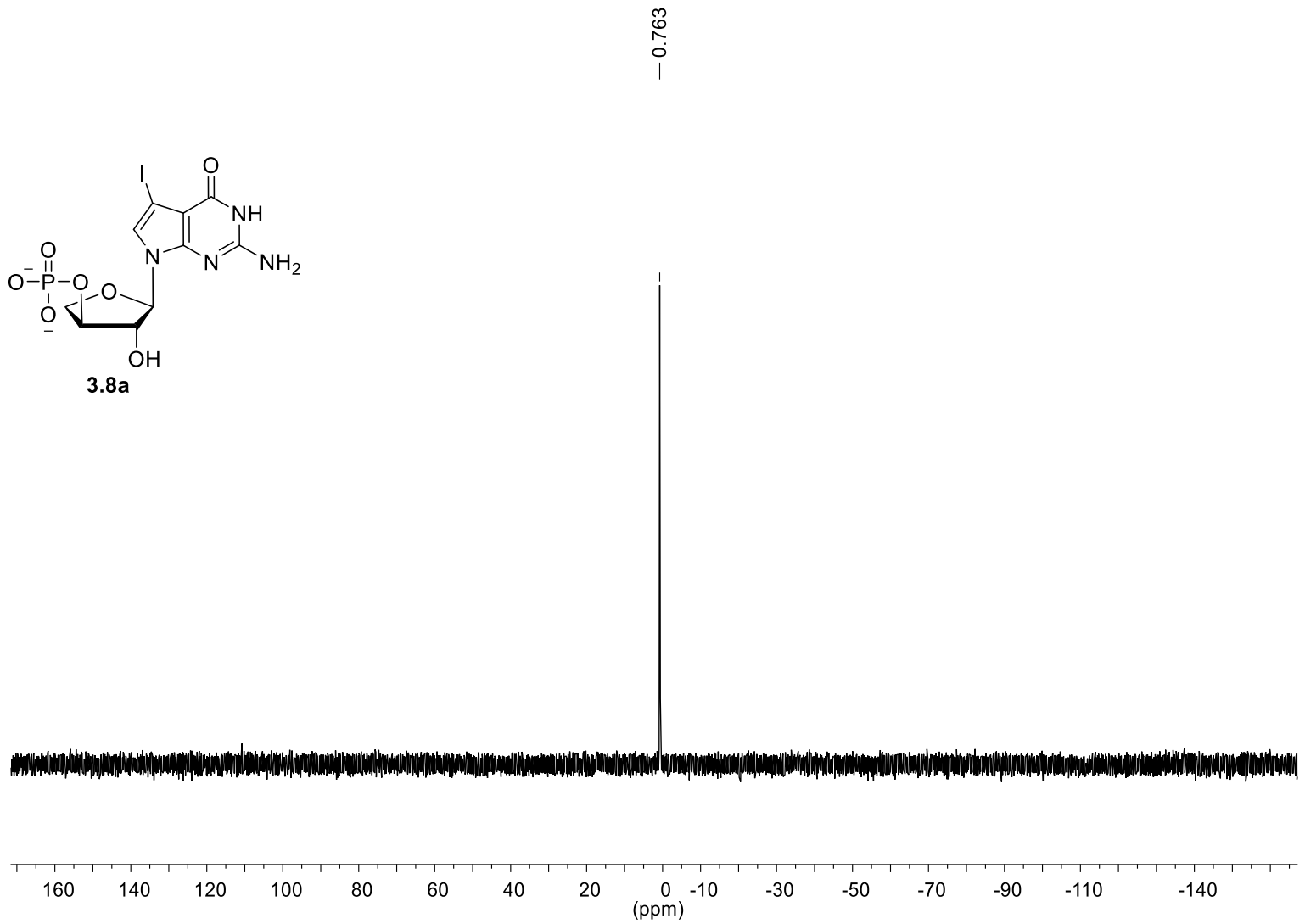
^{13}C NMR spectrum of compound **3.7c** (125.8 MHz, $CDCl_3$).



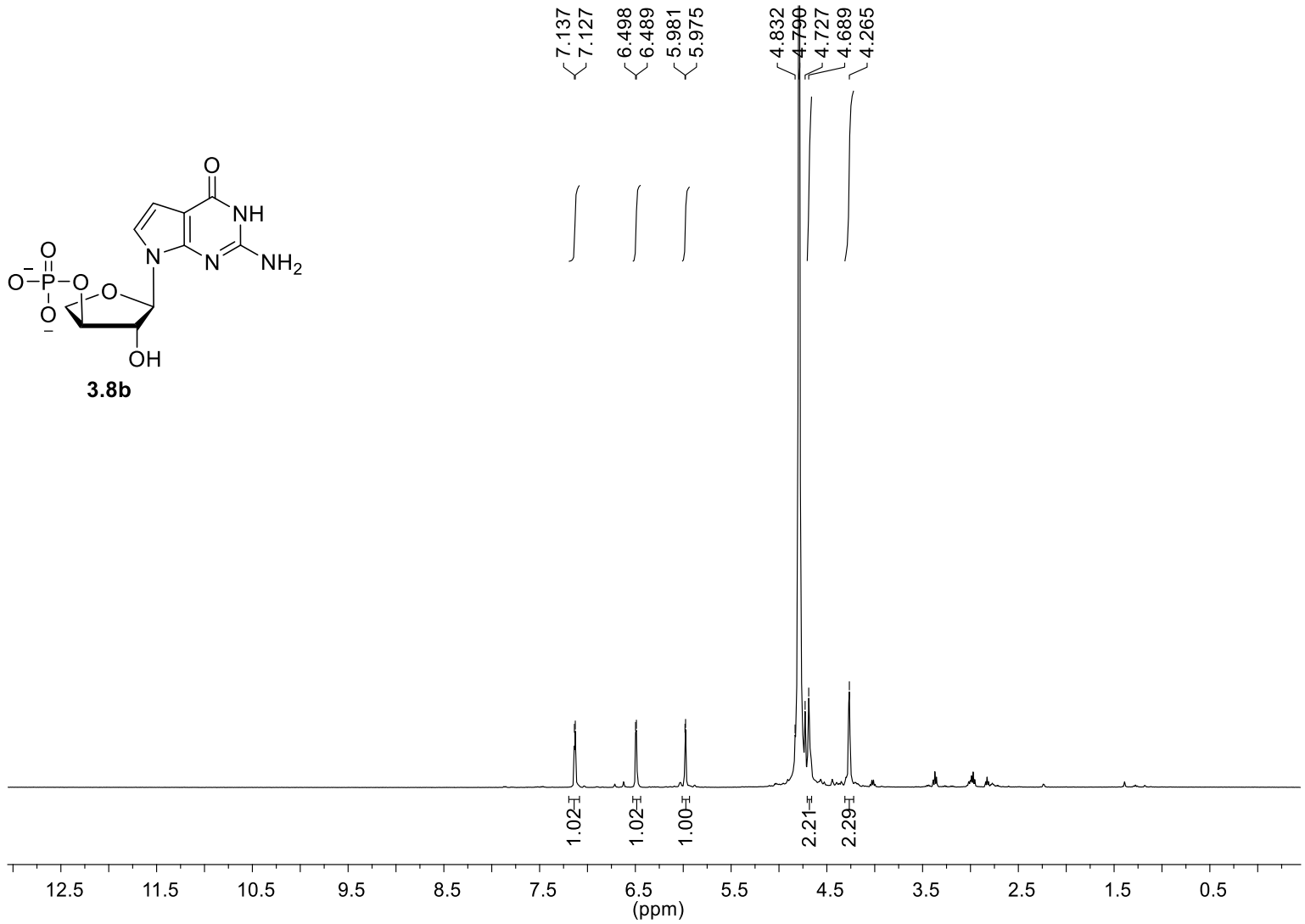
^{31}P NMR spectrum of compound **3.7c** (162 MHz, $CDCl_3$).



^{1}H NMR spectrum of compound **3.8a** (400 MHz, D_2O).

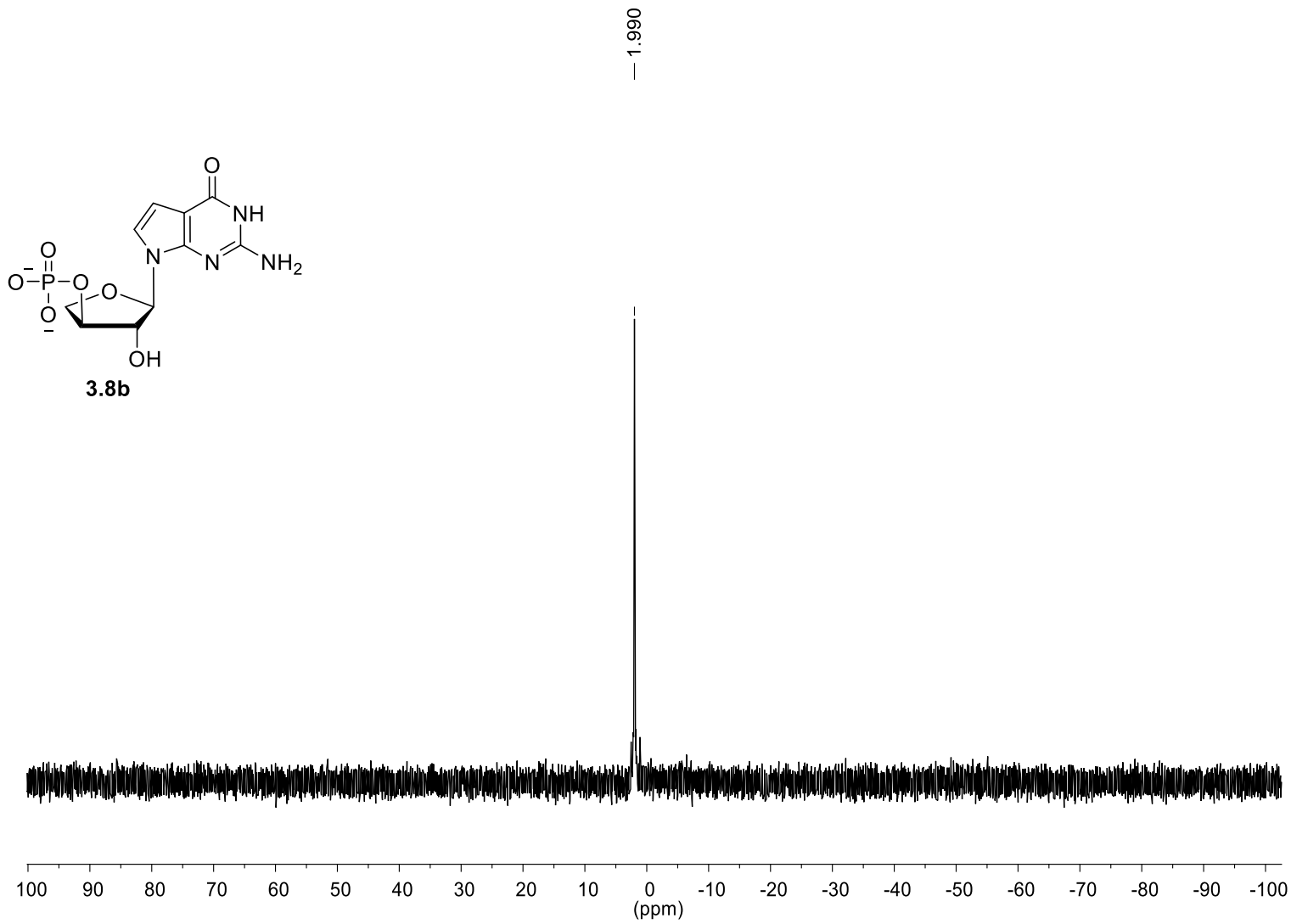


^{31}P NMR spectrum of compound **3.8a** (162 MHz, D_2O).

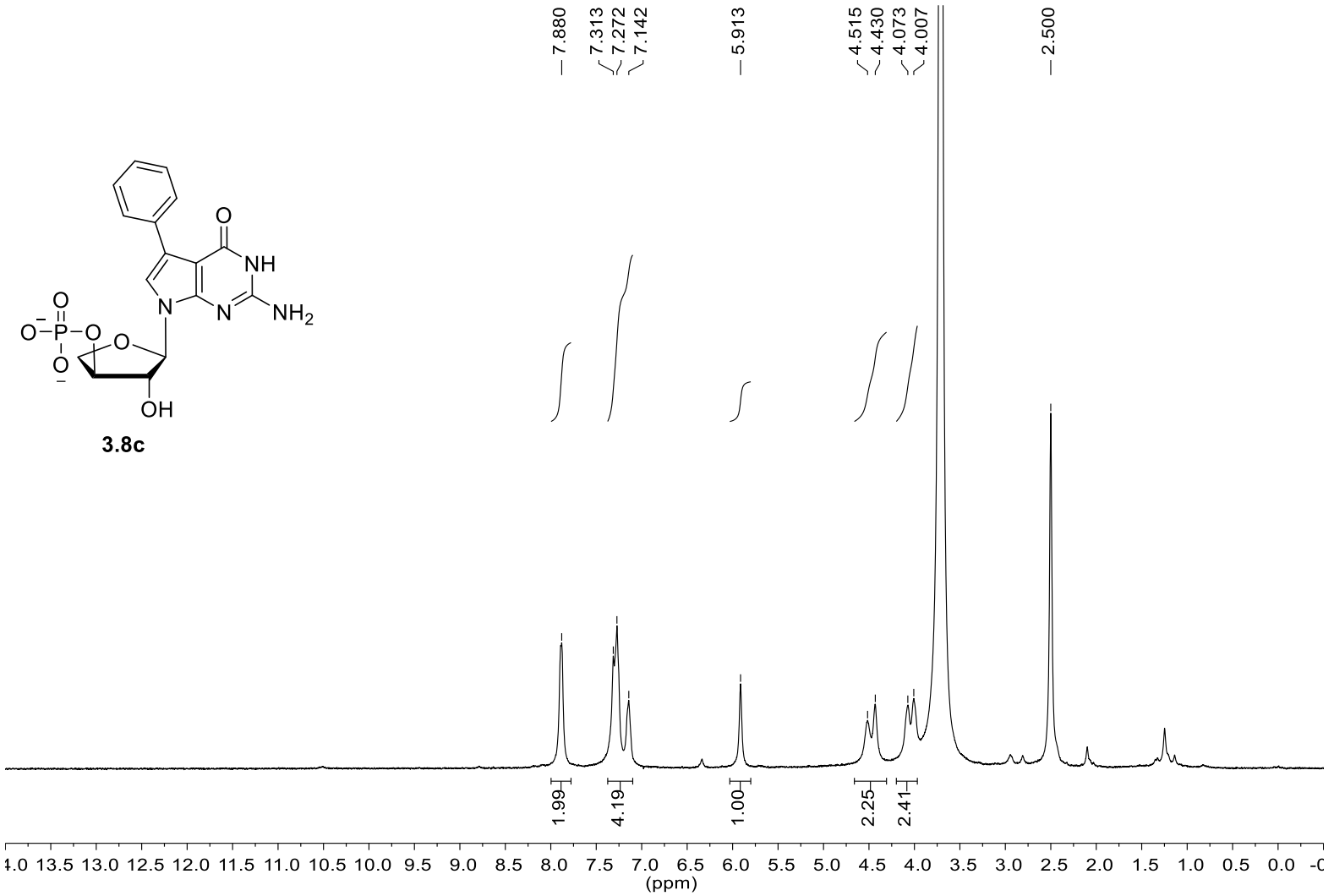


3.8b

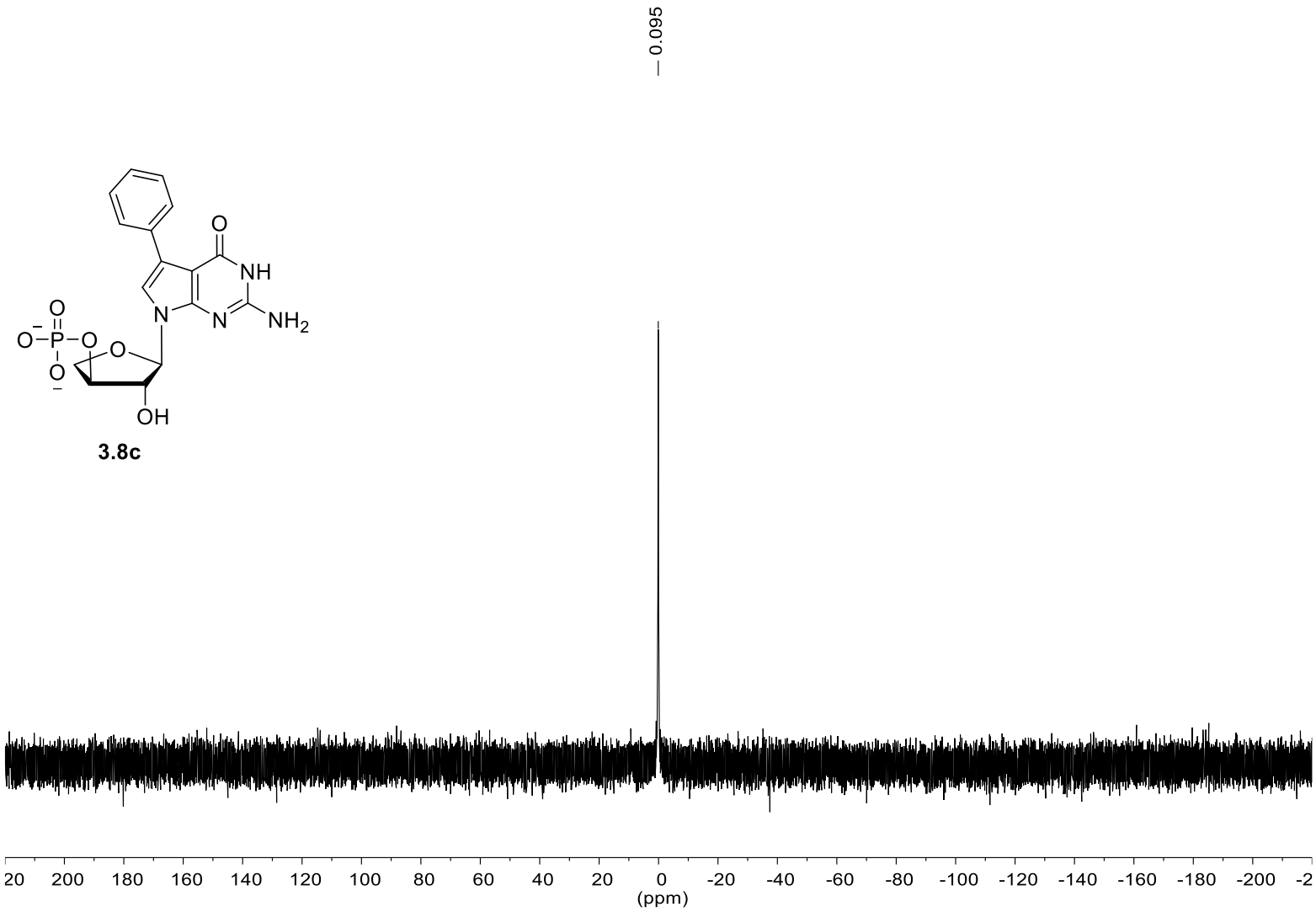
^{1}H NMR spectrum of compound **3.8b** (400 MHz, D_2O).



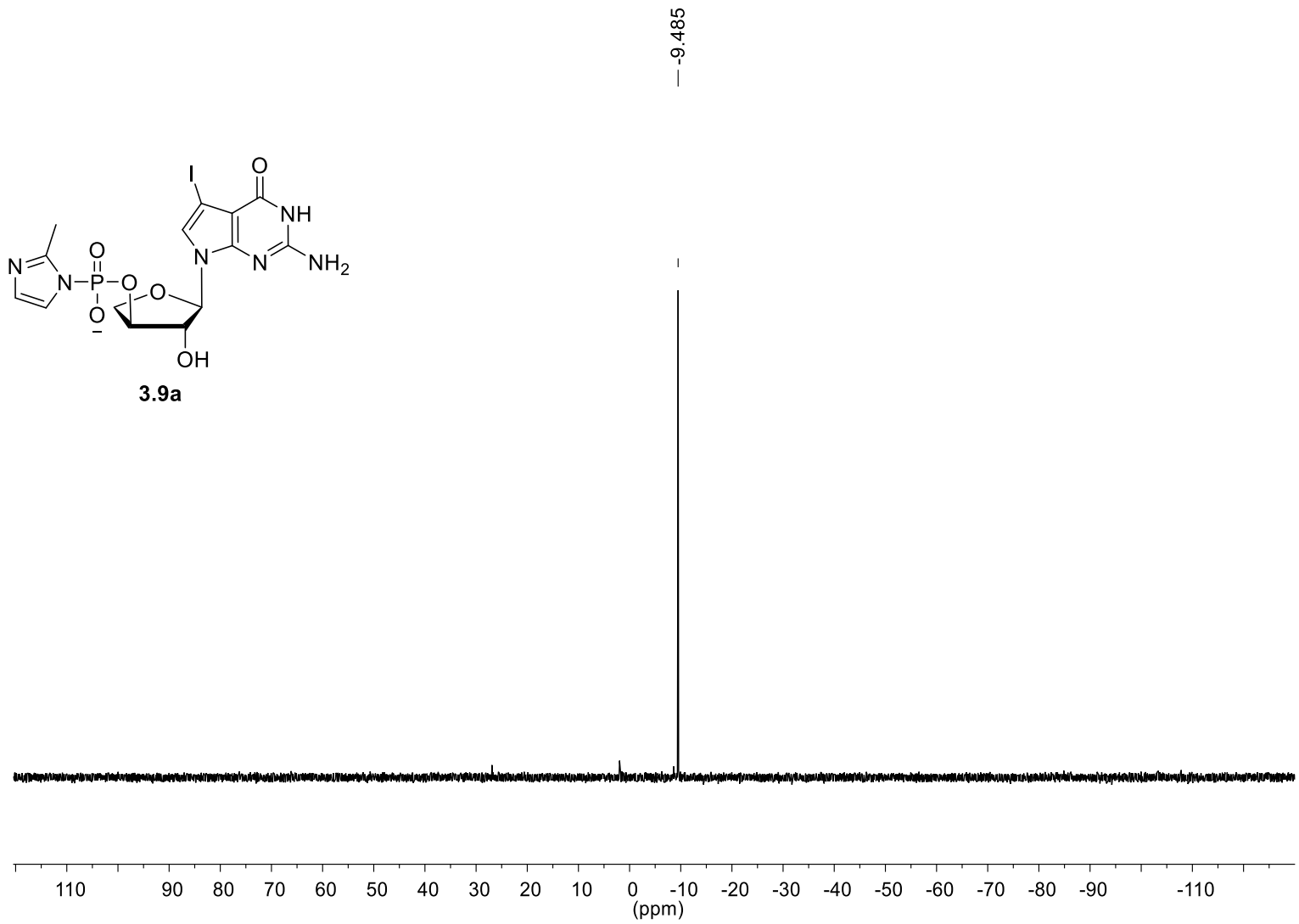
^{31}P NMR spectrum of compound **3.8b** (162 MHz, D_2O).



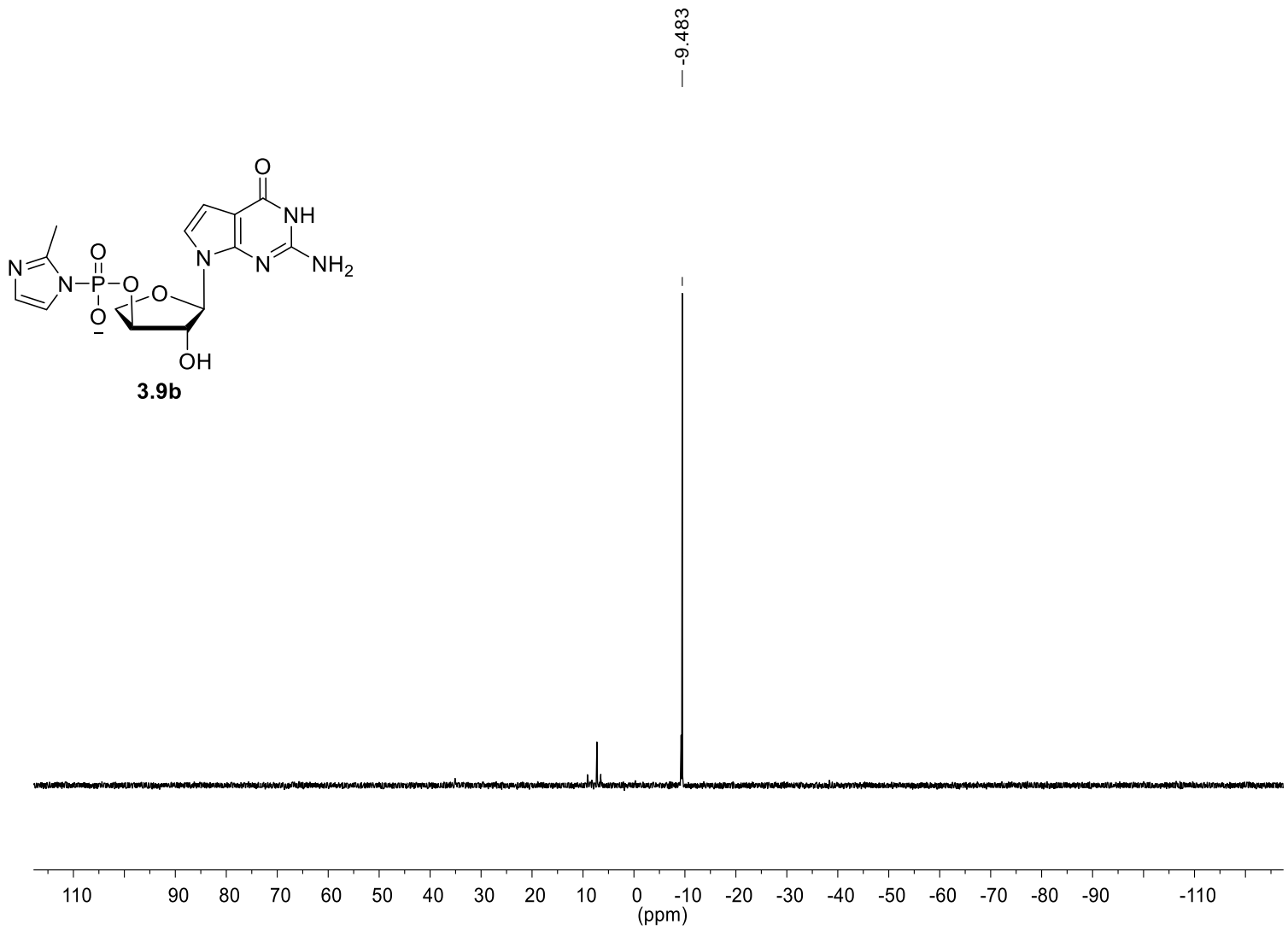
^{1}H NMR spectrum of compound **3.8c** (400 MHz, DMSO-d_6).



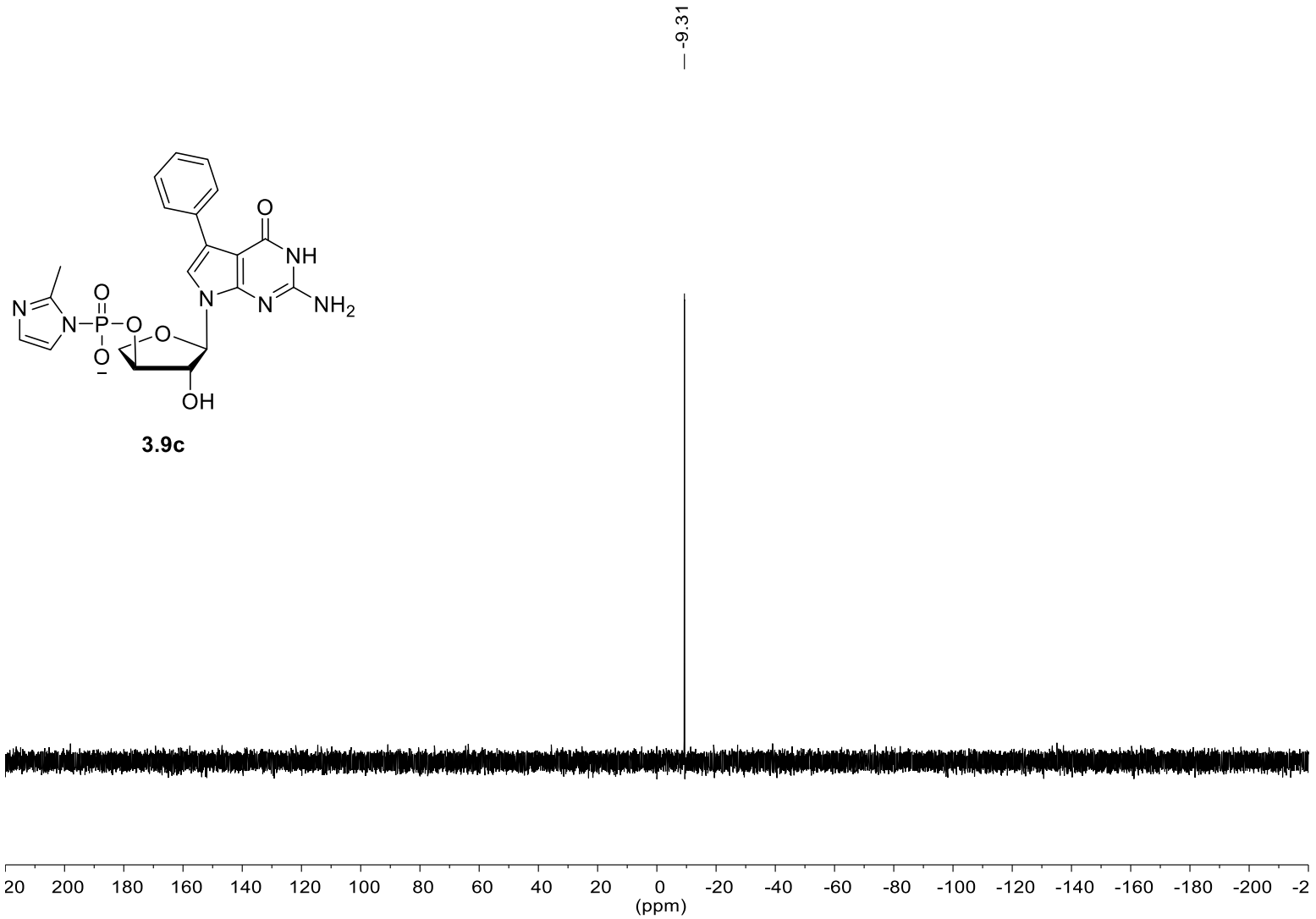
^{31}P NMR spectrum of compound **3.8c** (162 MHz, DMSO-d_6).



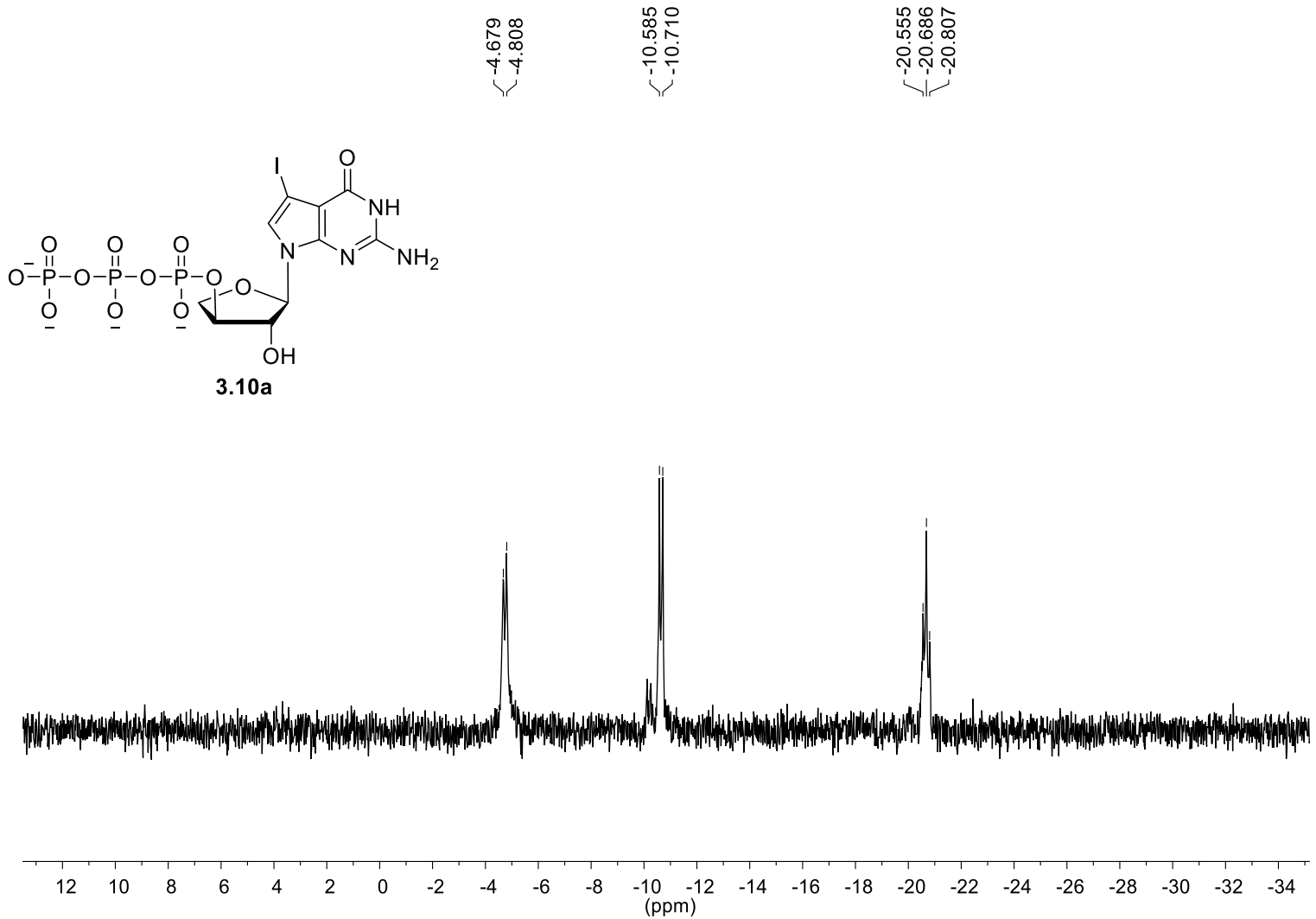
^{31}P NMR spectrum of compound **3.9a** (162 MHz, DMSO-d_6).



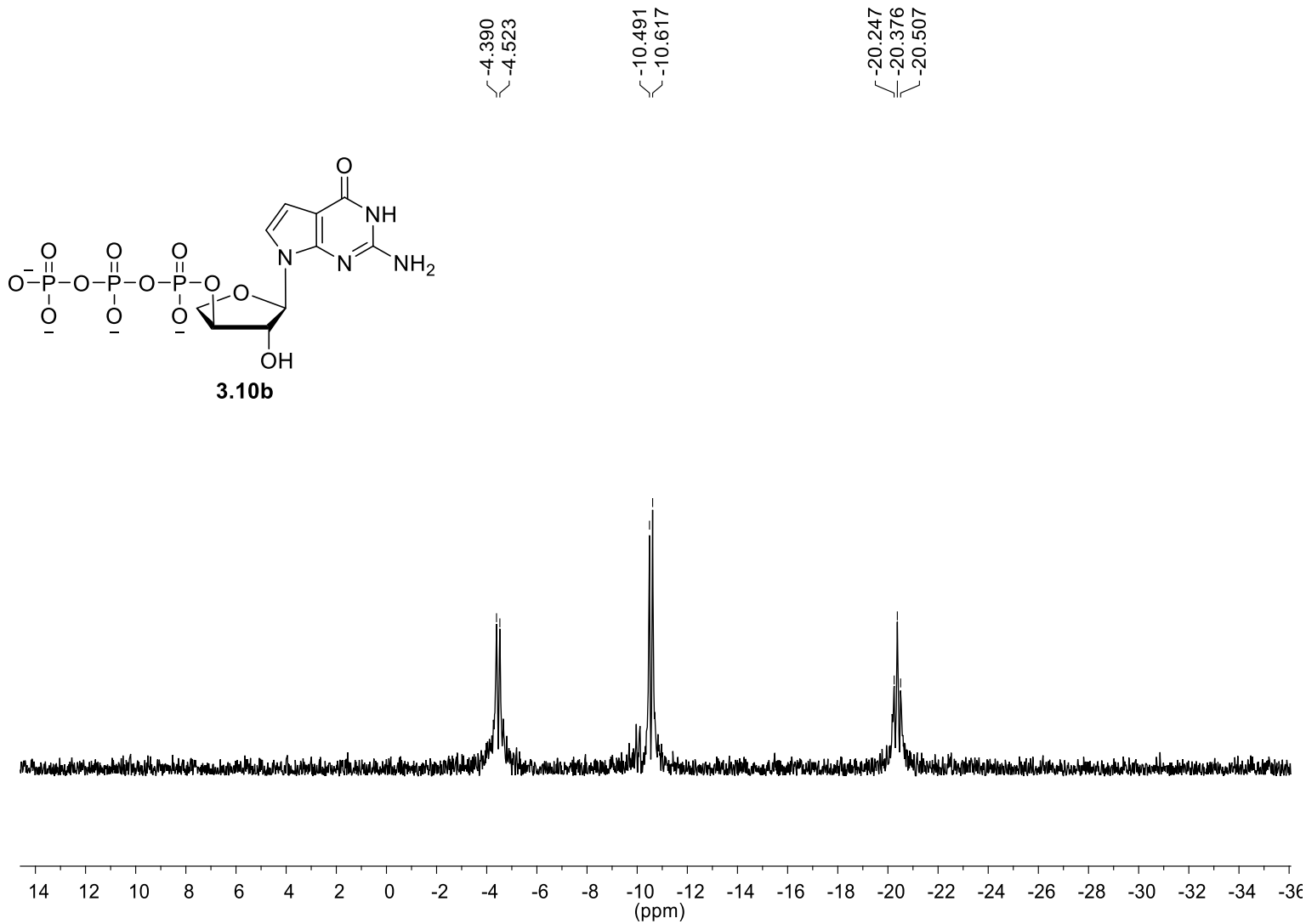
^{31}P NMR spectrum of compound **3.9b** (162 MHz, DMSO-d_6).



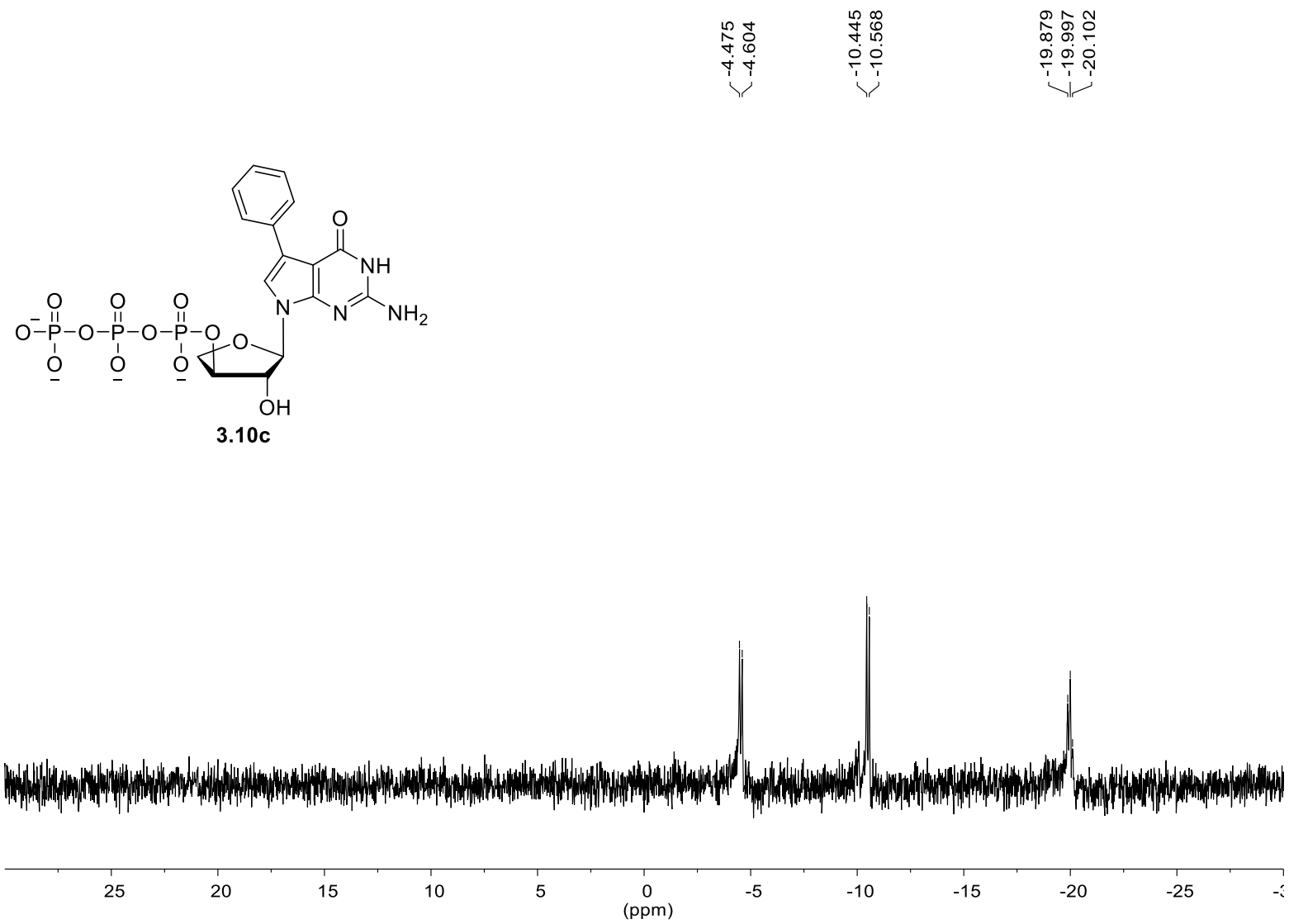
^{31}P NMR spectrum of compound **3.9c** (162 MHz, D_2O).



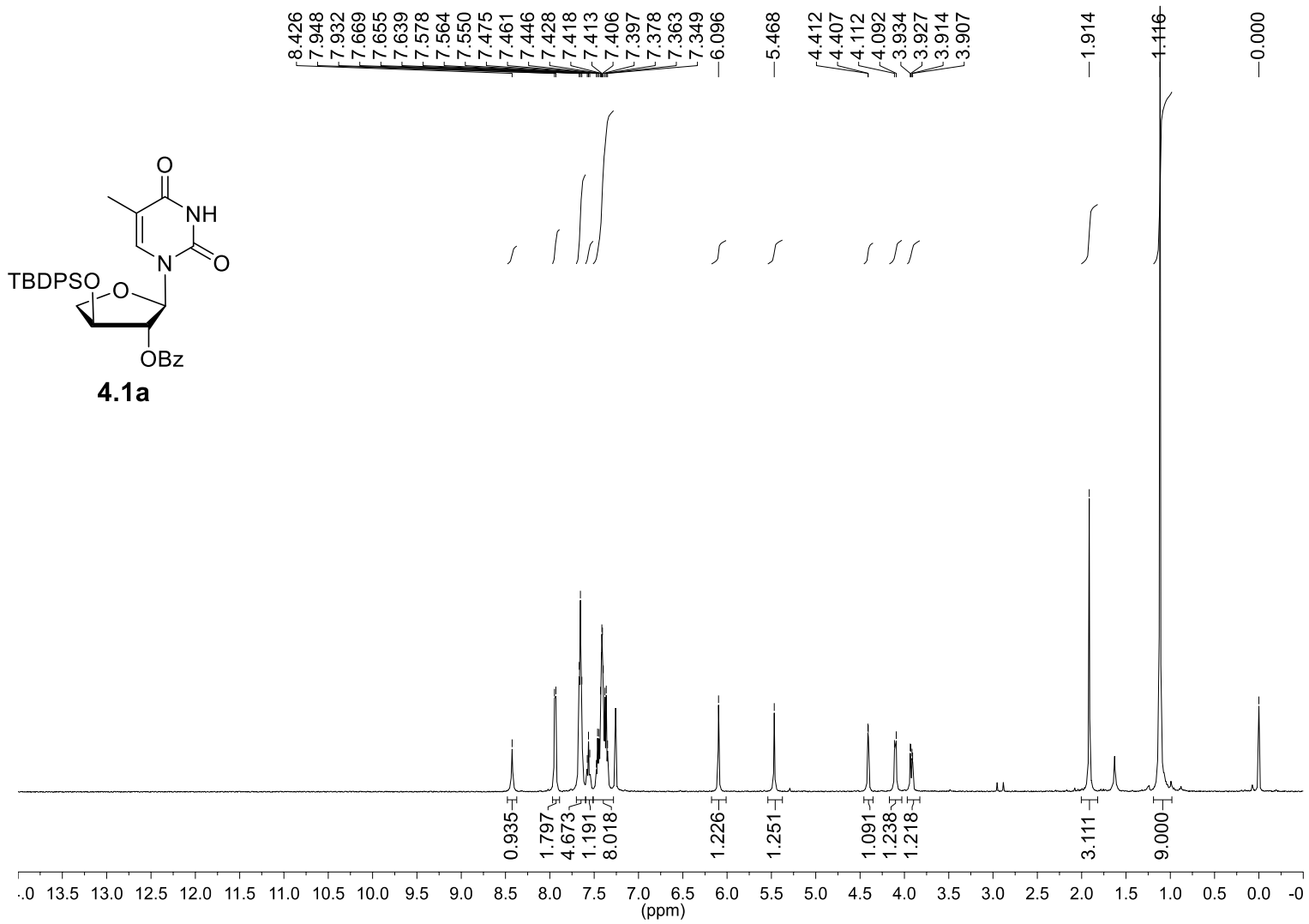
^{31}P NMR spectrum of compound **3.10a** (162 MHz, D_2O).



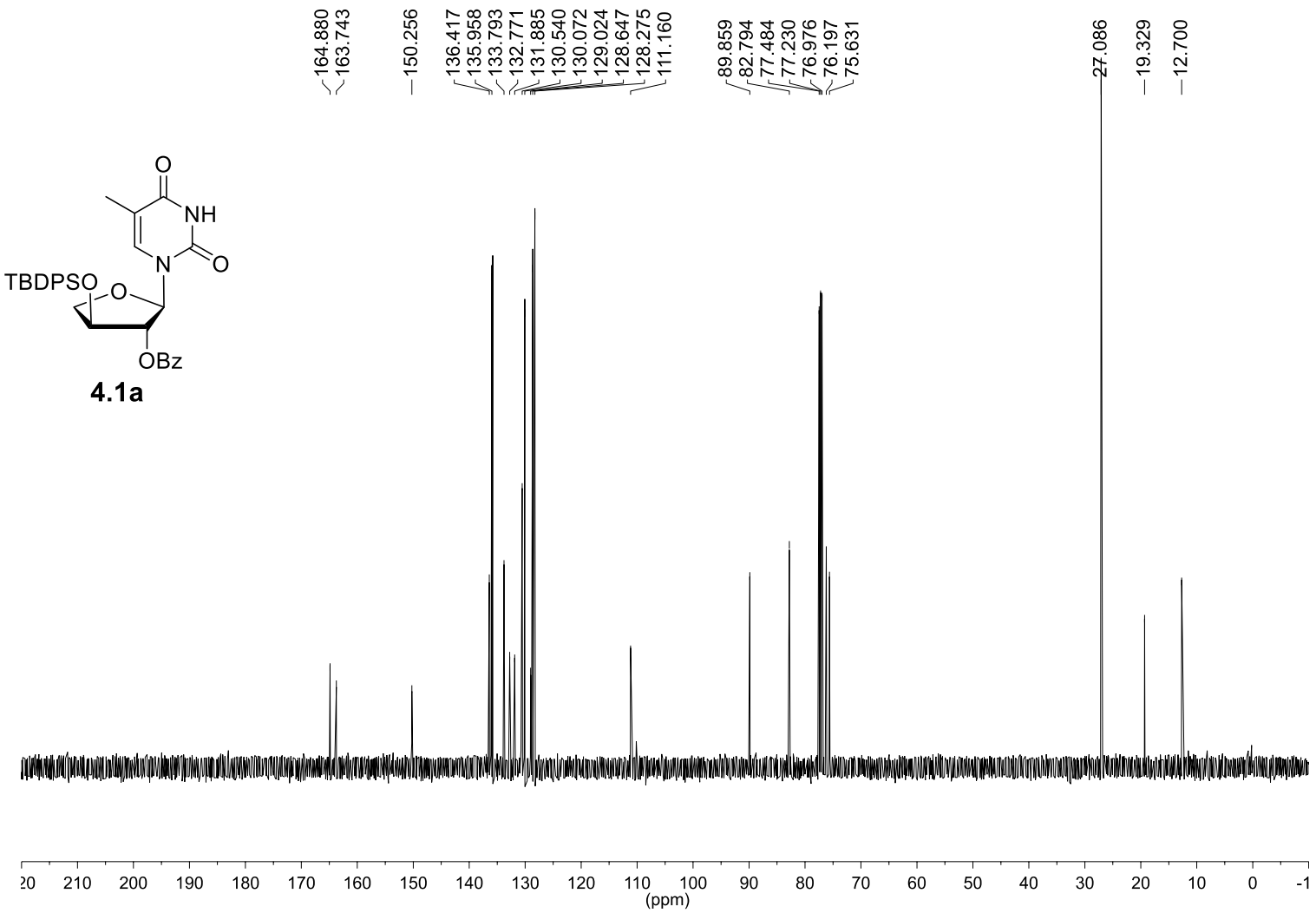
^{31}P NMR spectrum of compound **3.10b** (162 MHz, D_2O).



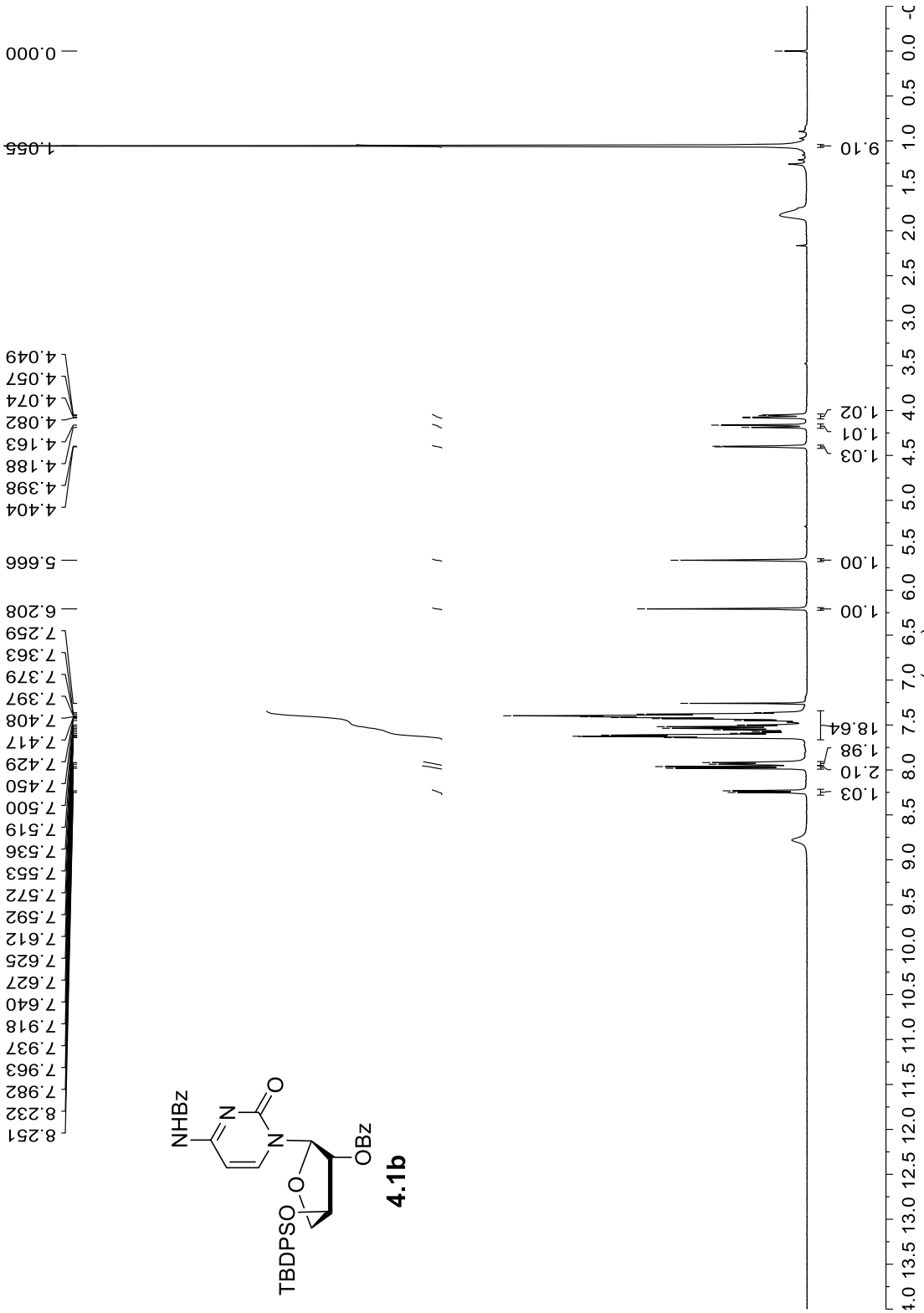
^{31}P NMR spectrum of compound **3.10c** (162 MHz, D_2O).



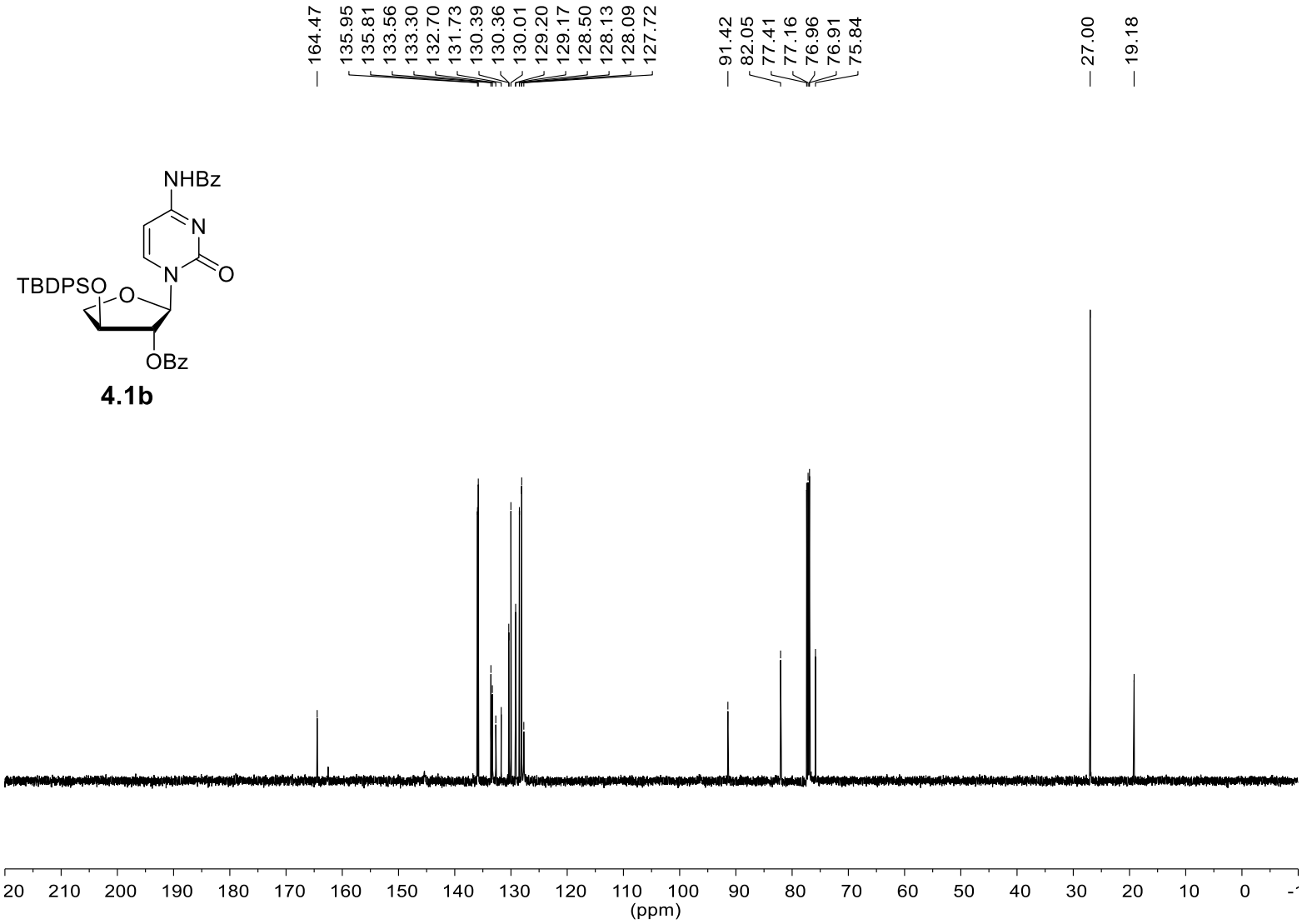
^{1}H NMR spectrum of compound **4.1a** (400 MHz, $CDCl_3$)



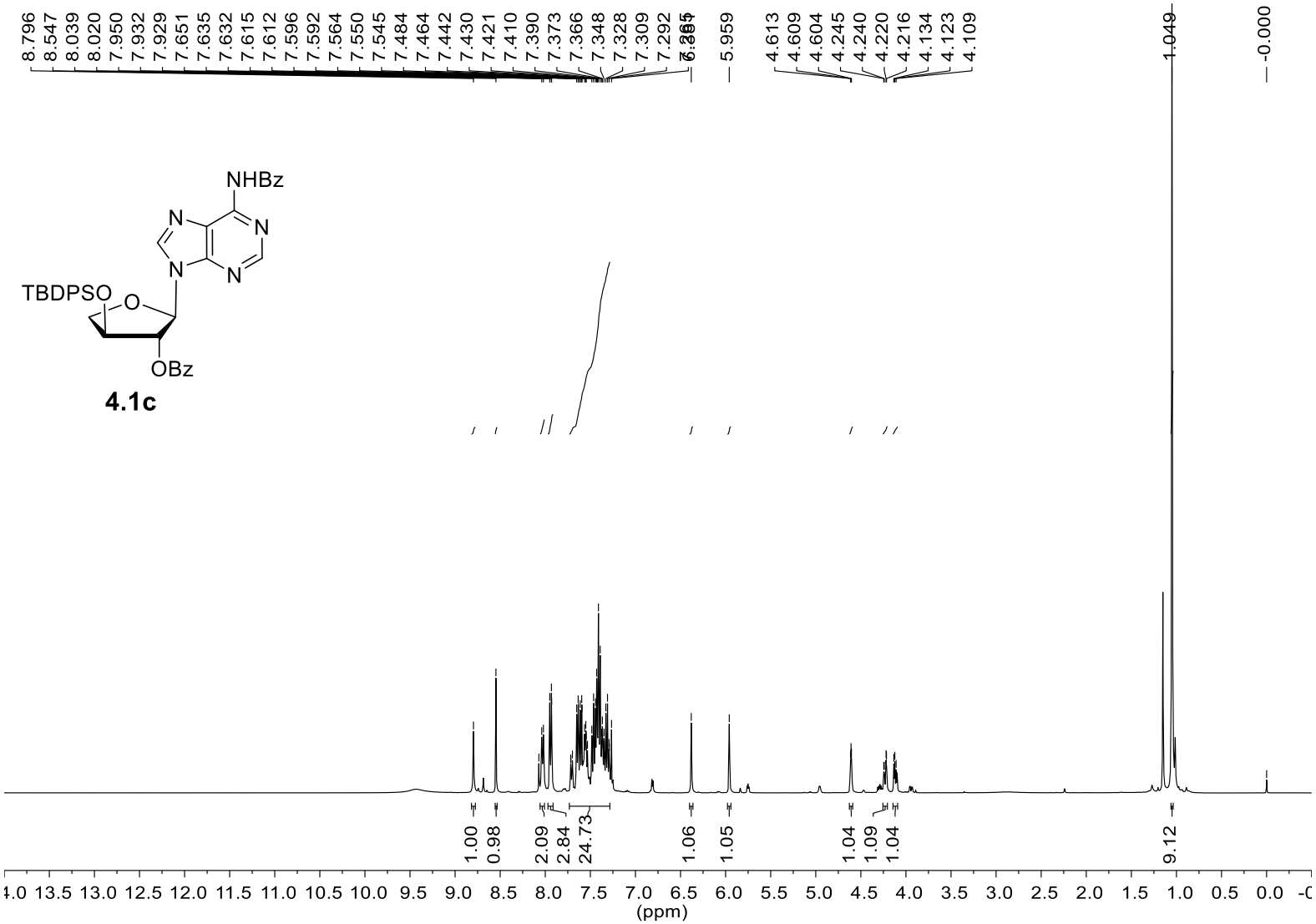
^{13}C NMR spectrum of compound **4.1a** (125.8 MHz, $CDCl_3$)



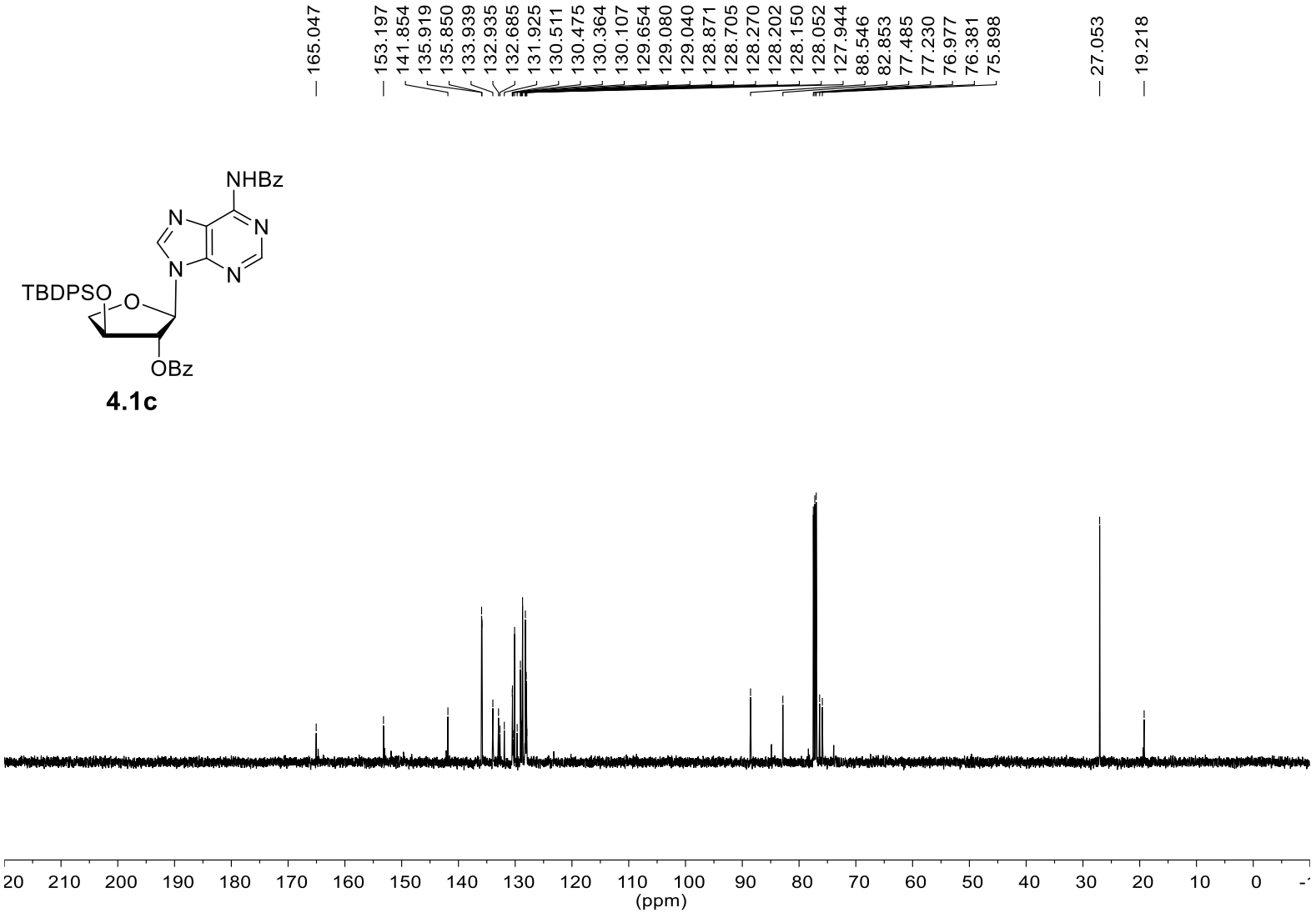
^{1}H NMR spectrum of compound **4.1b** (400 MHz, $CDCl_3$)



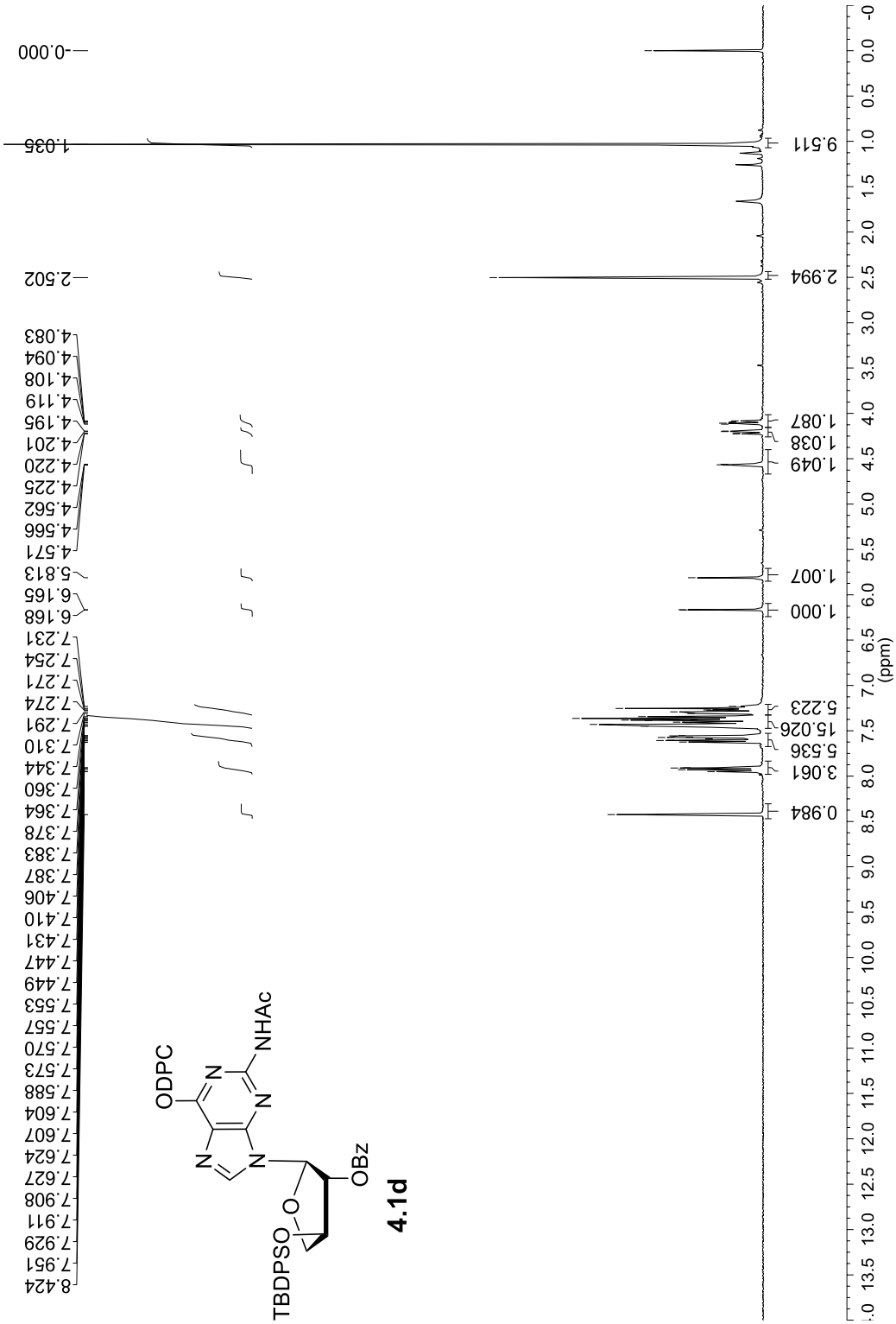
^{13}C NMR spectrum of compound **4.1b** (125.8 MHz, $CDCl_3$)



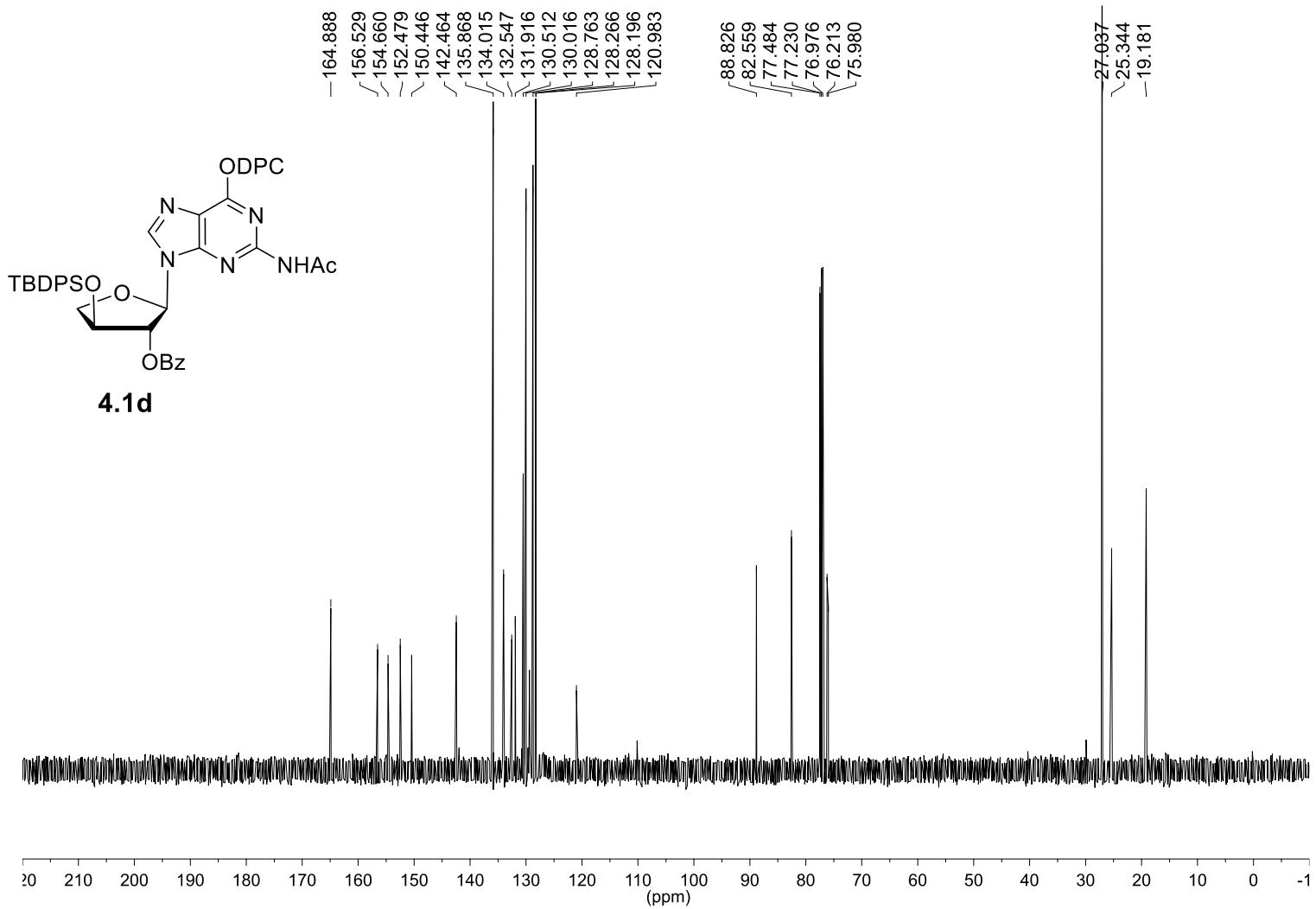
^{1}H NMR spectrum of compound **4.1c** (400 MHz, $CDCl_3$)



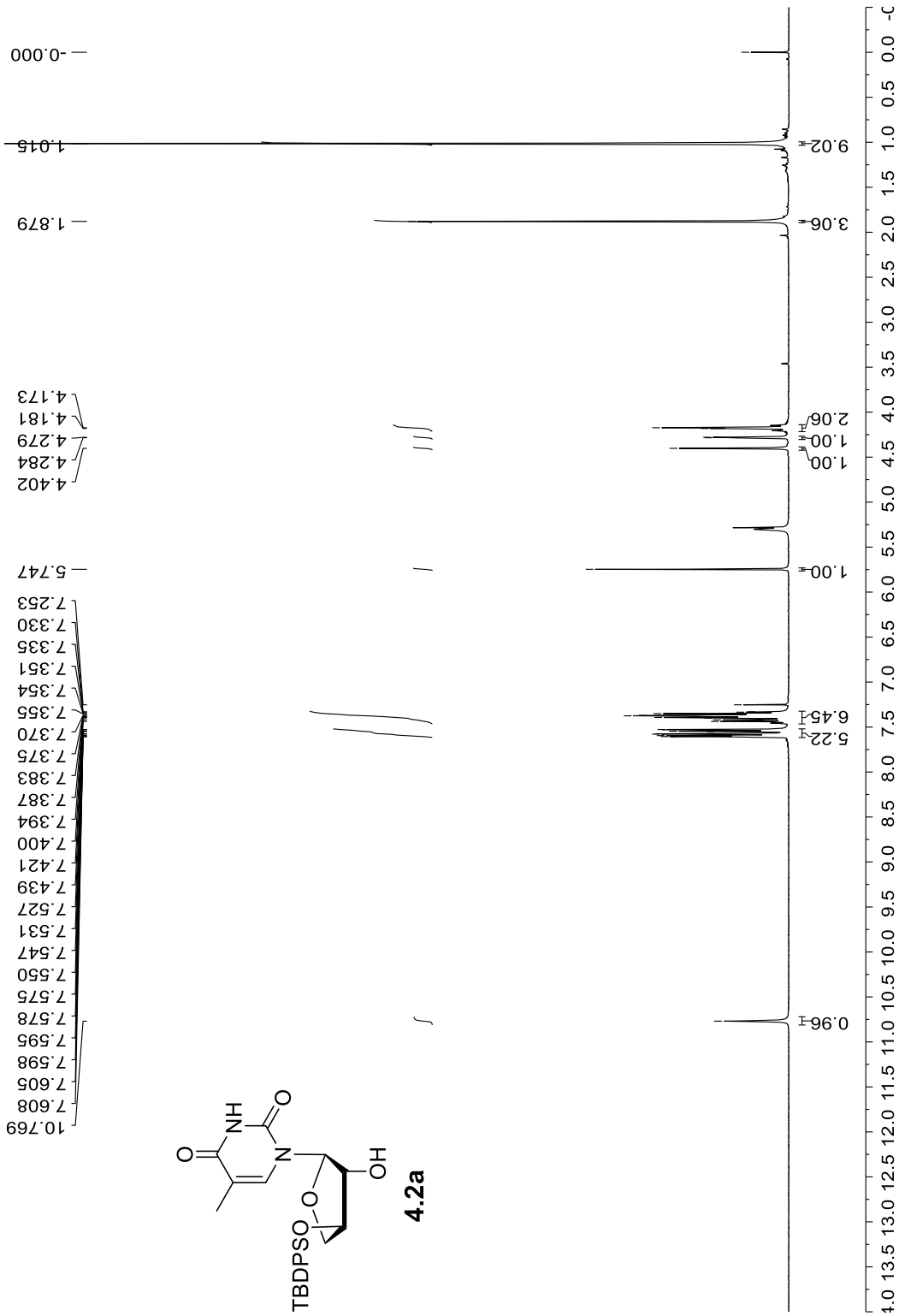
^{13}C NMR spectrum of compound **4.1c** (125.8 MHz, $CDCl_3$)



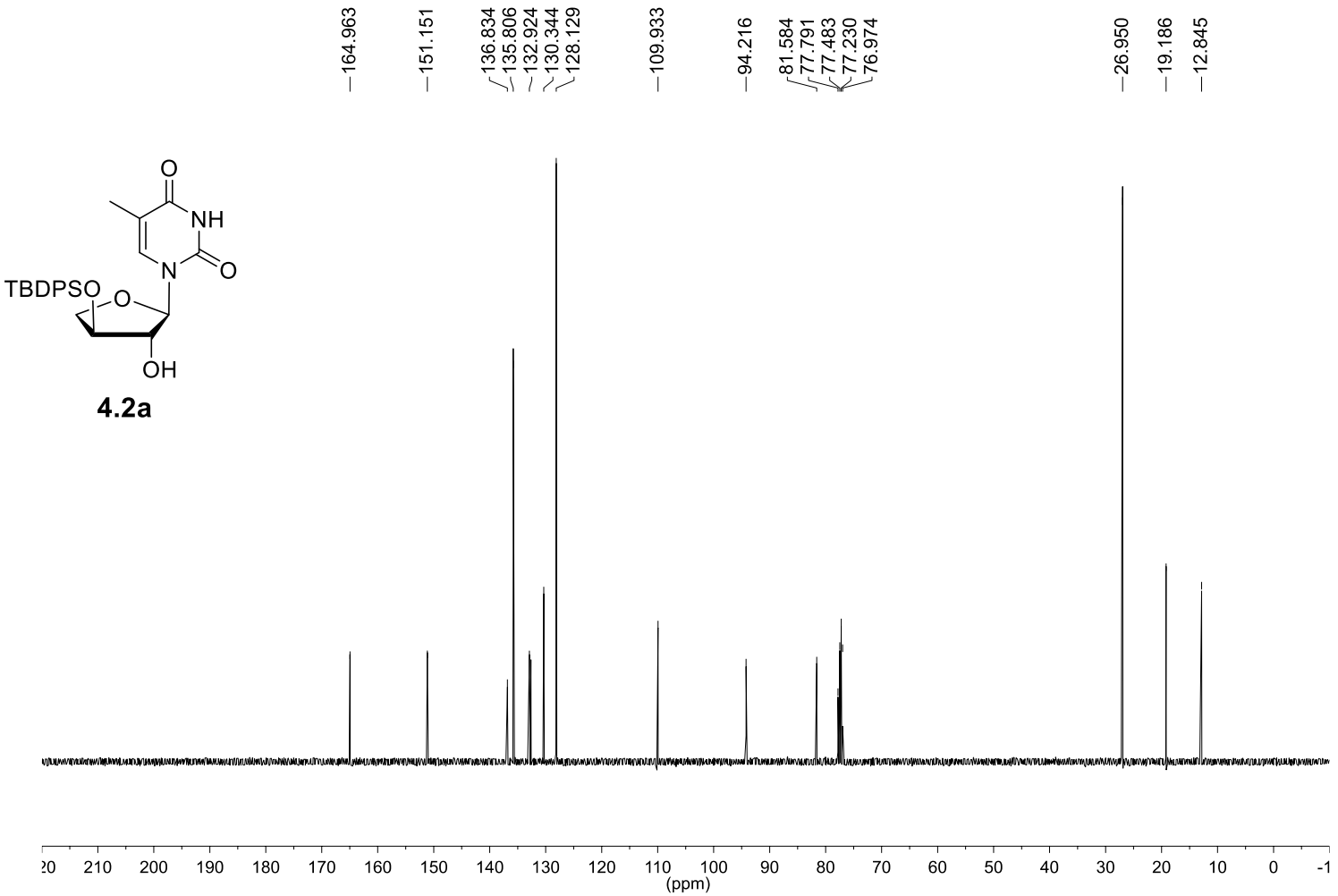
^{1}H NMR spectrum of compound **4.1d** (400 MHz, $CDCl_3$)



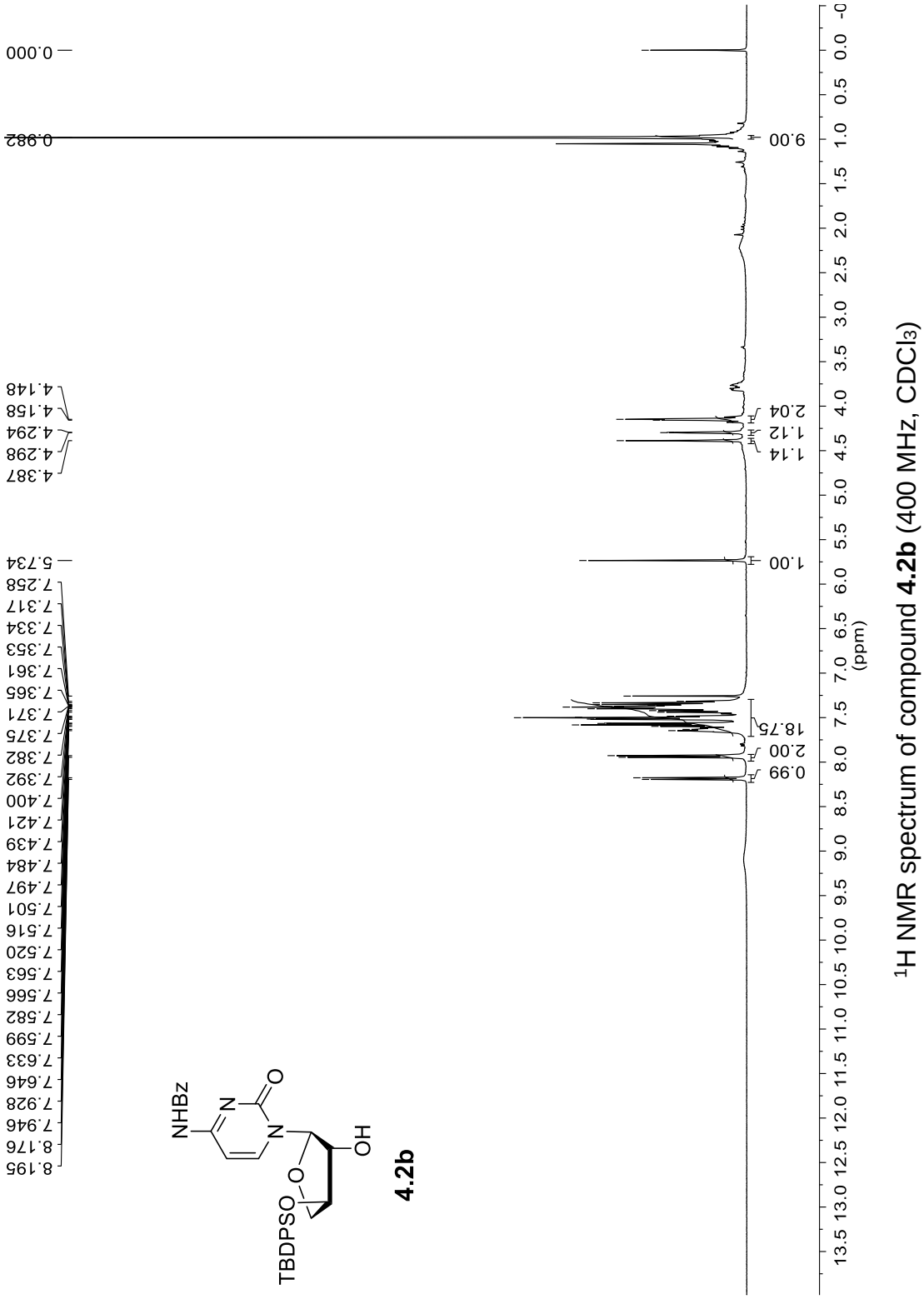
^{13}C NMR spectrum of compound **4.1d** (125.8 MHz, $CDCl_3$)



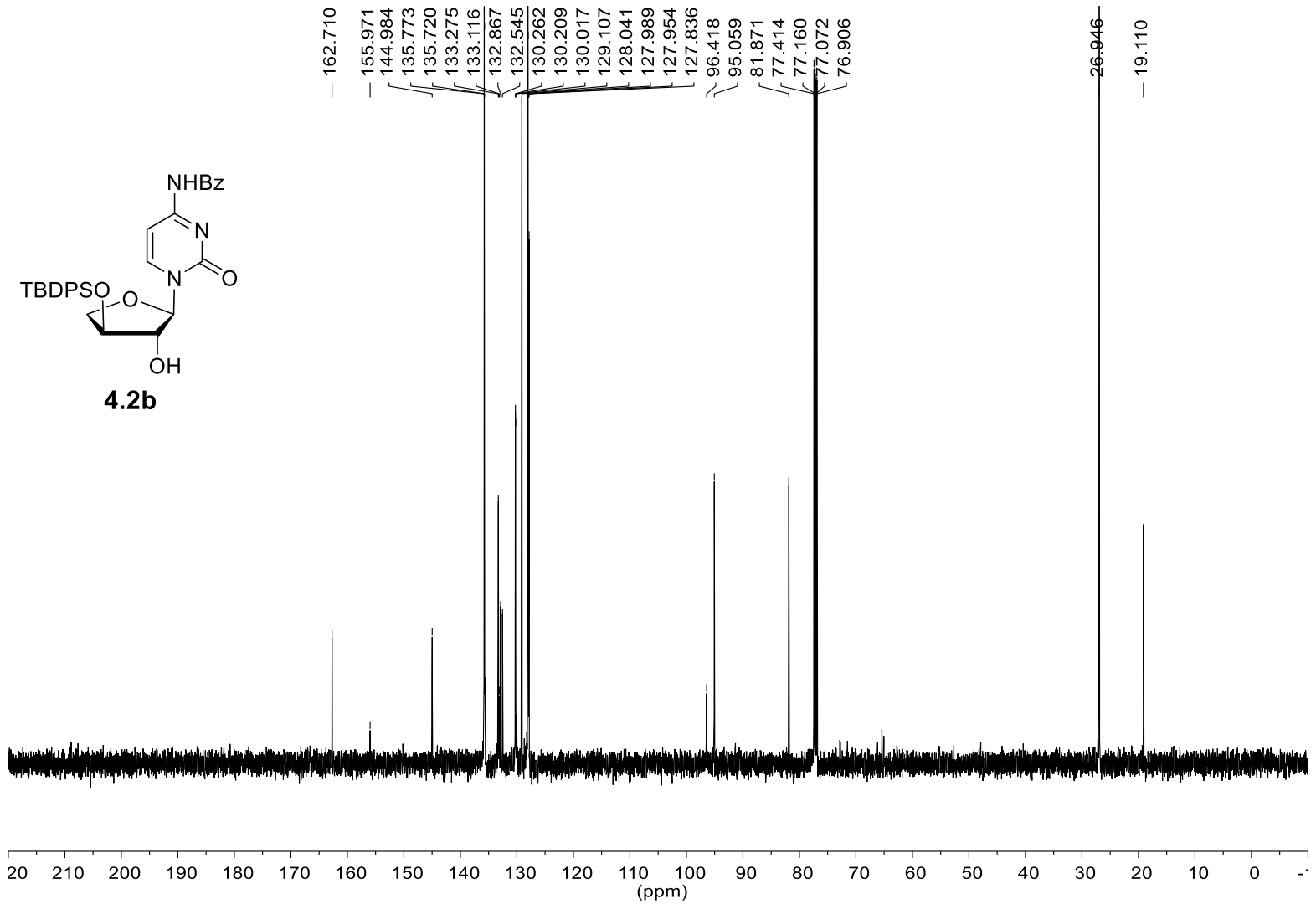
^{1}H NMR spectrum of compound **4.2a** (400 MHz, $CDCl_3$)



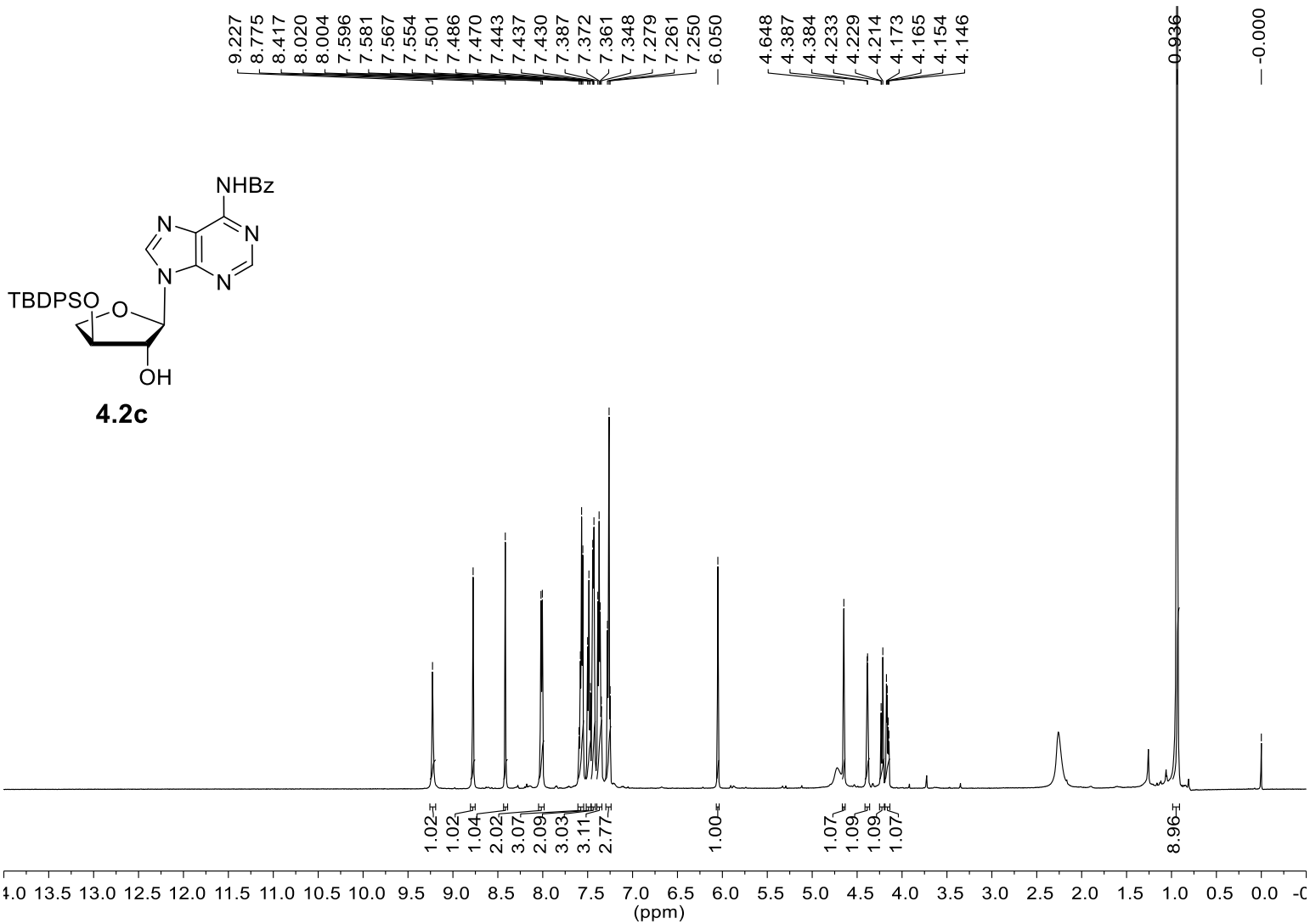
^{13}C NMR spectrum of compound **4.2a** (125.8 MHz, CD_3OD)



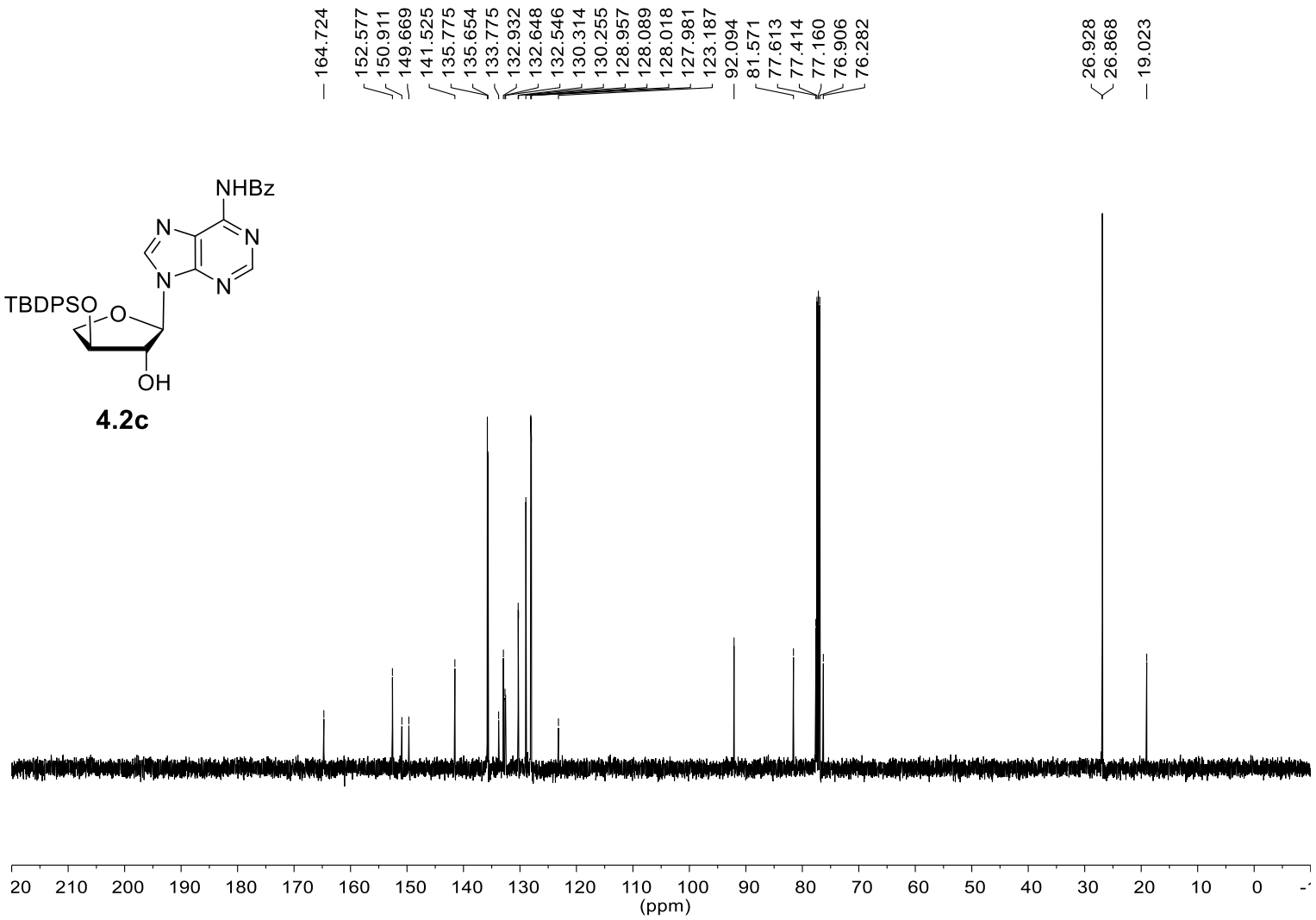
^{1}H NMR spectrum of compound **4.2b** (400 MHz, $CDCl_3$)



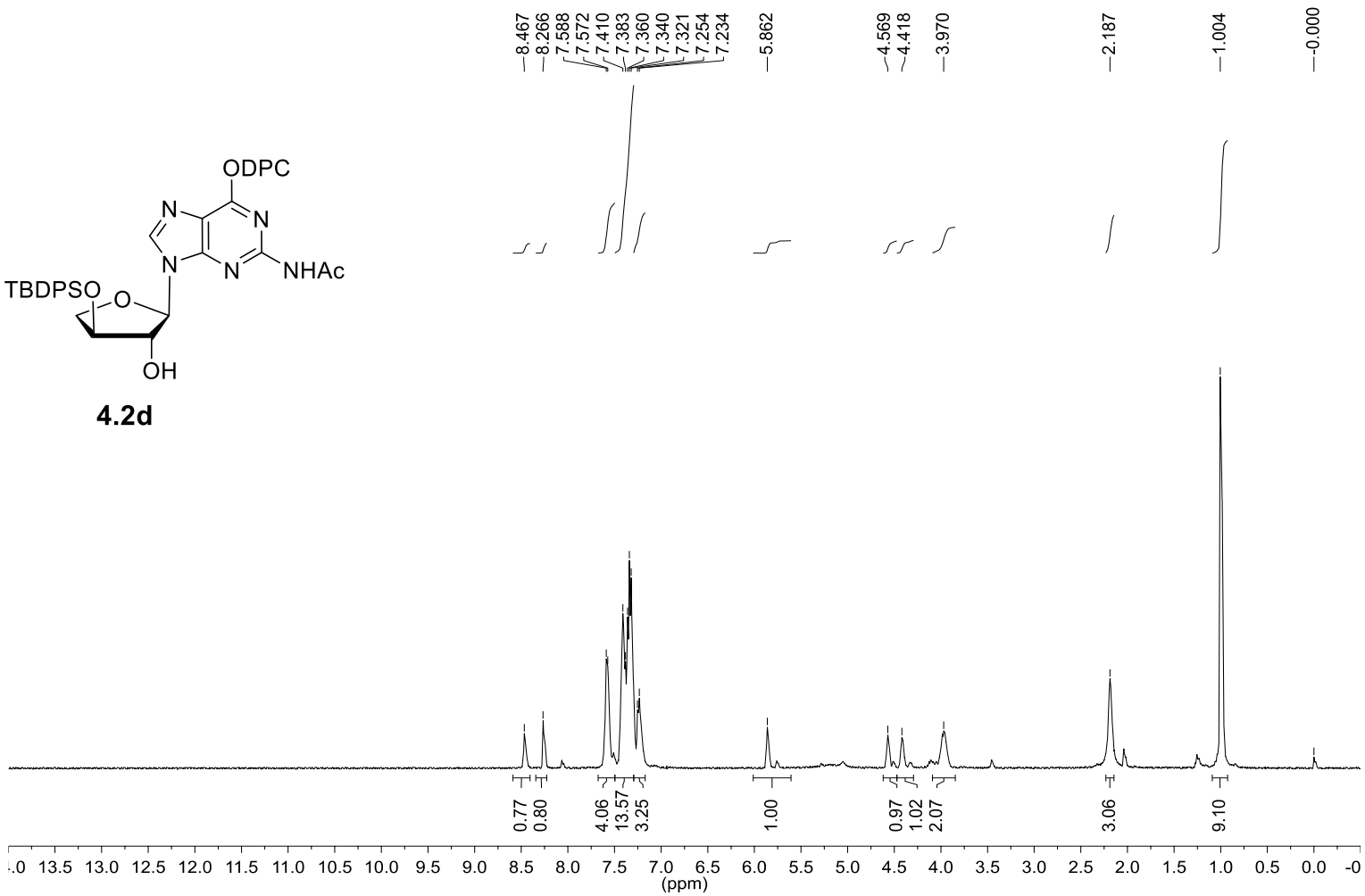
^{13}C NMR spectrum of compound **4.2b** (125.8 MHz, $CDCl_3$)



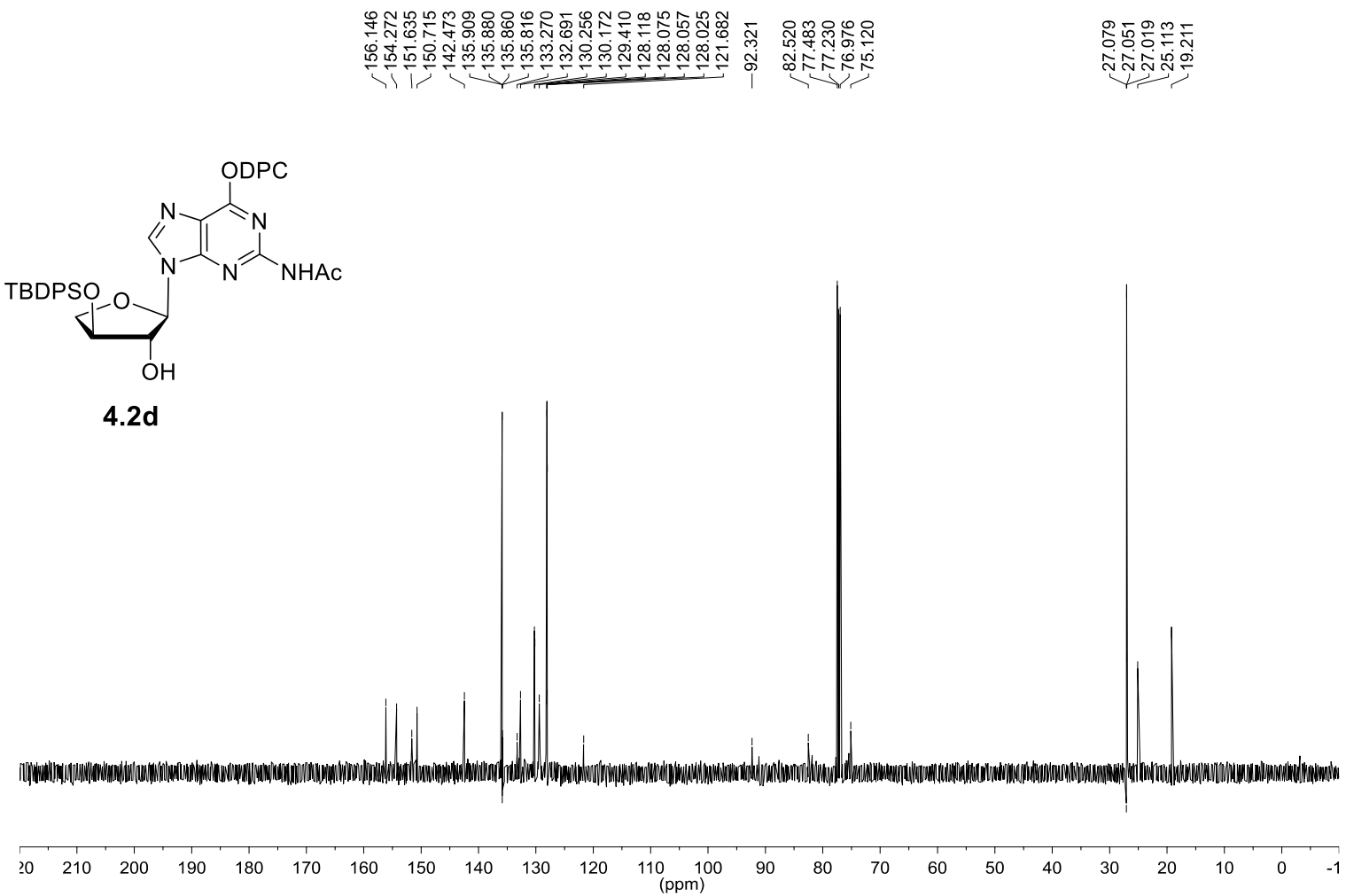
^{1}H NMR spectrum of compound **4.2c** (400 MHz, $CDCl_3$)



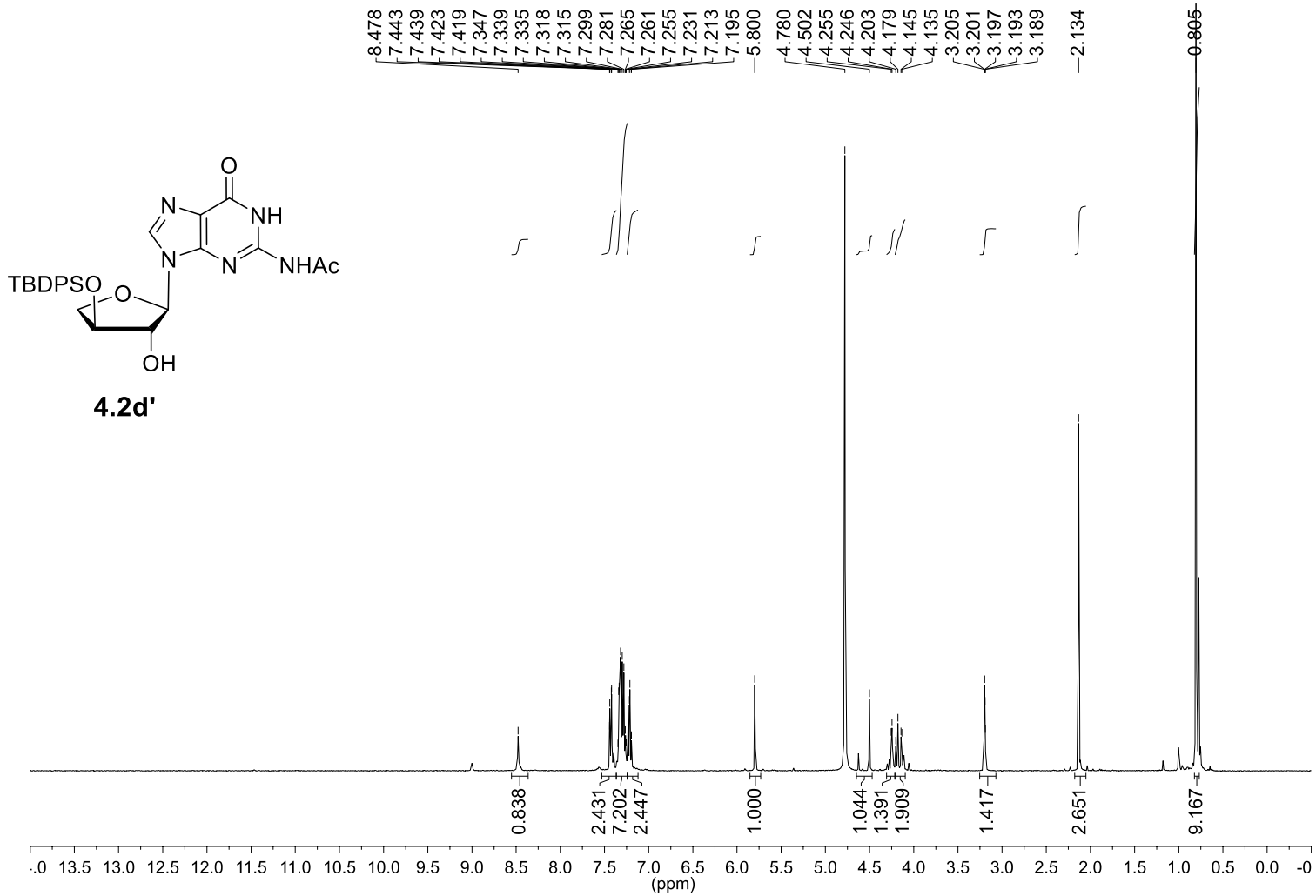
^{13}C NMR spectrum of compound **4.2c** (125.8 MHz, $CDCl_3$)



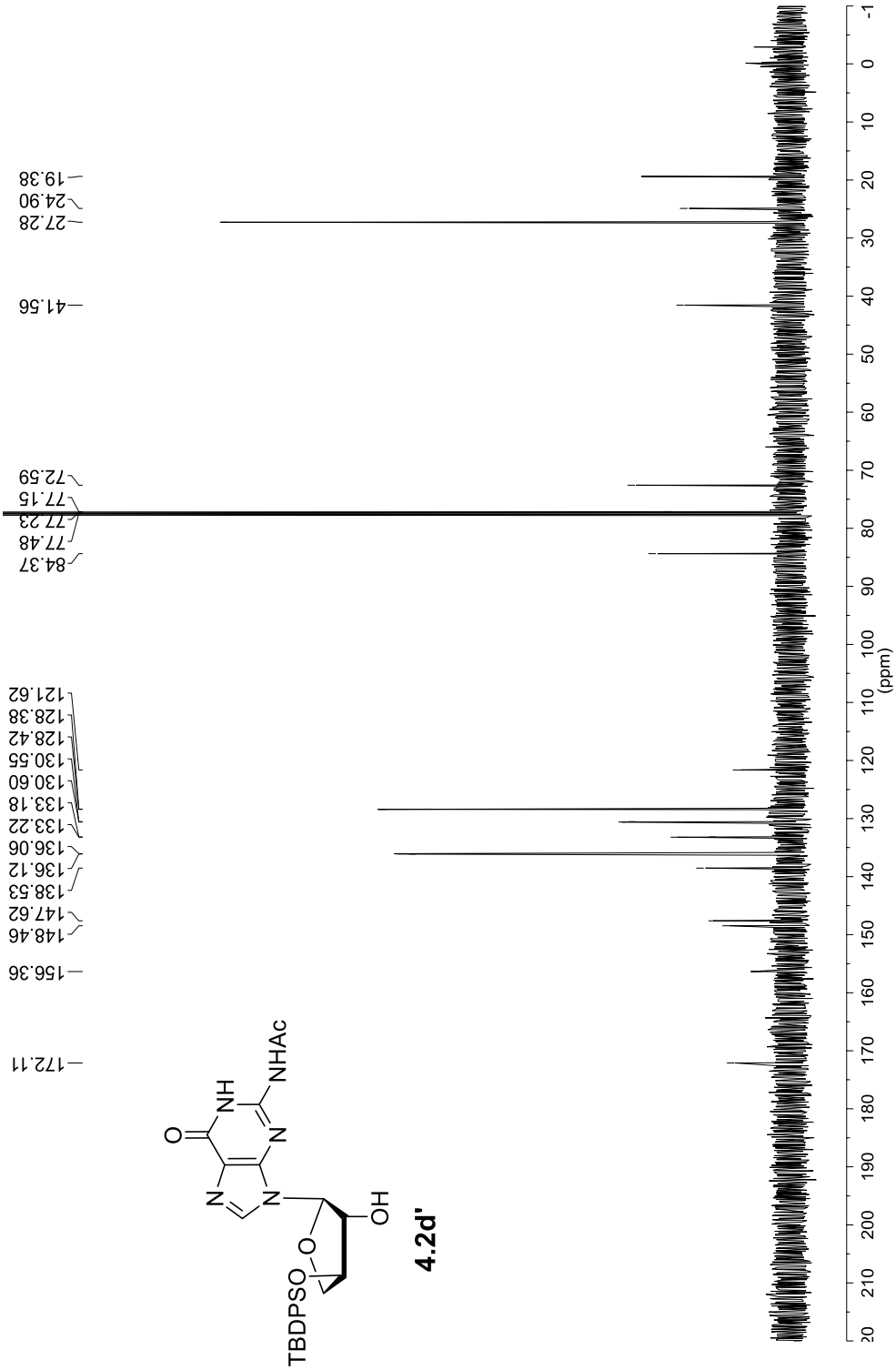
^{1}H NMR spectrum of compound **4.2d** (400 MHz, $CDCl_3$)



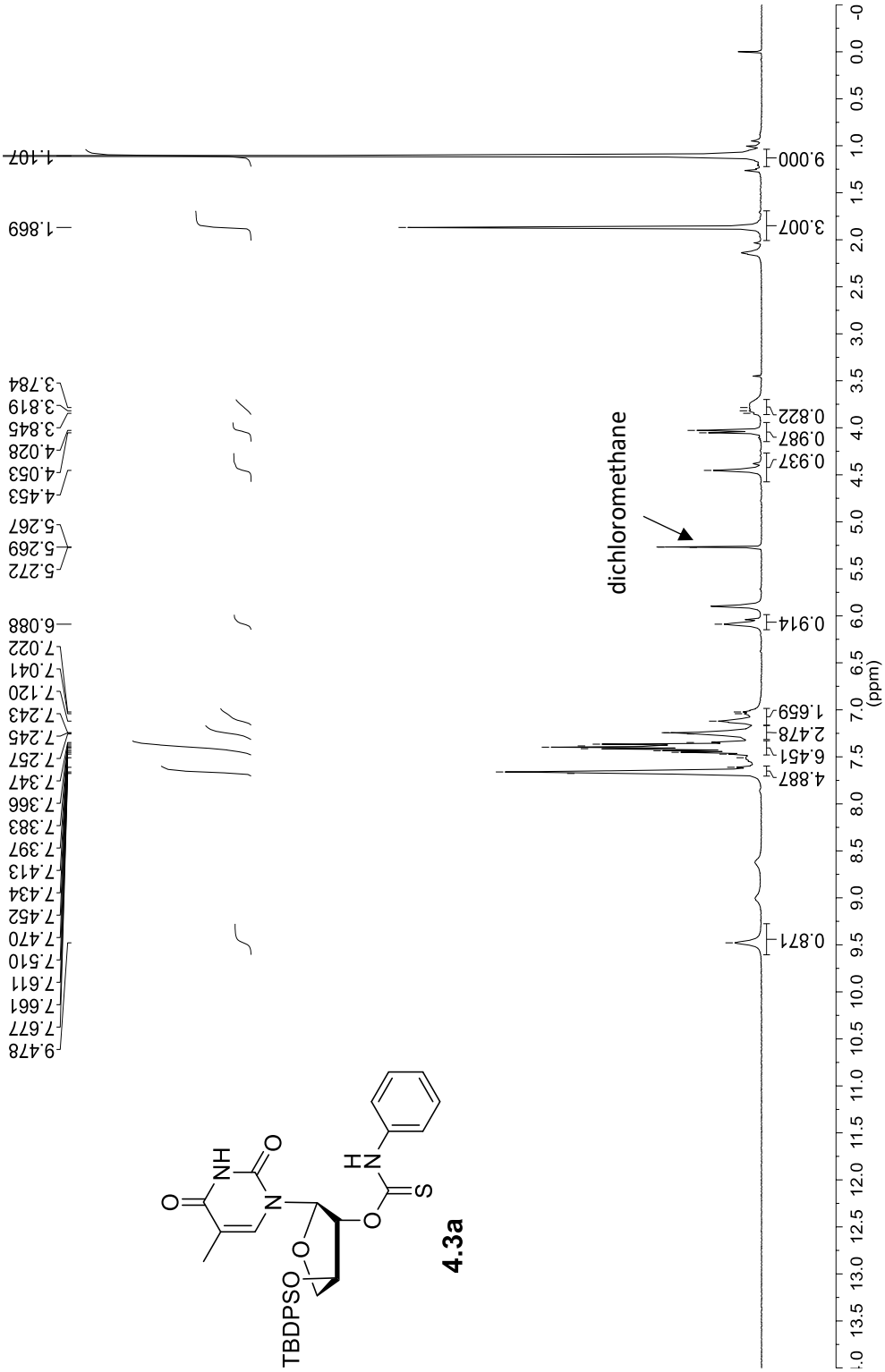
^{13}C NMR spectrum of compound **4.2d** (125.8 MHz, $CDCl_3$)



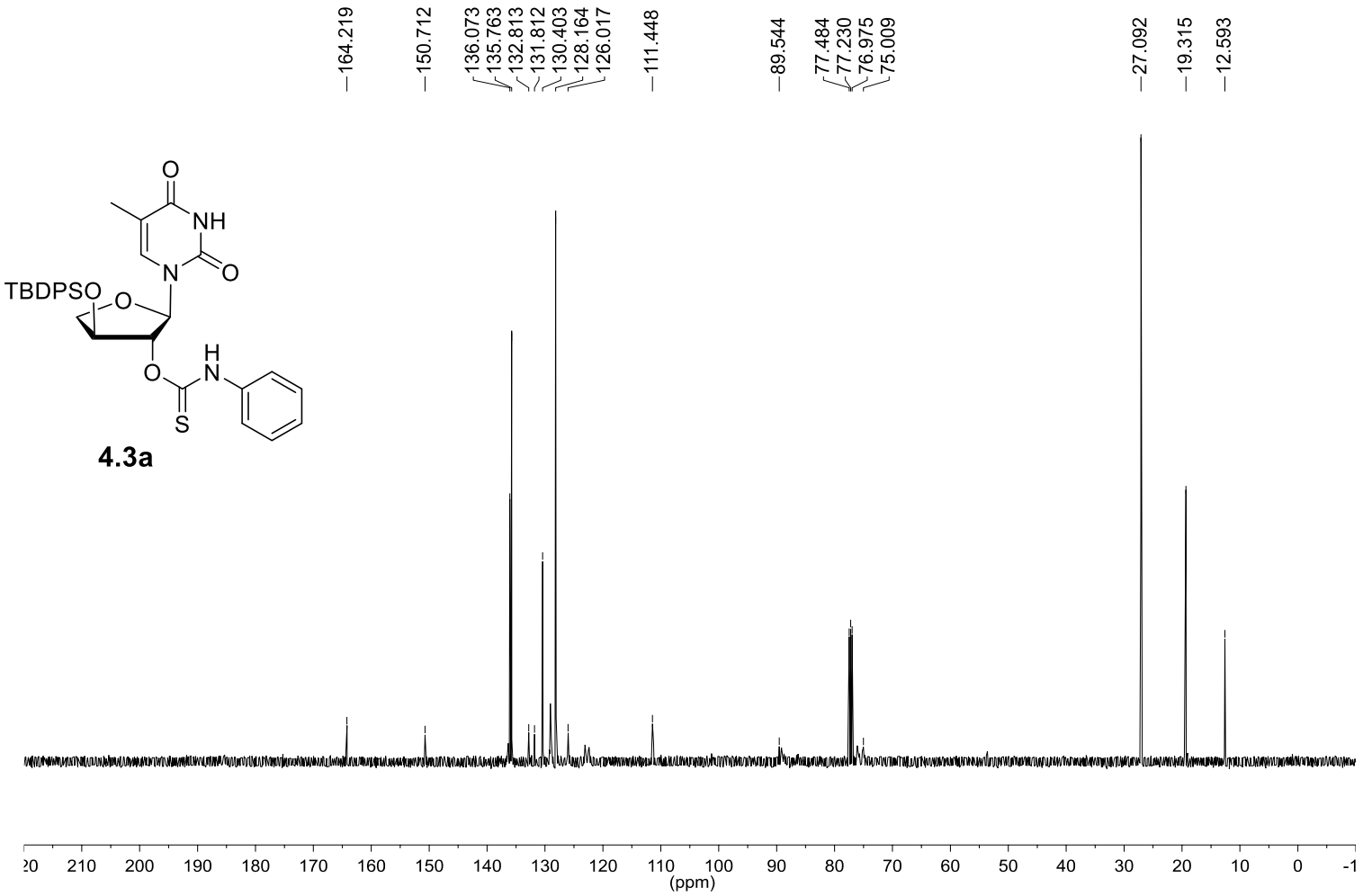
^{1}HNMR spectrum of compound **4.2d'** (400 MHz, CD_3OD)



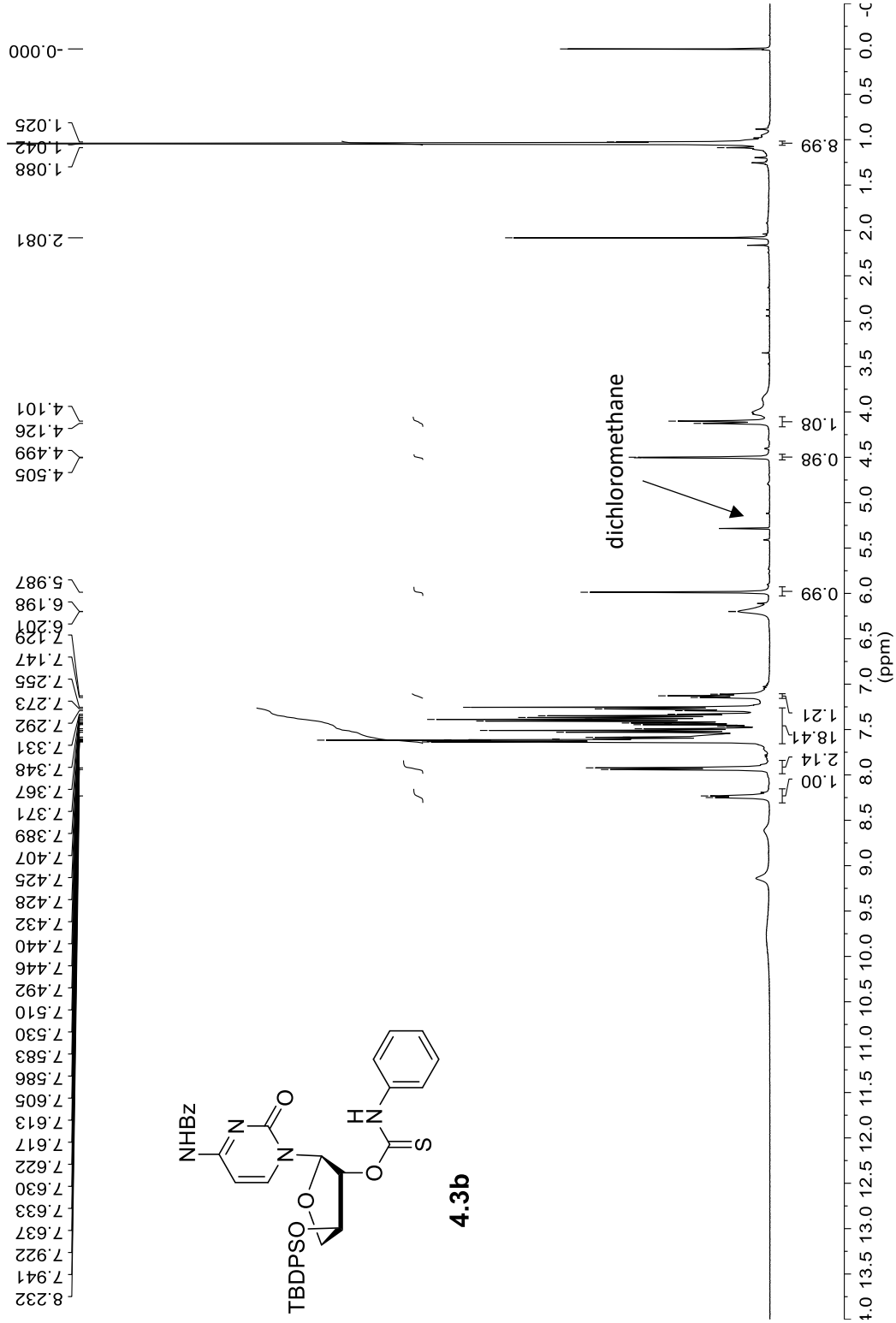
^{13}C NMR spectrum of compound **4.2d'** (125.8 MHz, CD_3OD)



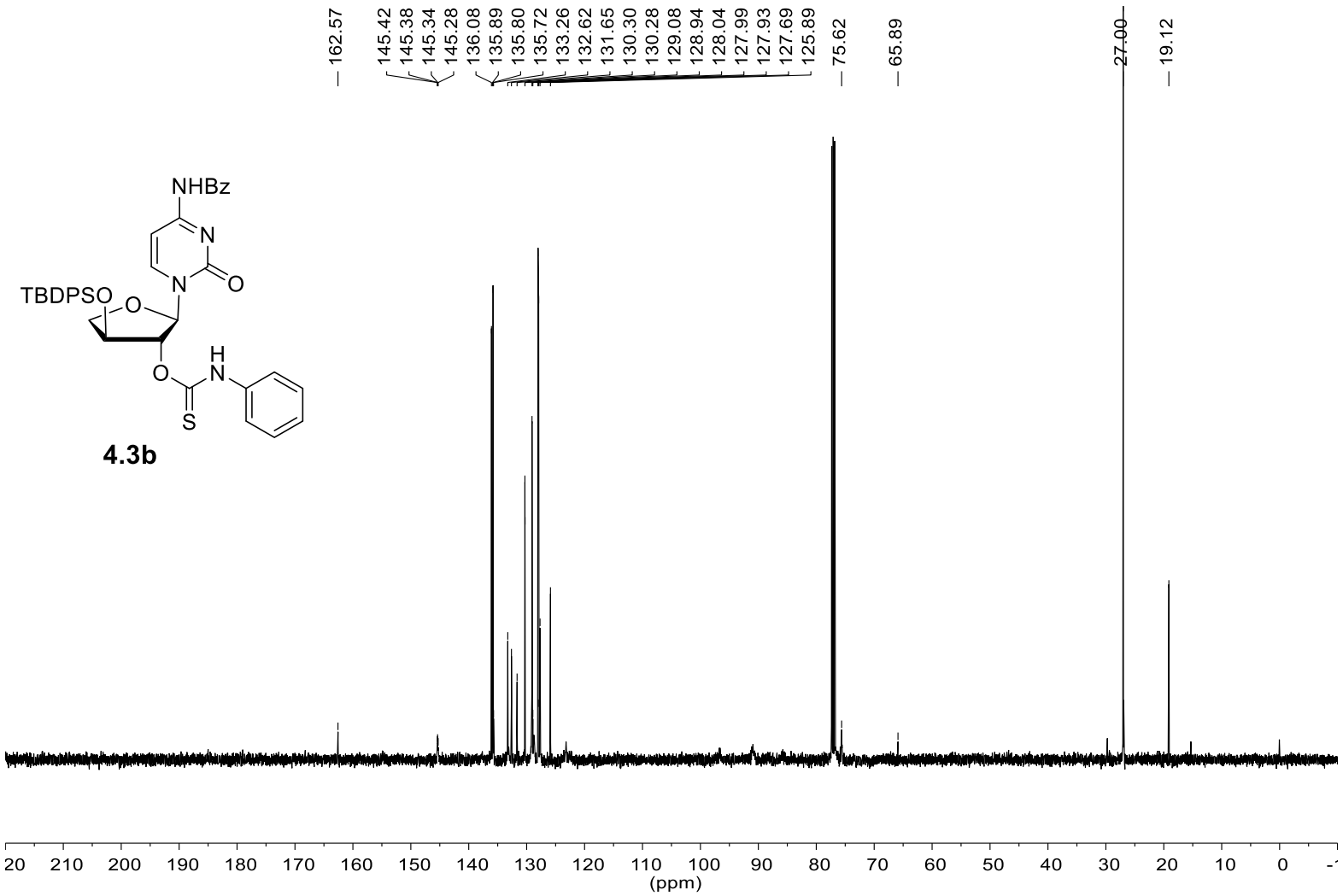
^{1}H NMR spectrum of compound **4.3a** (400 MHz, $CDCl_3$)



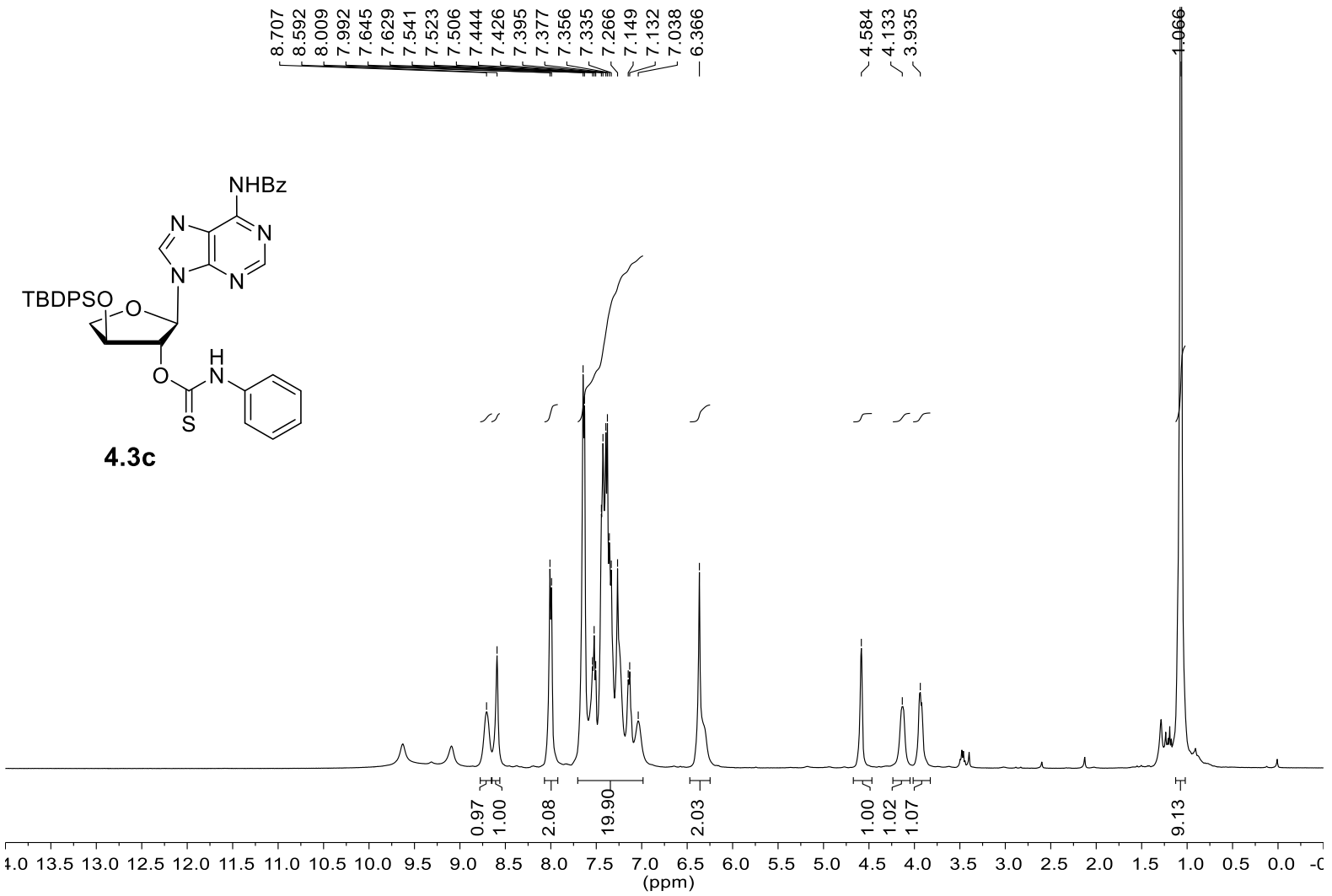
^{13}C NMR spectrum of compound **4.3a** (125.8 MHz, $CDCl_3$)



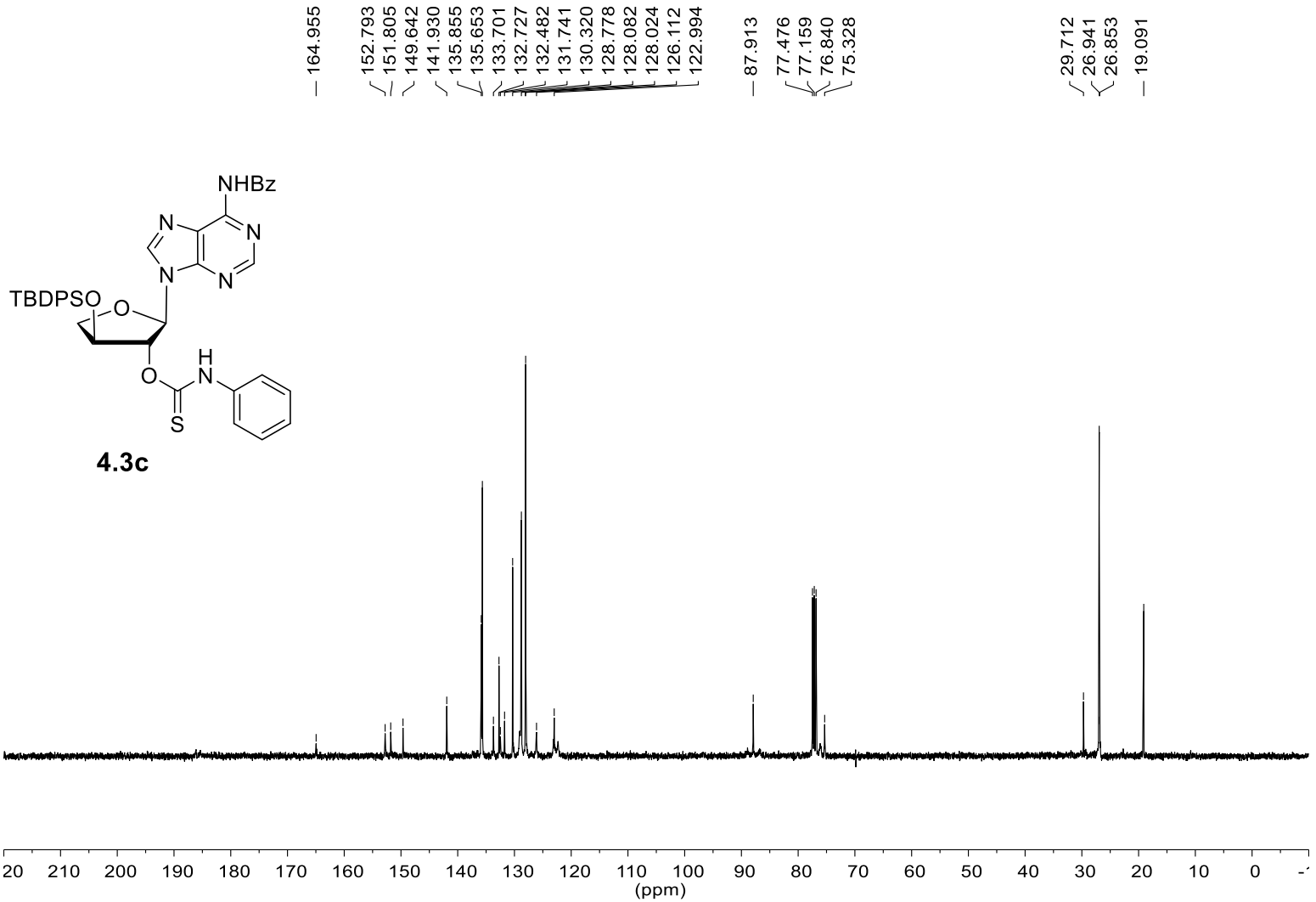
^{1}H NMR spectrum of compound **4.3b** (400 MHz, $CDCl_3$)



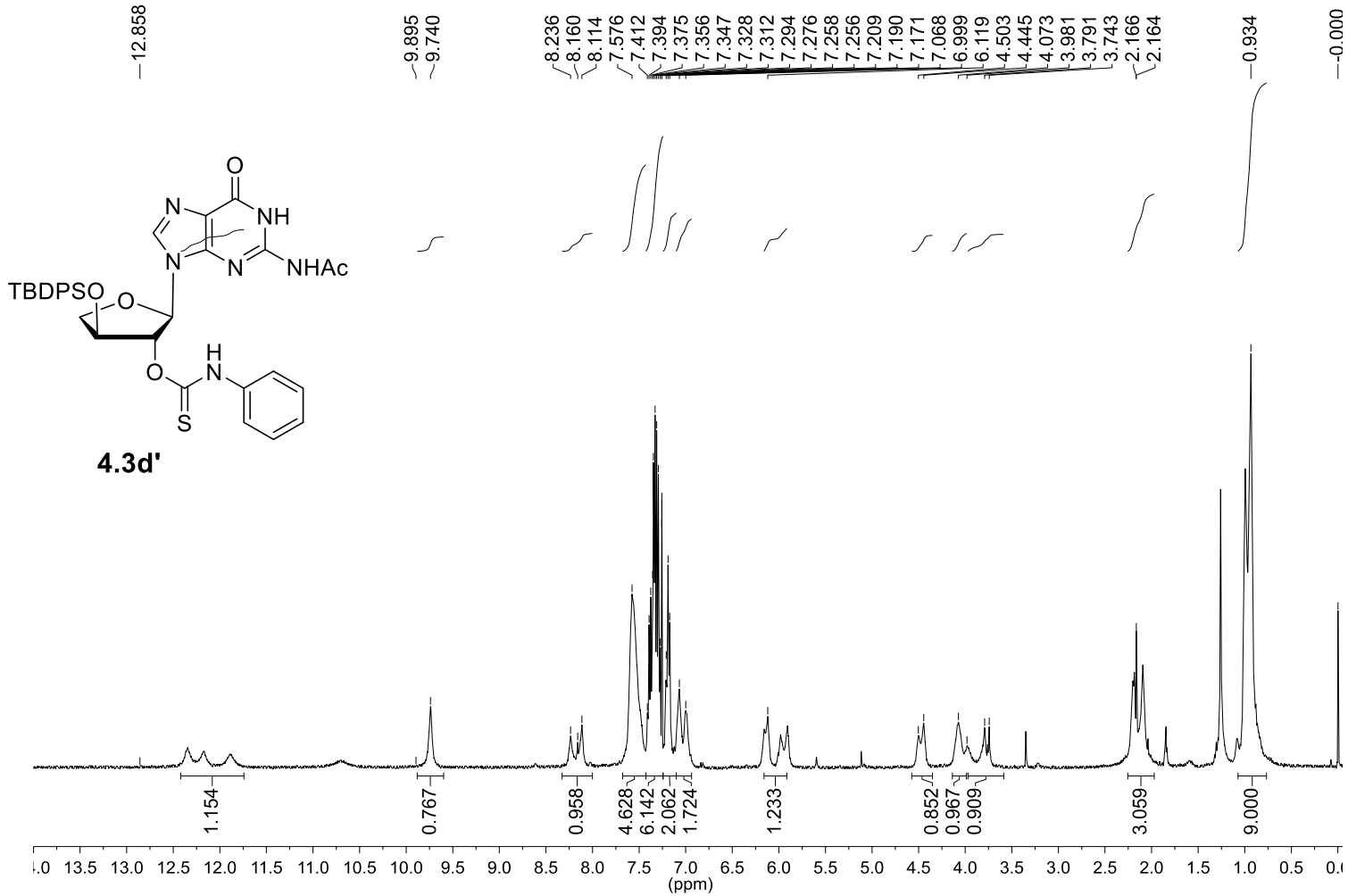
^{13}C NMR spectrum of compound **4.3b** (125.8 MHz, $CDCl_3$)



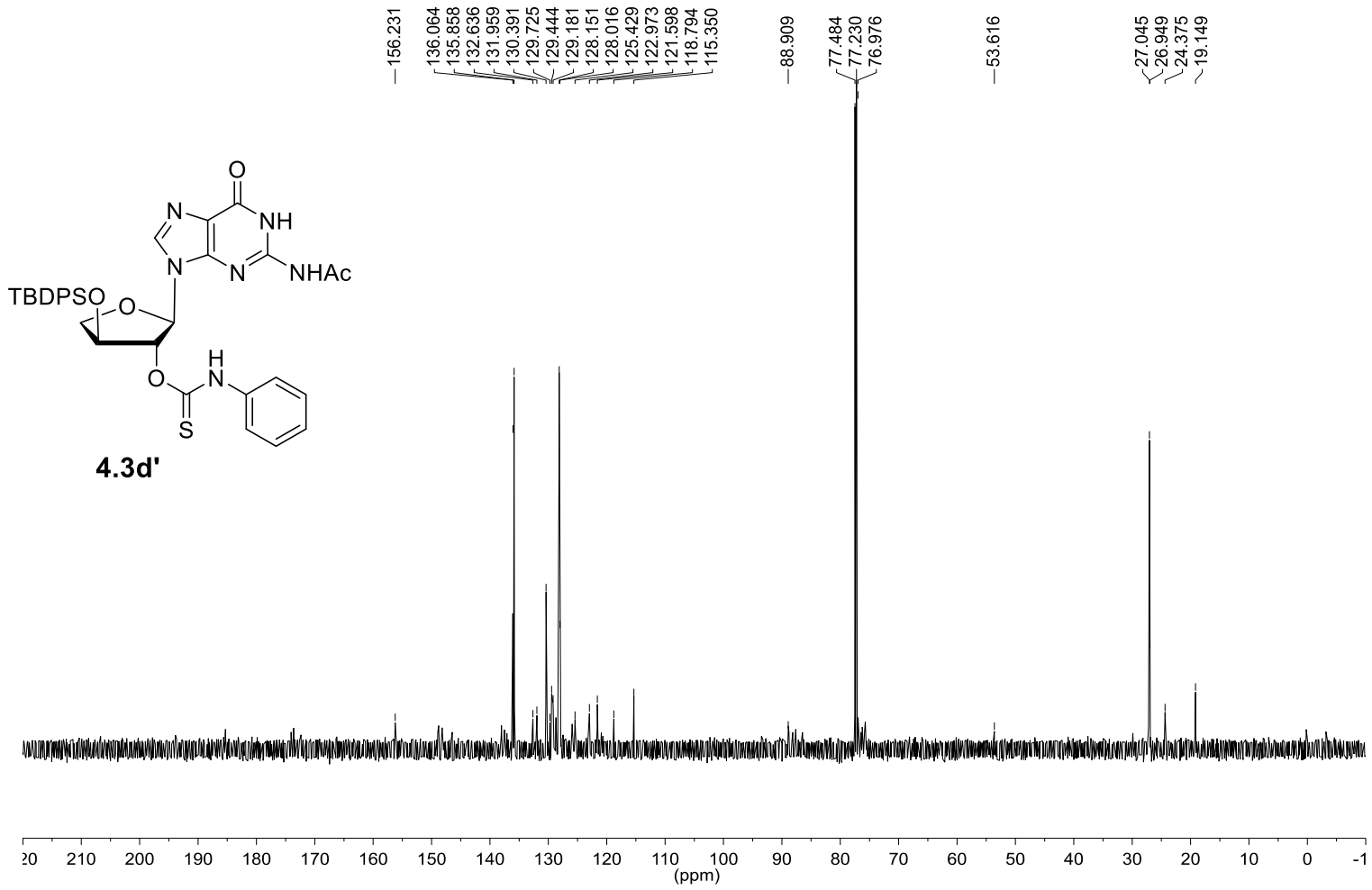
^{1}H NMR spectrum of compound **4.3c** (400 MHz, $CDCl_3$)



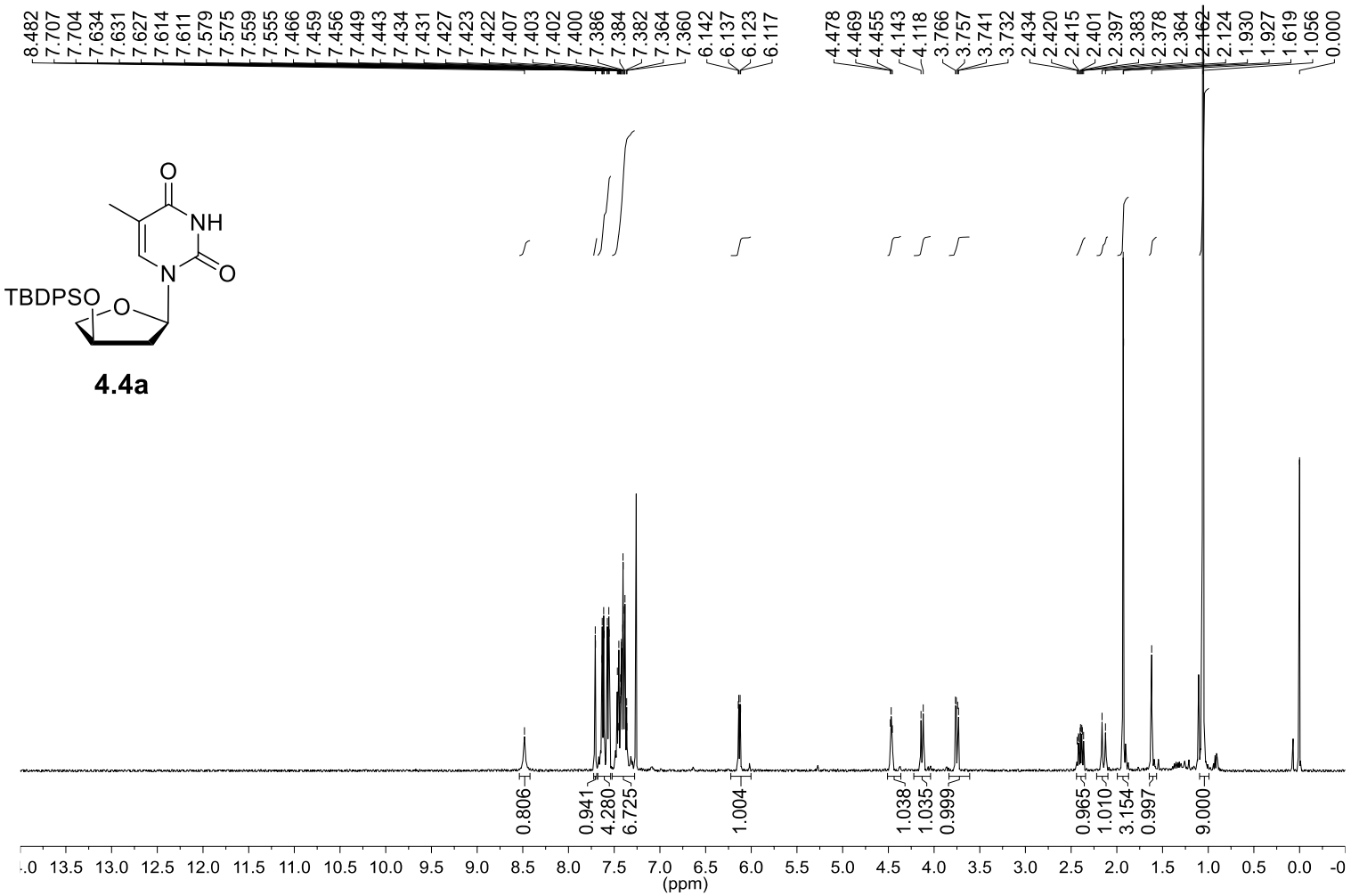
^{13}C NMR spectrum of compound **4.3c** (100.6 MHz, $CDCl_3$)



^{1}H NMR spectrum of compound **4.3d'** (400 MHz, $CDCl_3$)

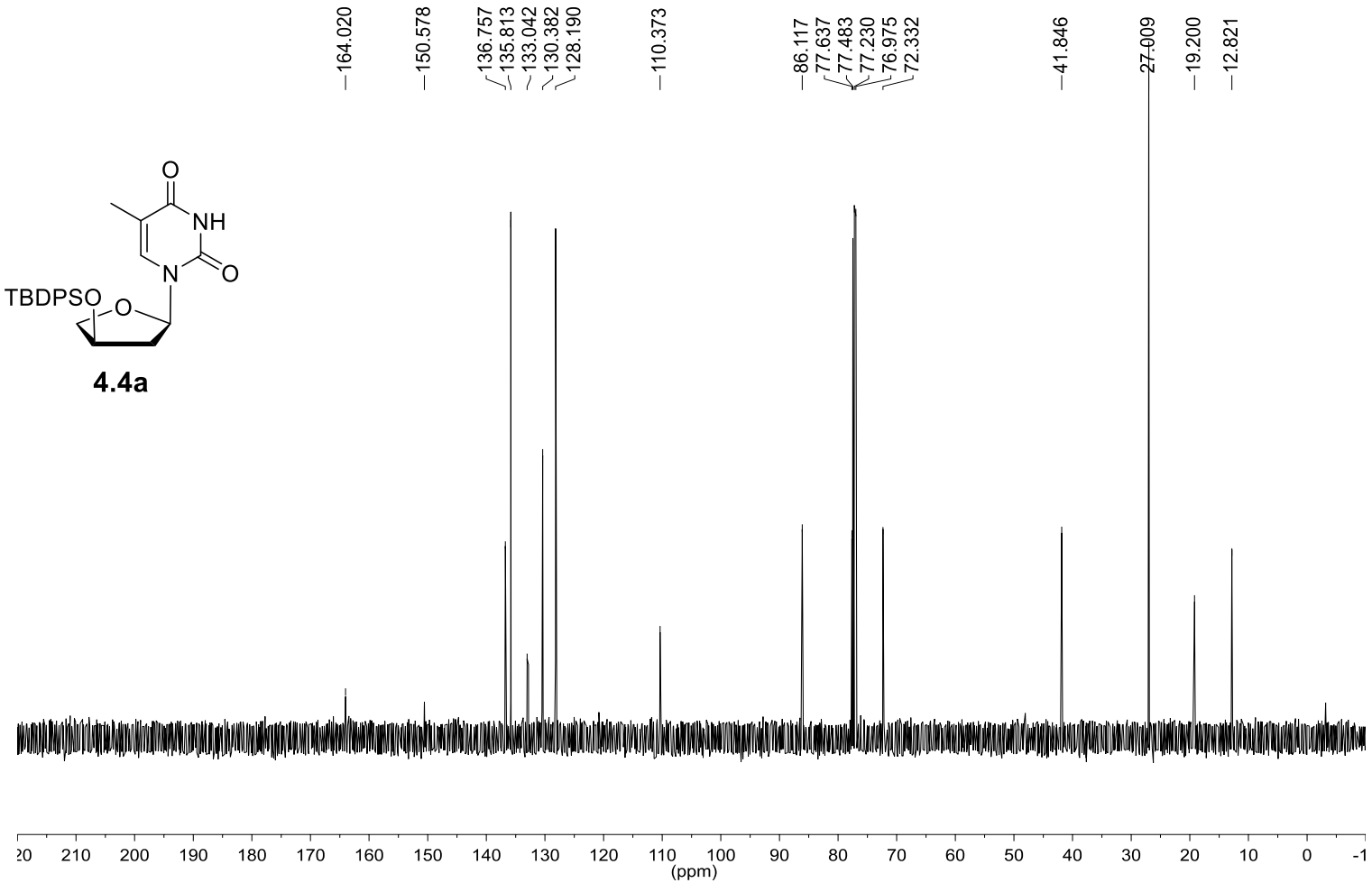


^{13}C NMR spectrum of compound **3d'** (125.8 MHz, $CDCl_3$)

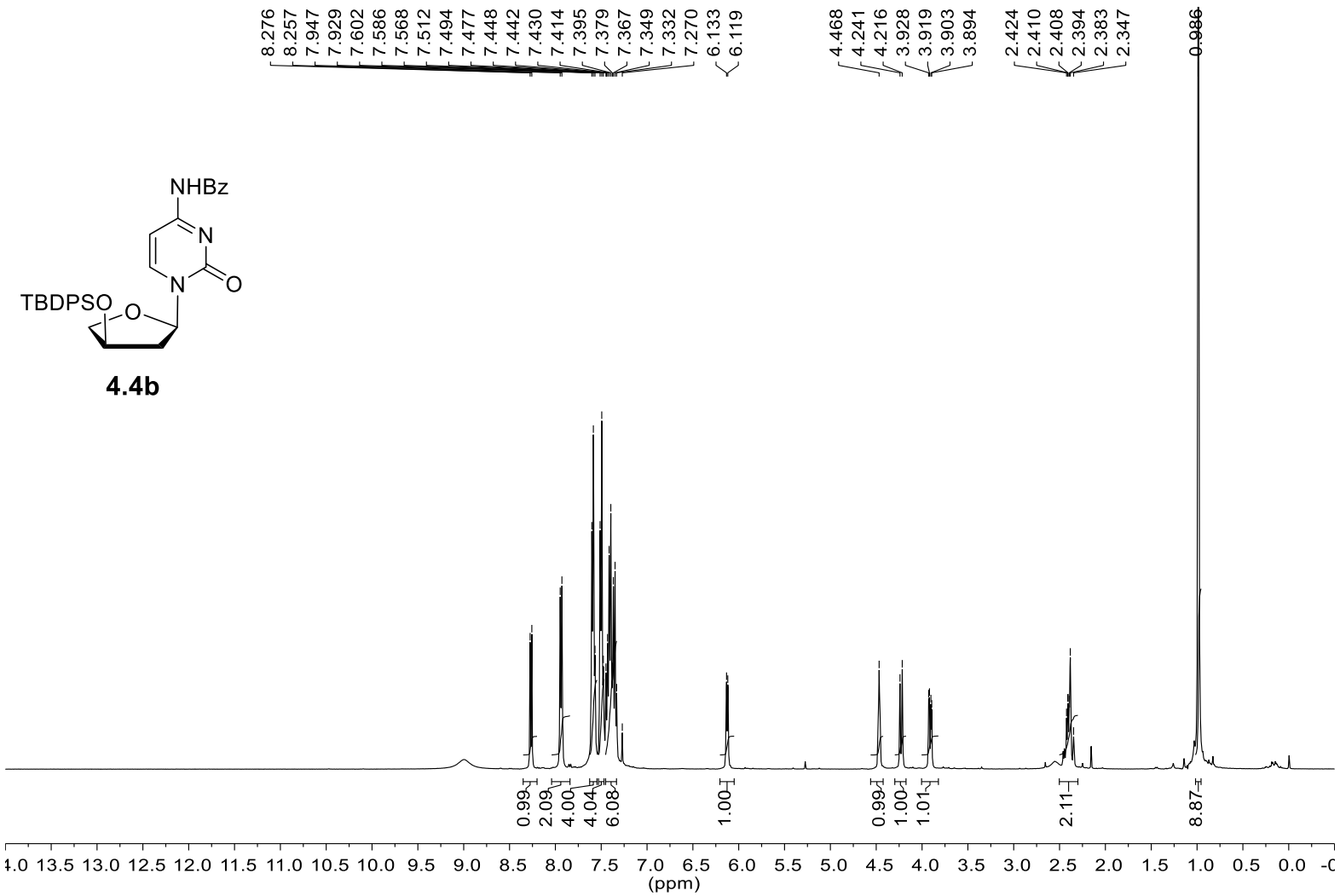


4.4a

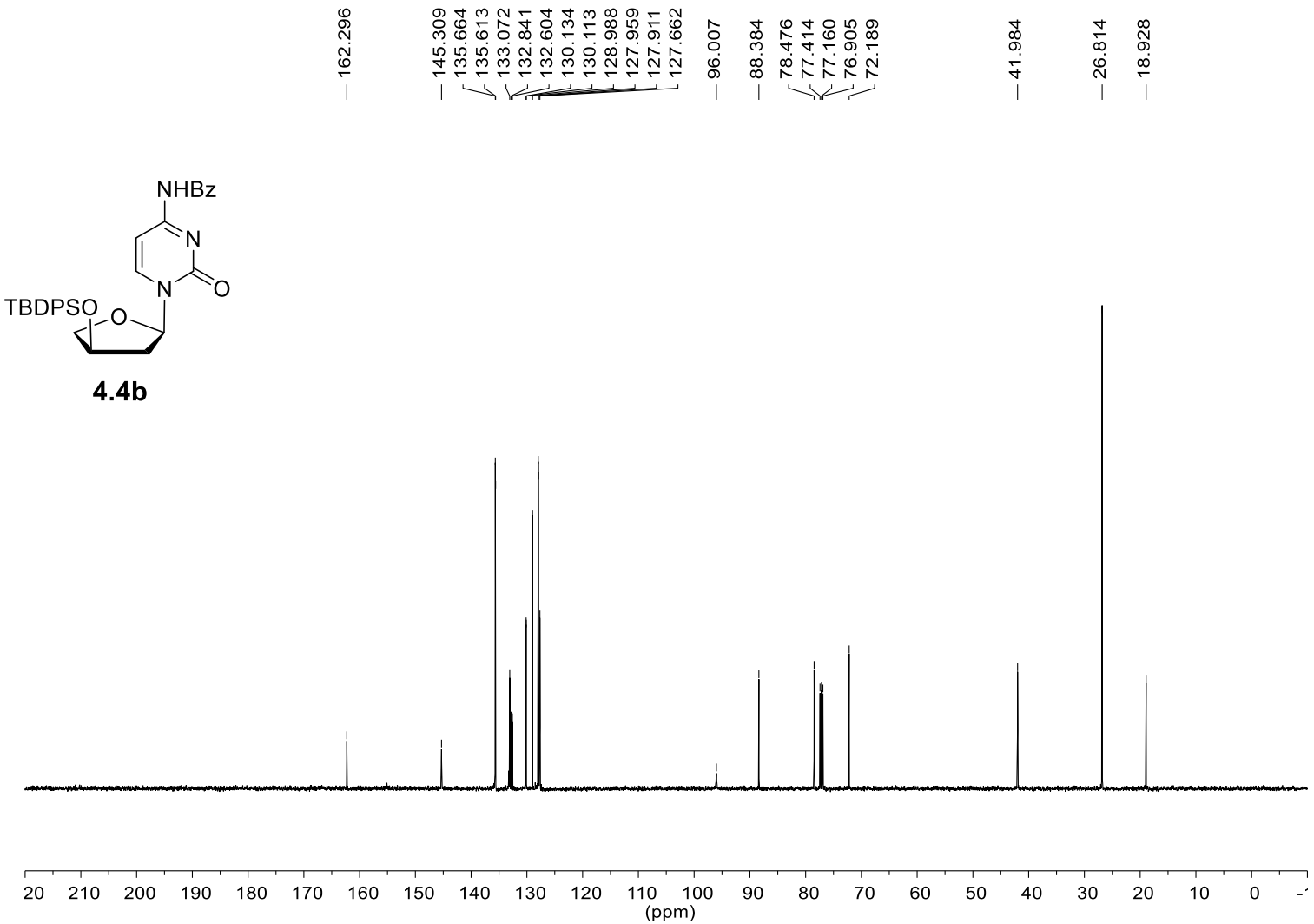
^{1}H NMR spectrum of compound **4.4a** (400 MHz, DMSO-d_6)



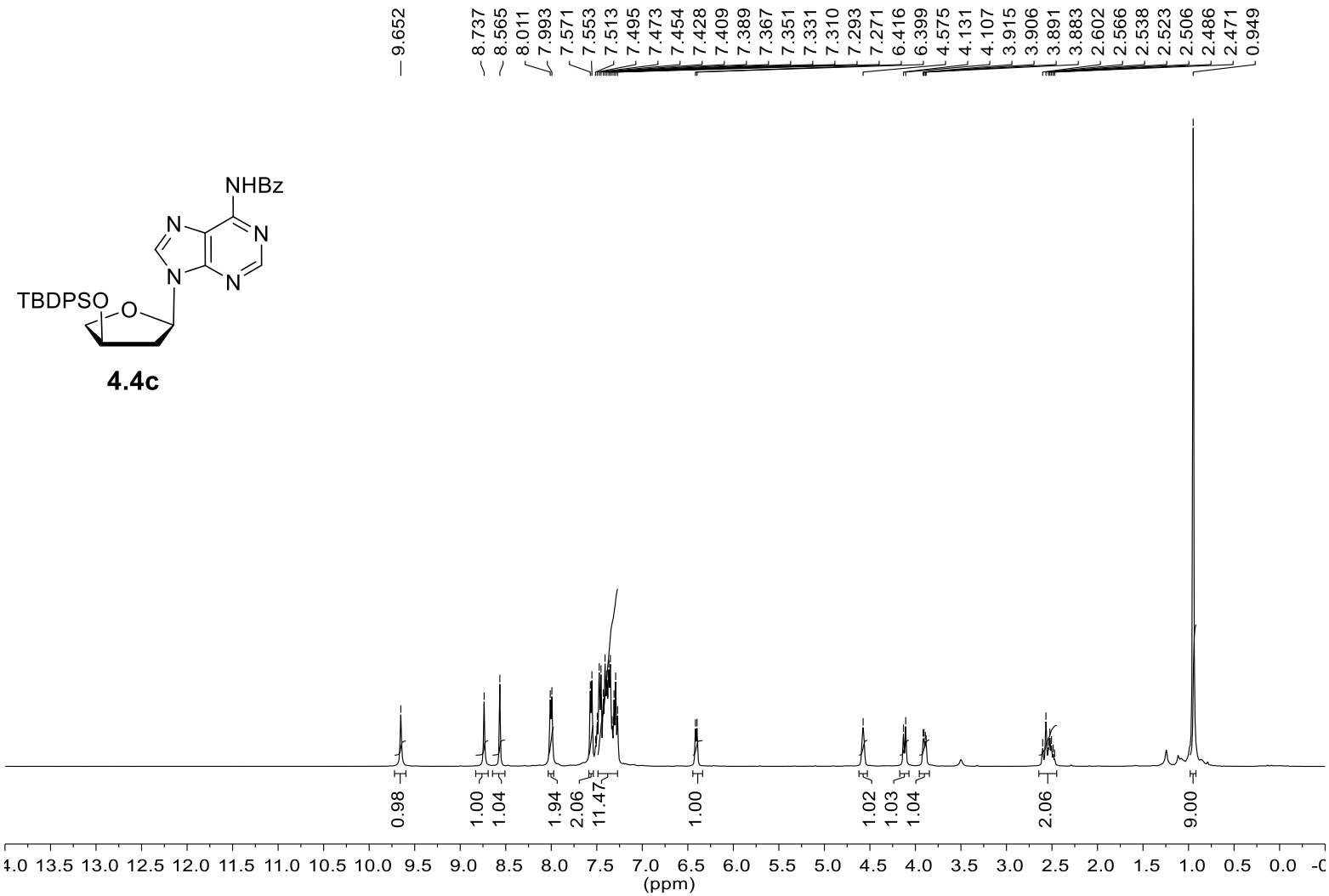
^{13}C NMR spectrum of compound **4.4a** (125.8 MHz, DMSO-d_6)



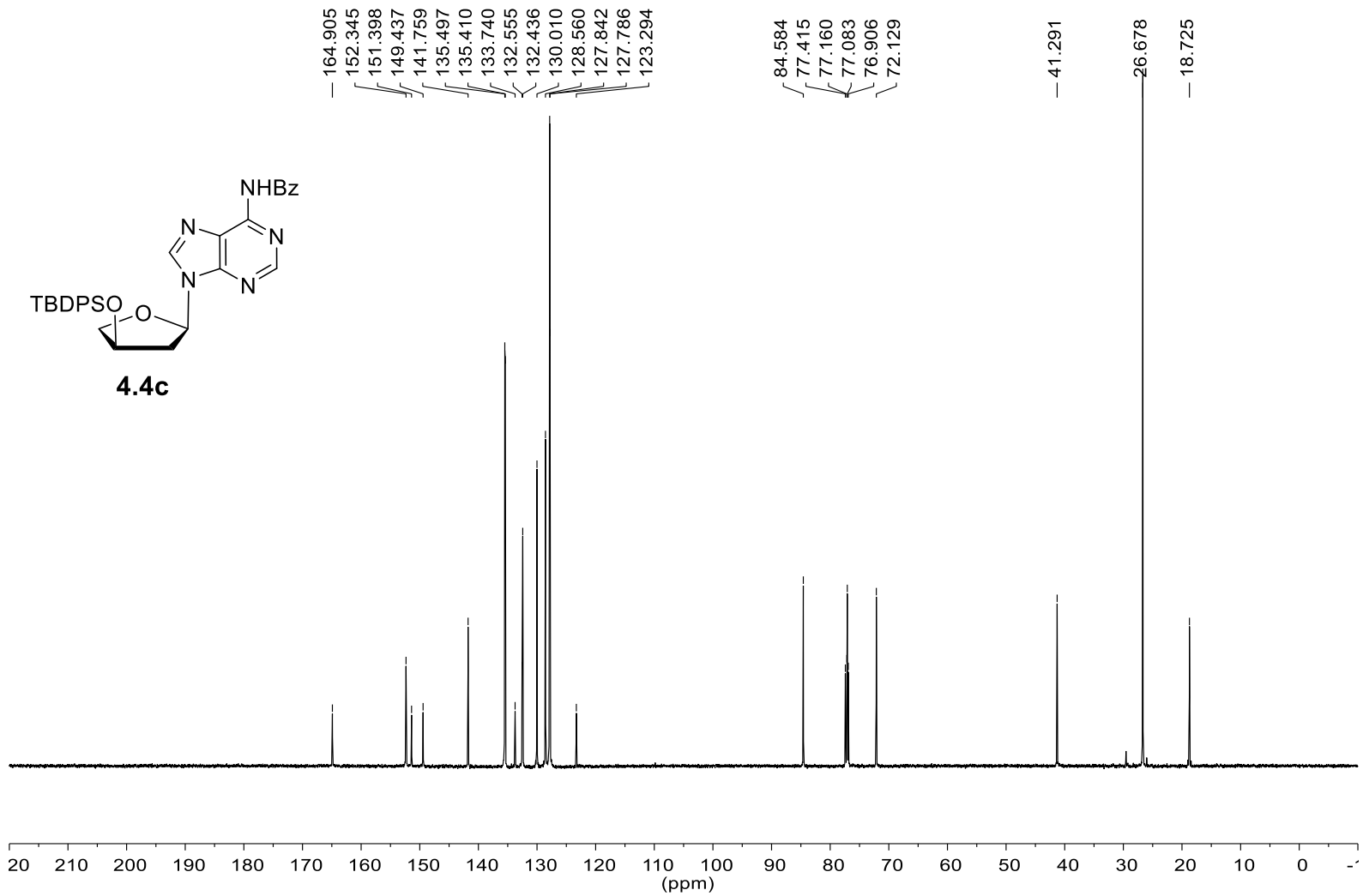
^{1}H NMR spectrum of compound **4.4b** (400 MHz, $CDCl_3$)



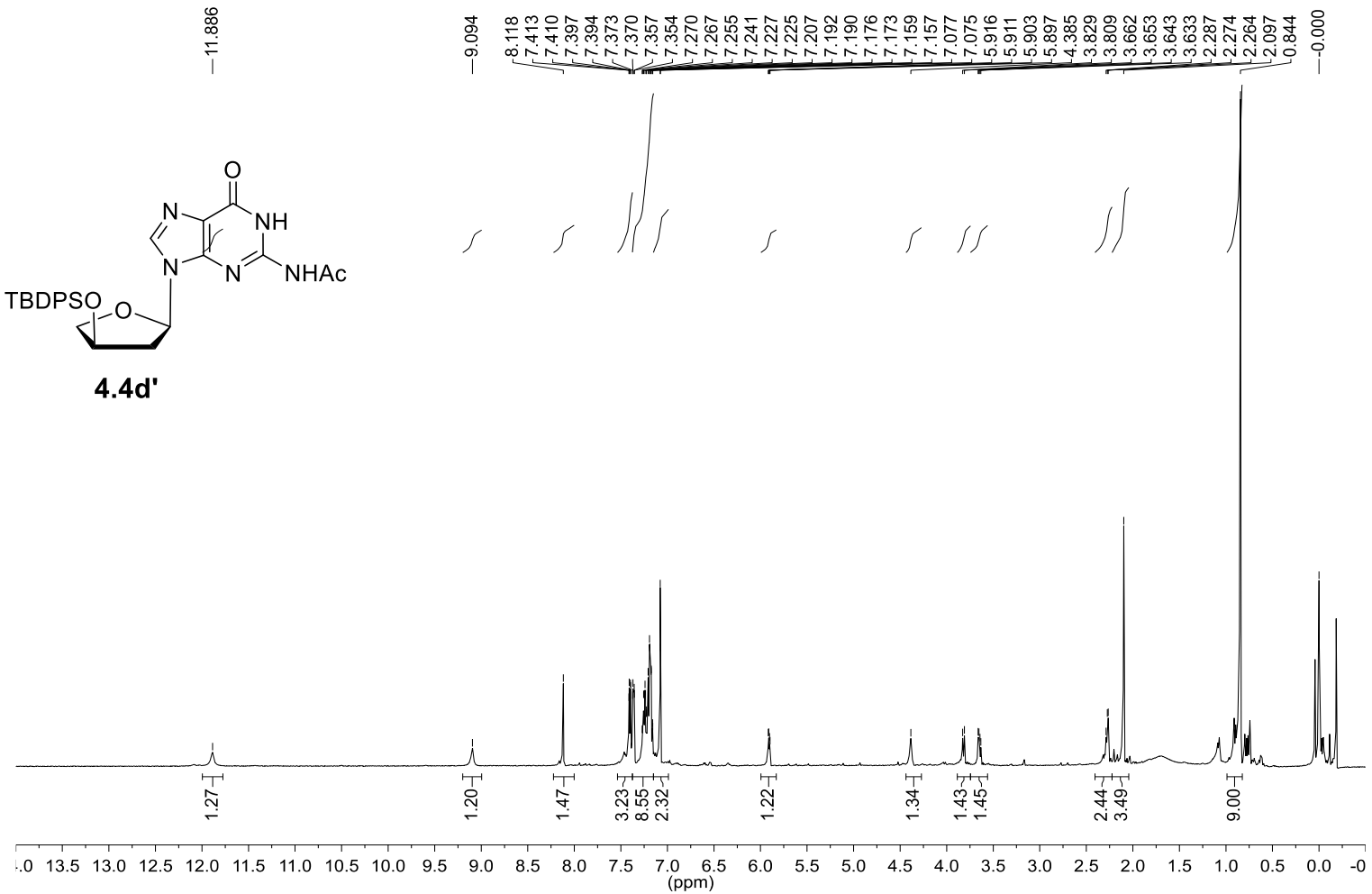
^{13}C NMR spectrum of compound **4.4b** (125.8 MHz, $CDCl_3$)



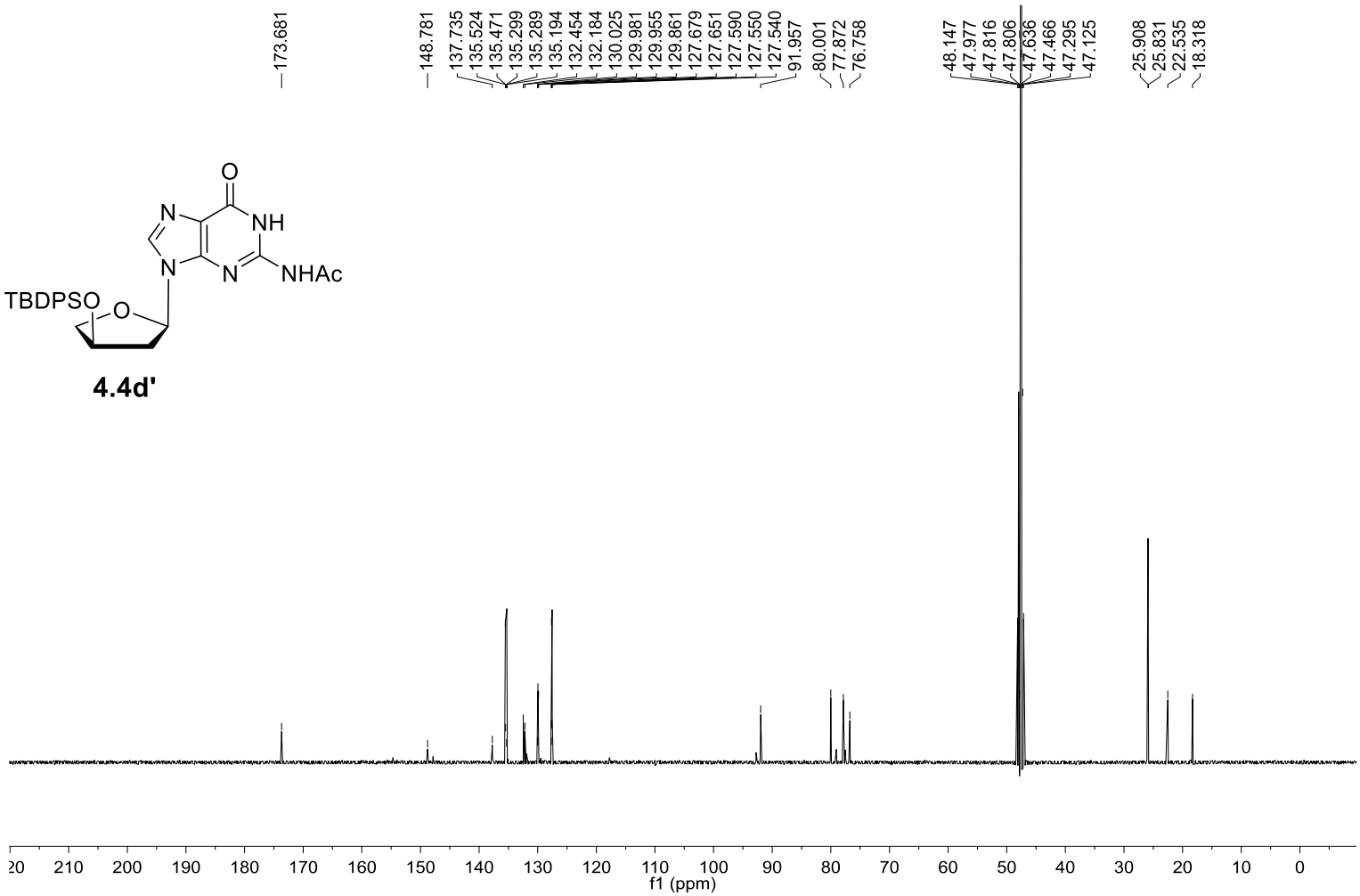
^{1}H NMR spectrum of compound **4.4c** (400 MHz, $CDCl_3$)



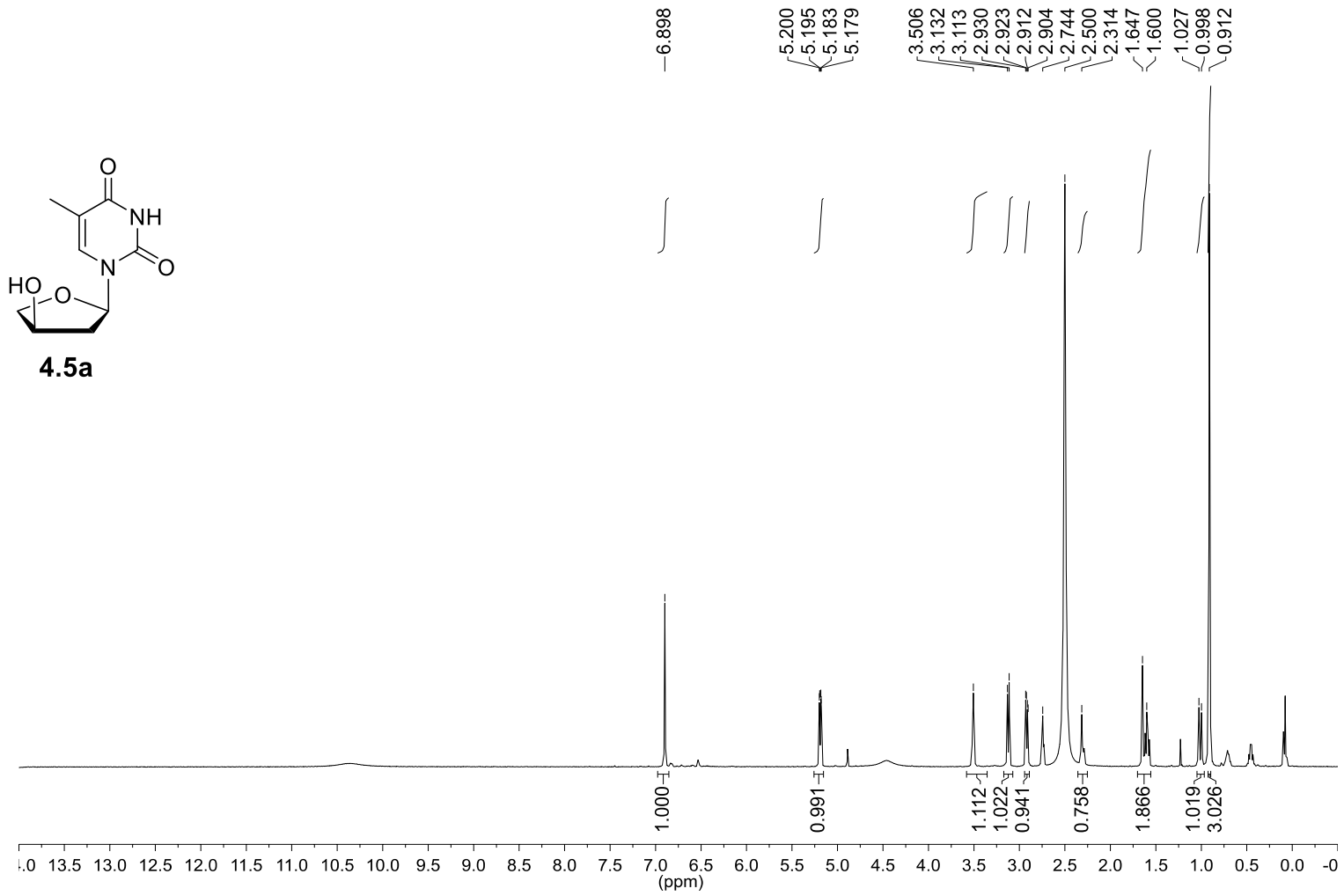
^{13}C NMR spectrum of compound **4.4c** (125.8 MHz, $CDCl_3$)



^{1}H NMR spectrum of compound **4.4d'** (400 MHz, CD_3OD)

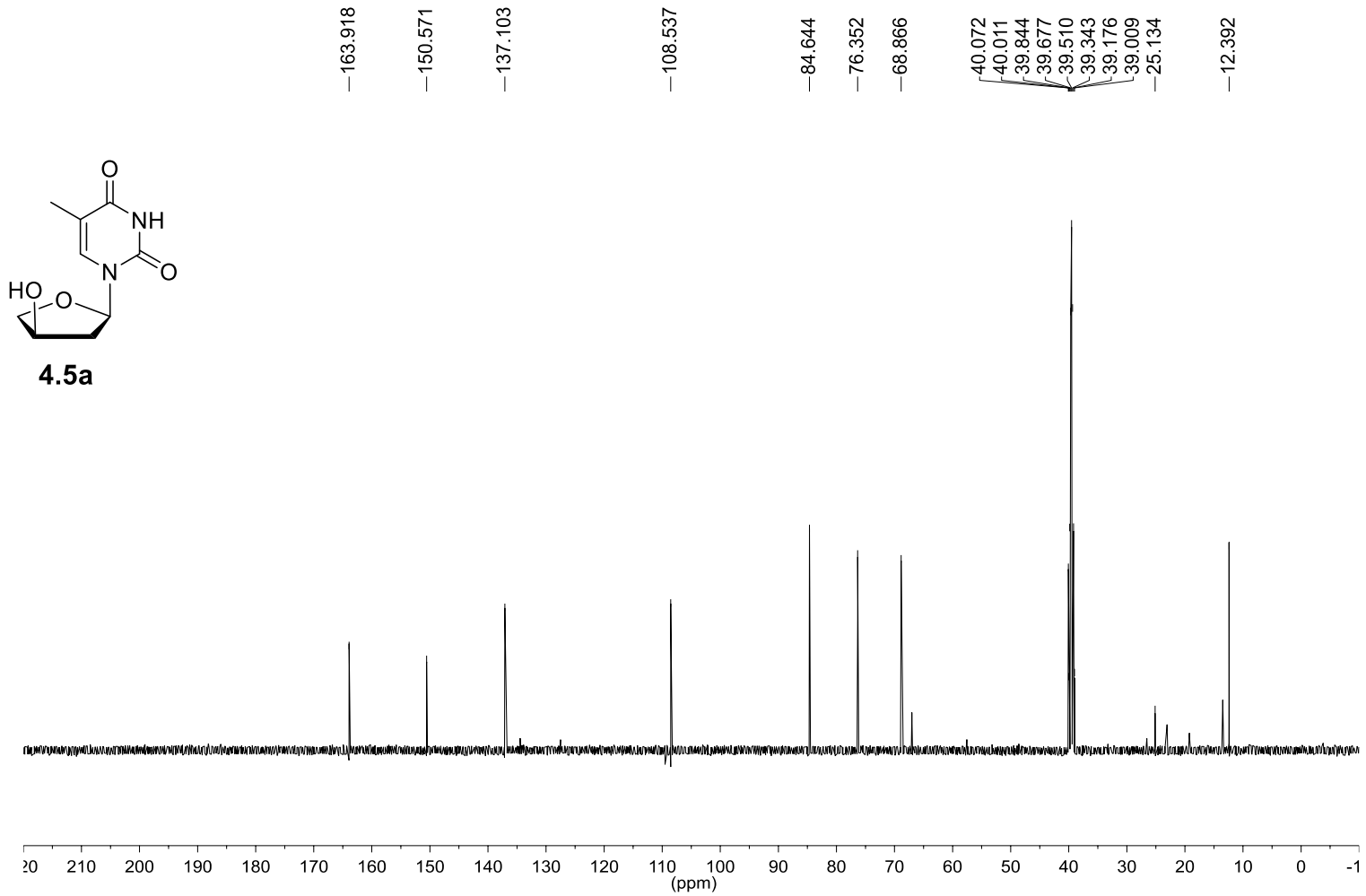


^{13}C NMR spectrum of compound **4.4d'** (125.8 MHz, CD_3OD)

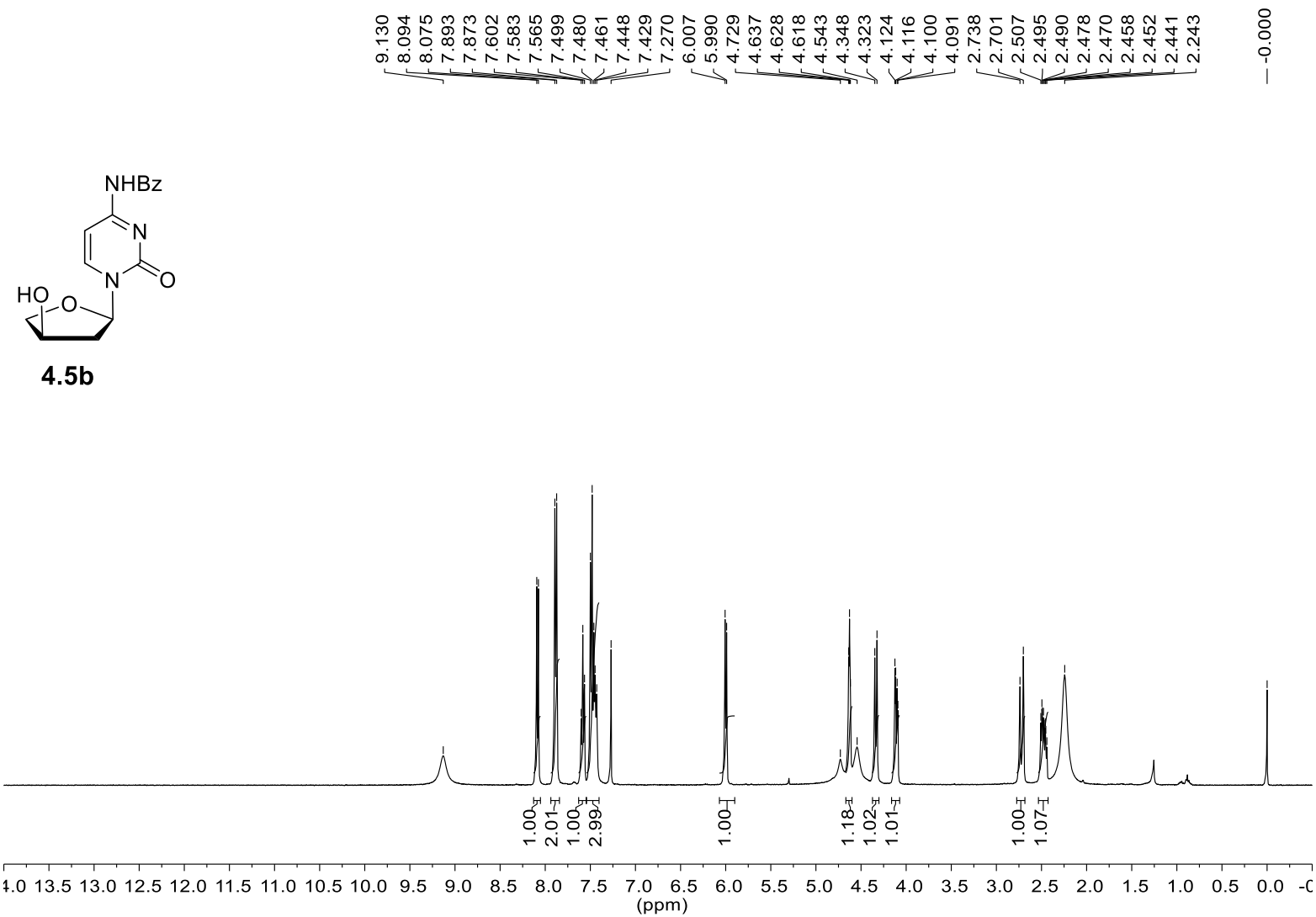


4.5a

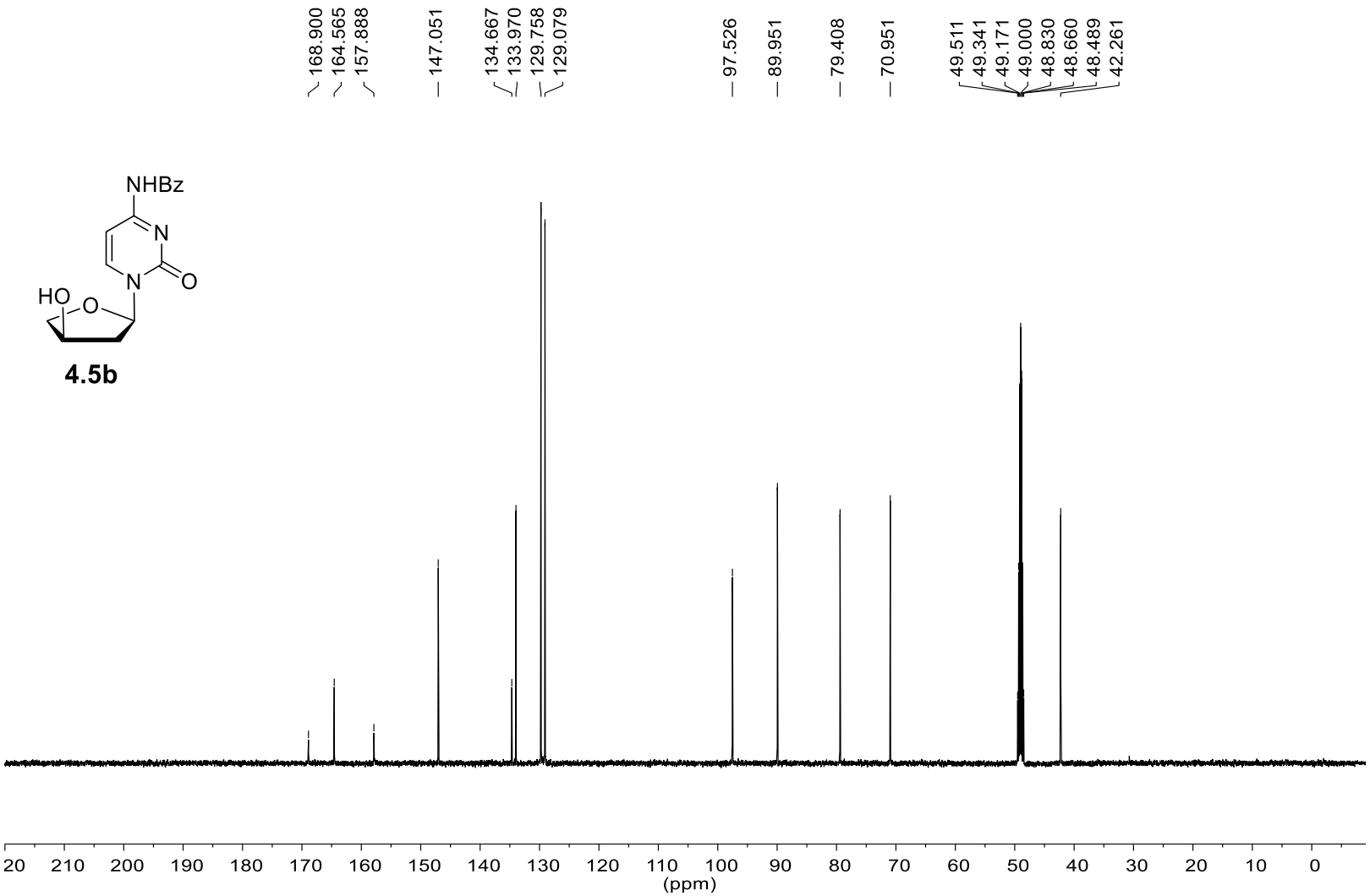
^{1}H NMR spectrum of compound **4.5a** (400 MHz, DMSO-d_6)



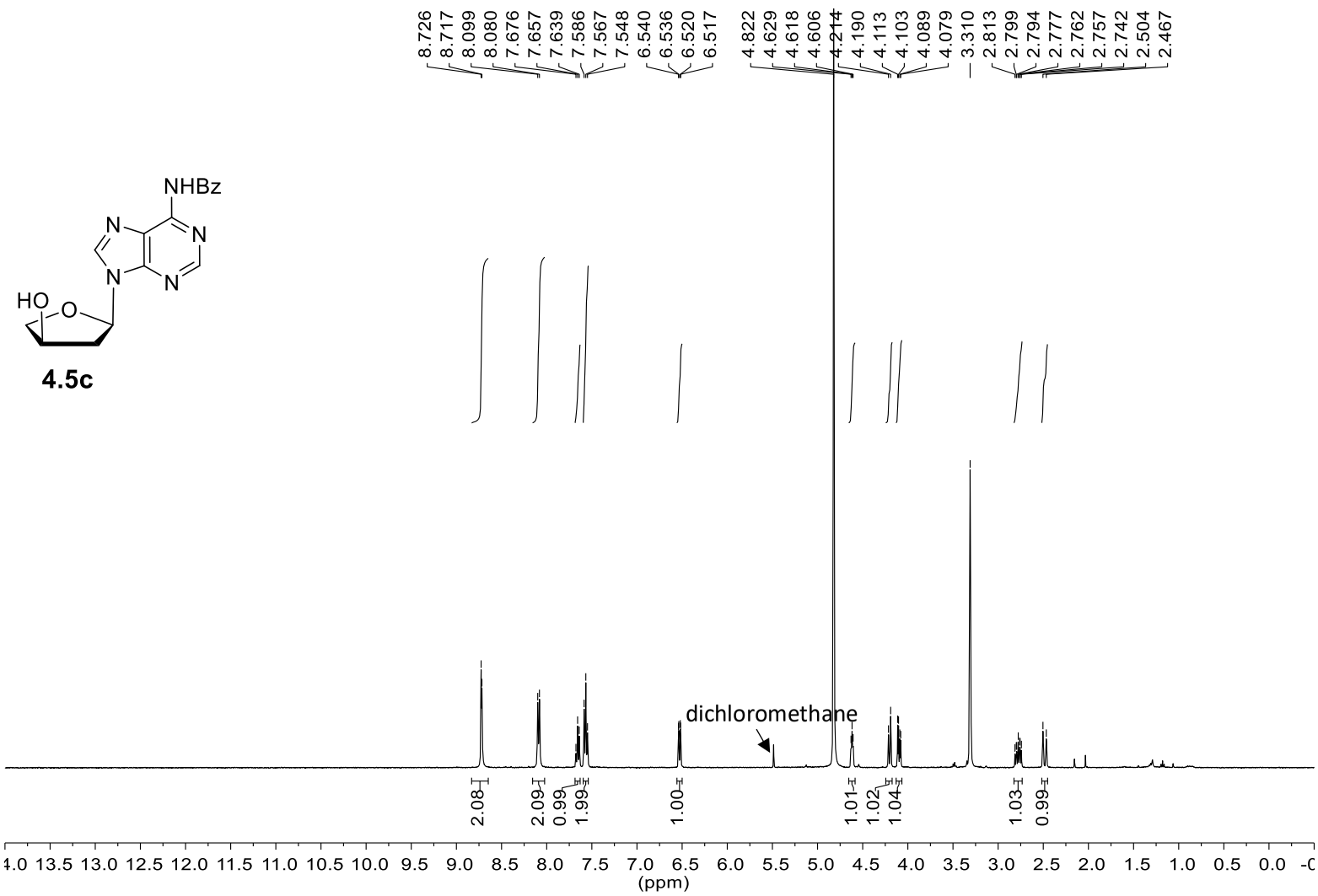
^{13}C NMR spectrum of compound **4.5a** (125.8 MHz, DMSO-d_6)



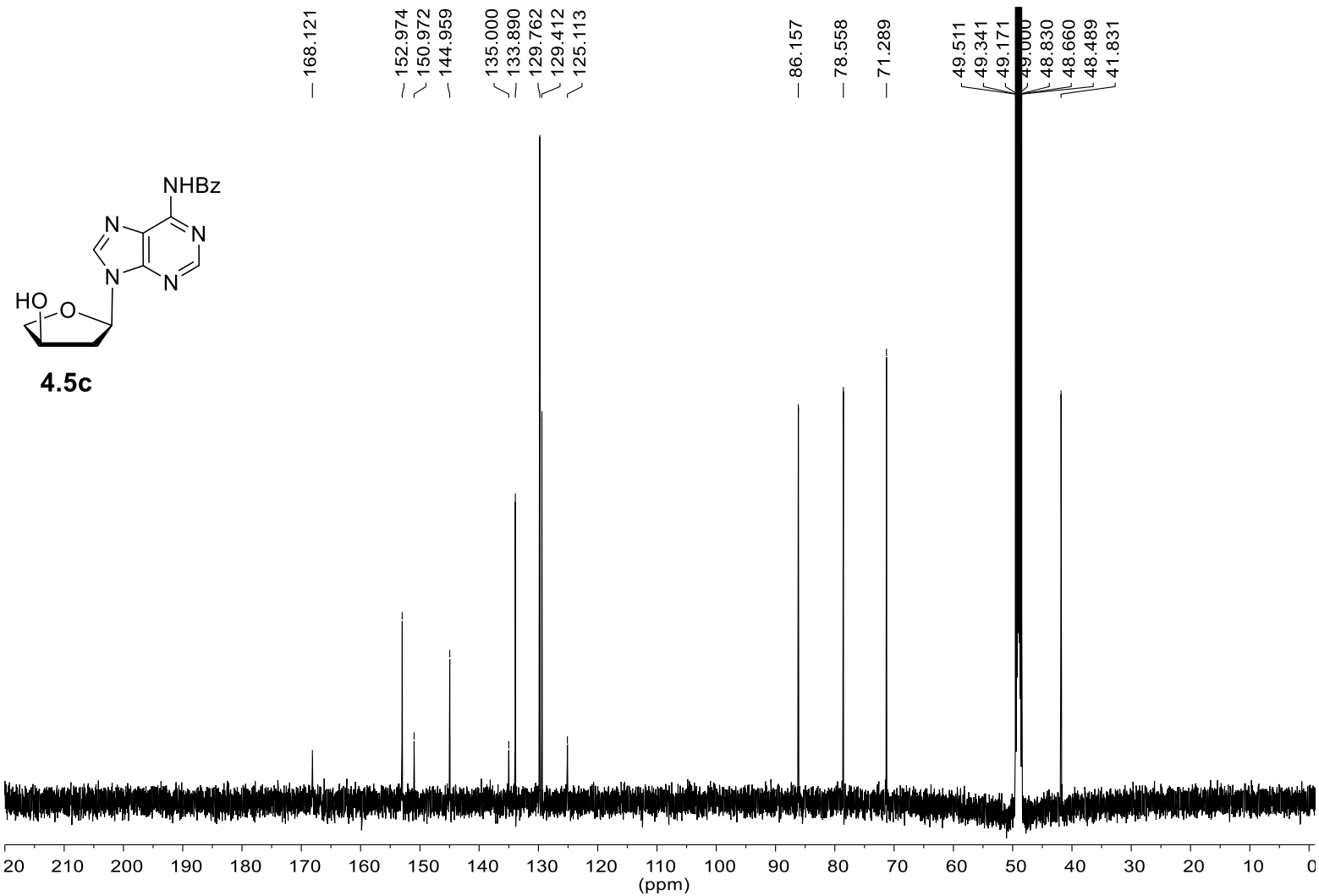
^{1}H NMR spectrum of compound **4.5b** (400 MHz, $CDCl_3$)



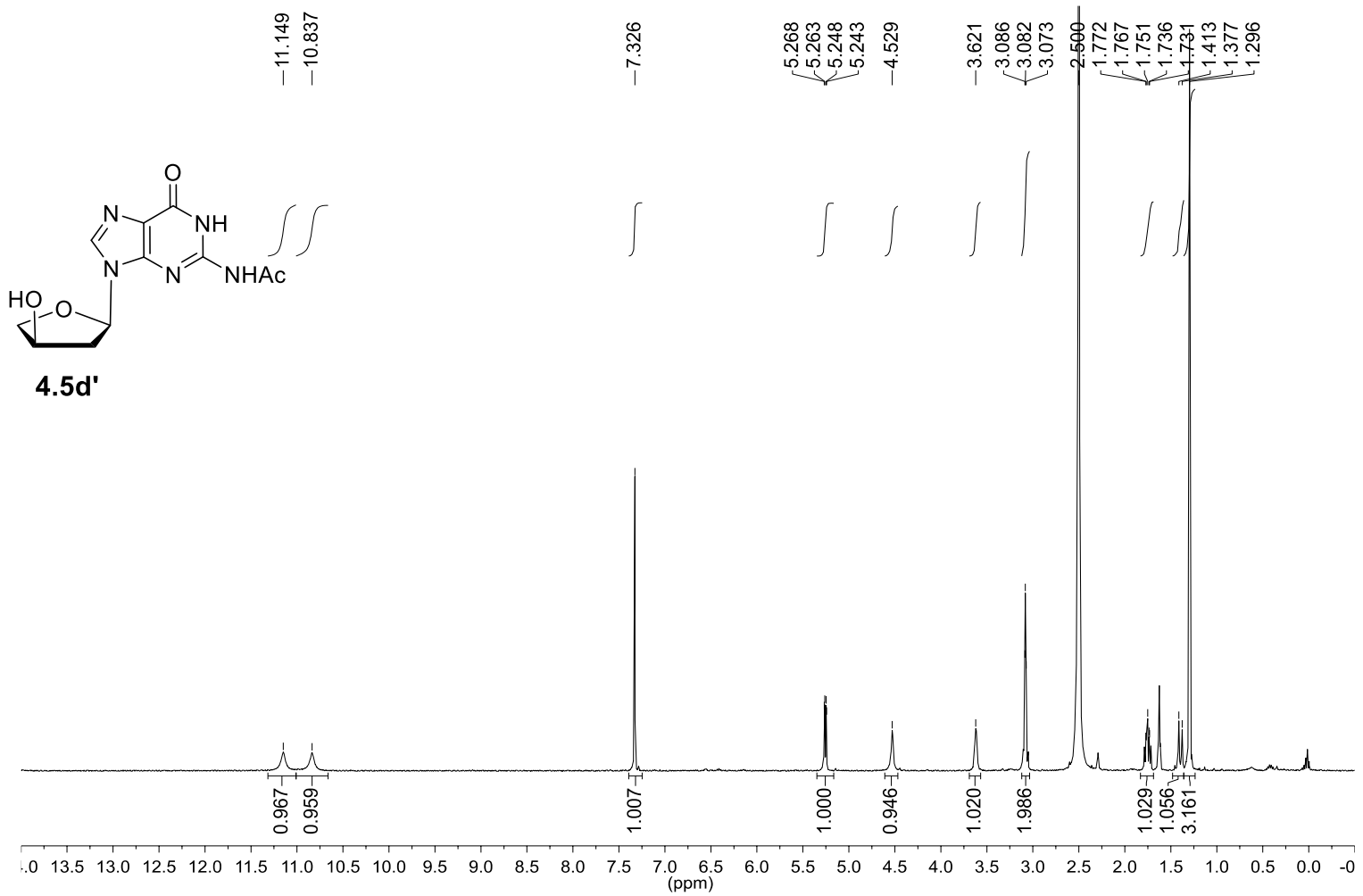
^{13}C NMR spectrum of compound **4.5b** (100.6 MHz, CD_3OD)



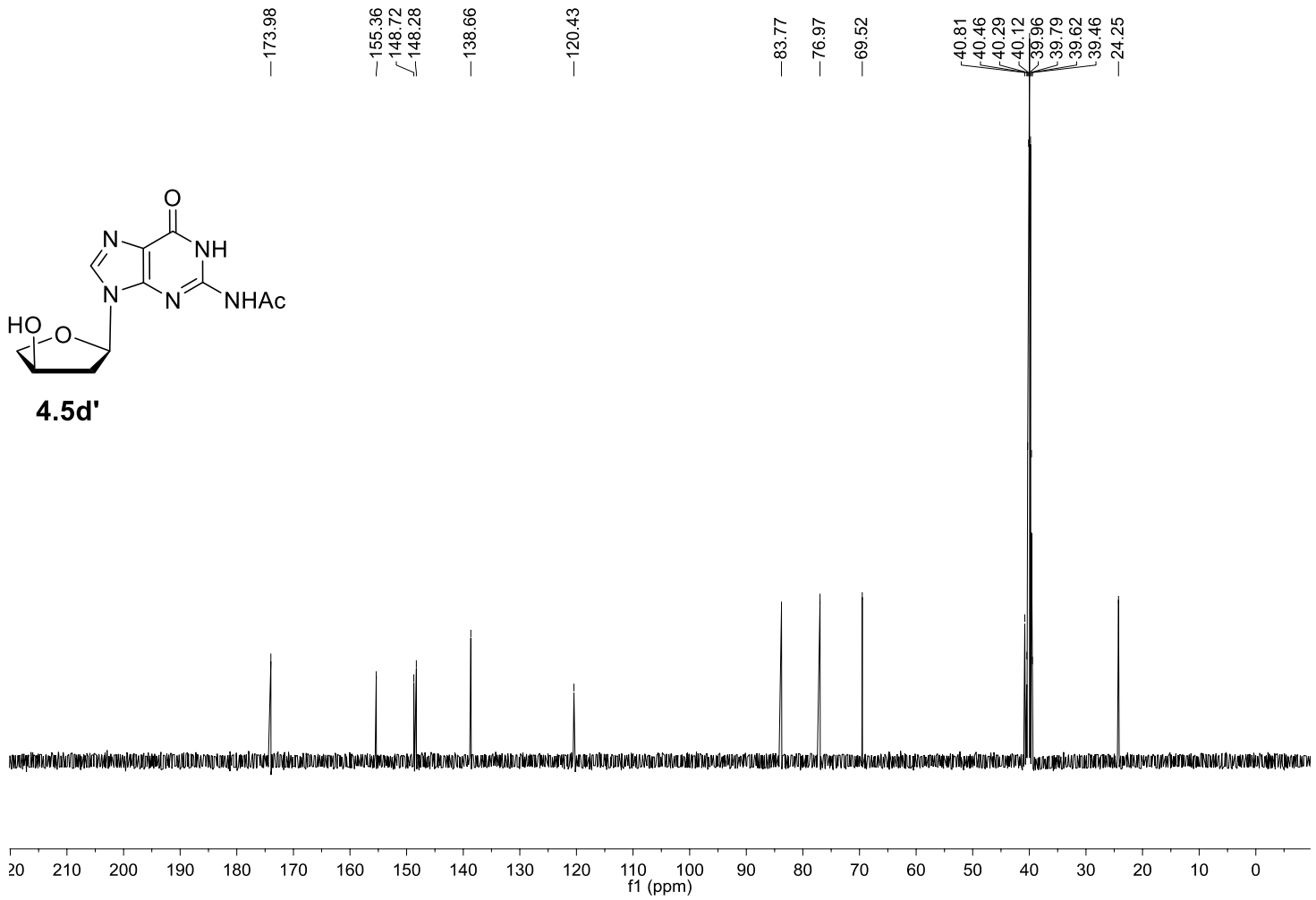
^{1}H NMR spectrum of compound **4.5c** (400 MHz, MeOH-d_4)



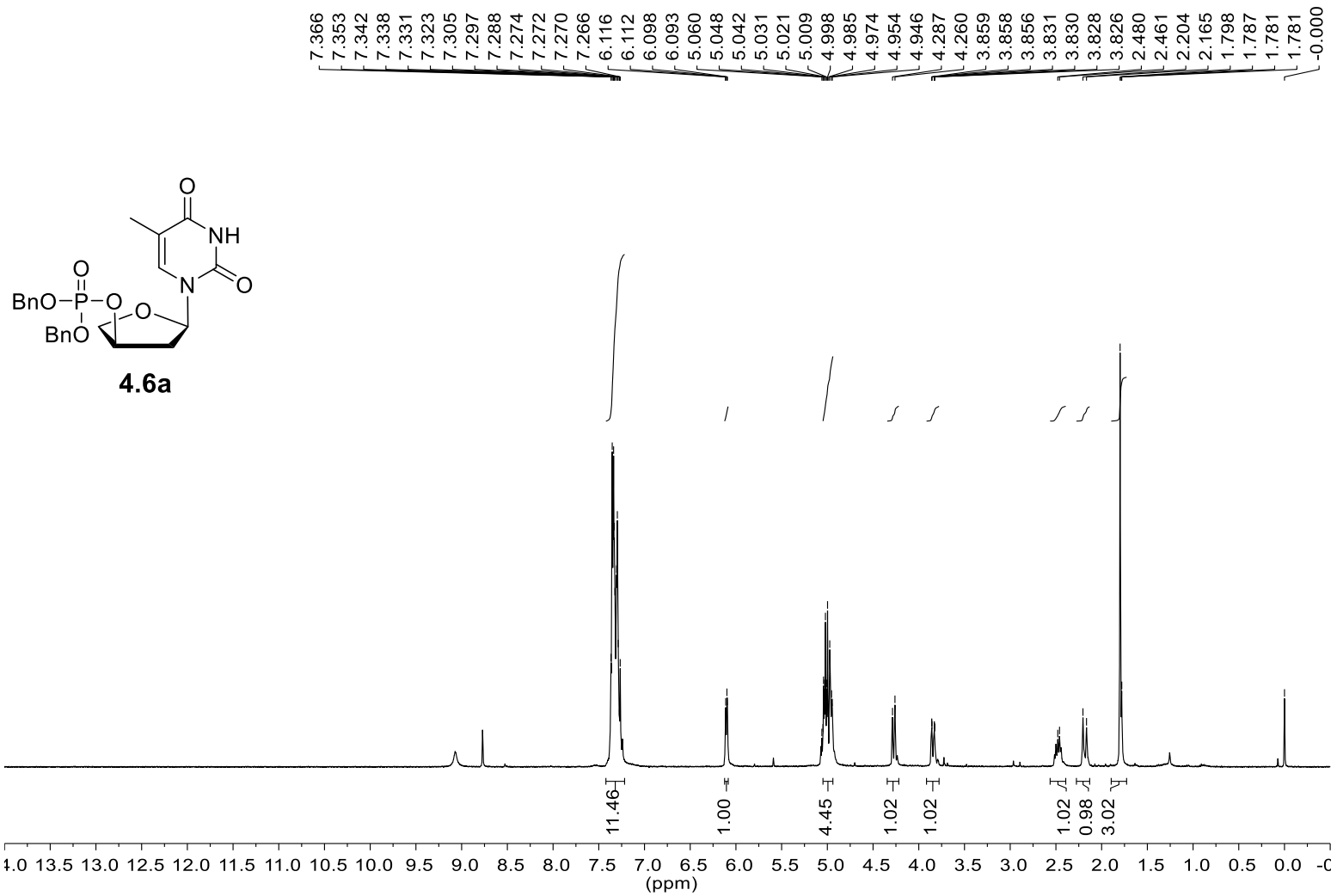
^{13}C NMR spectrum of compound **4.5c** (125.8 MHz, MeOH-d_4)



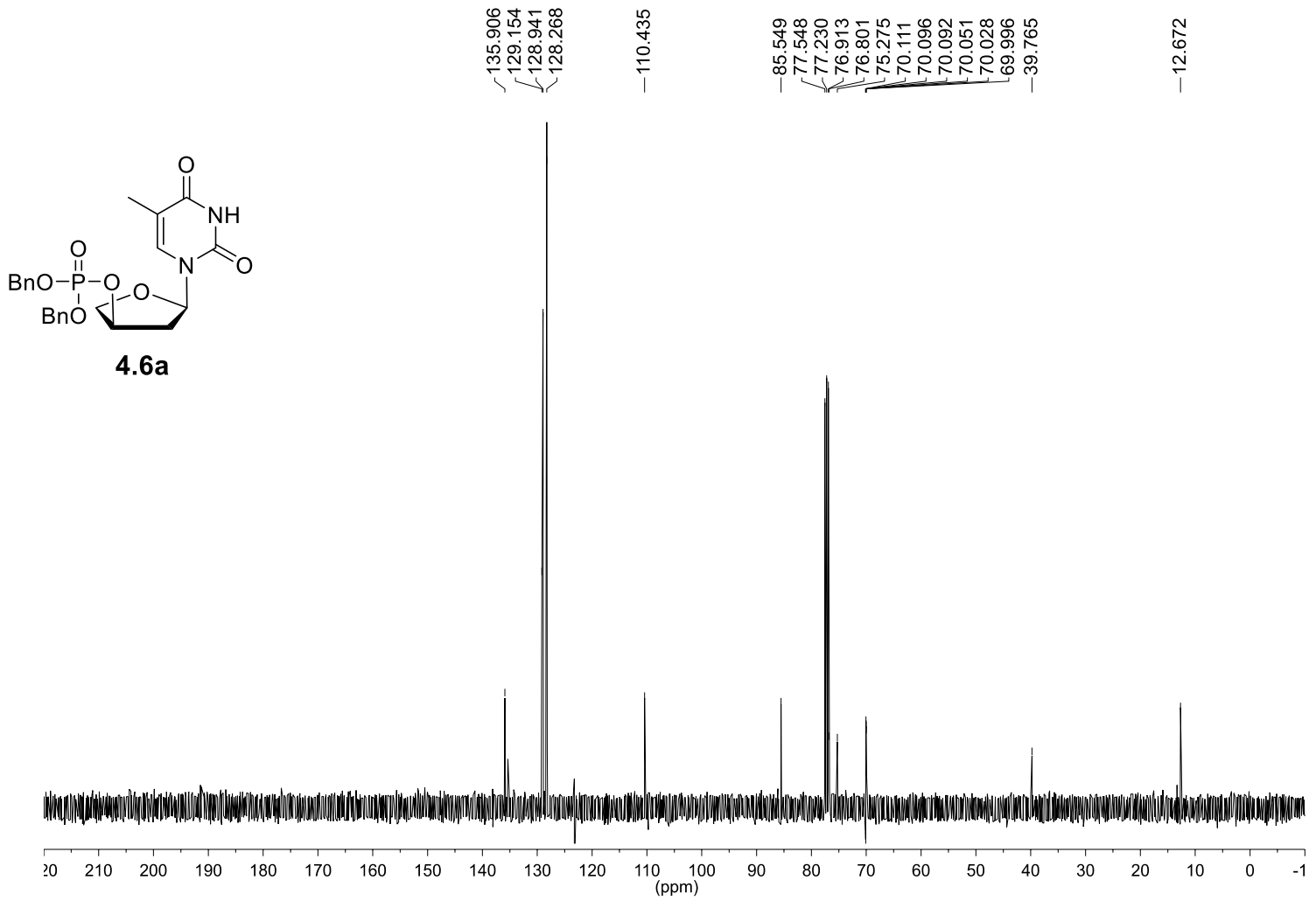
^{1}H NMR spectrum of compound **4.5d'** (400 MHz, DMSO-d_6)



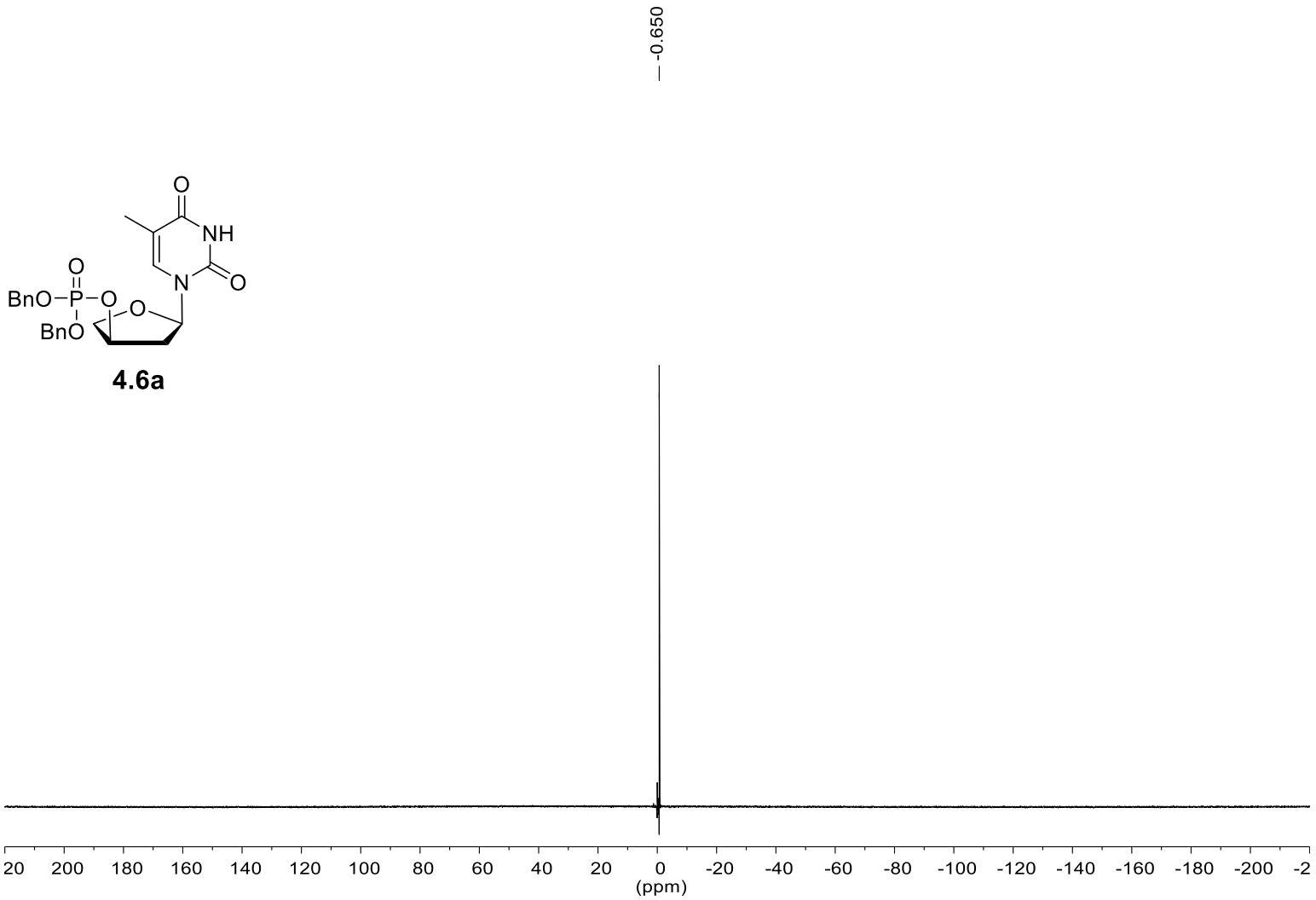
^{13}C NMR spectrum of compound **4.5d'** (125.8 MHz, DMSO-d_6)



^{1}H NMR spectrum of compound **4.6a** (400 MHz, $CDCl_3$)

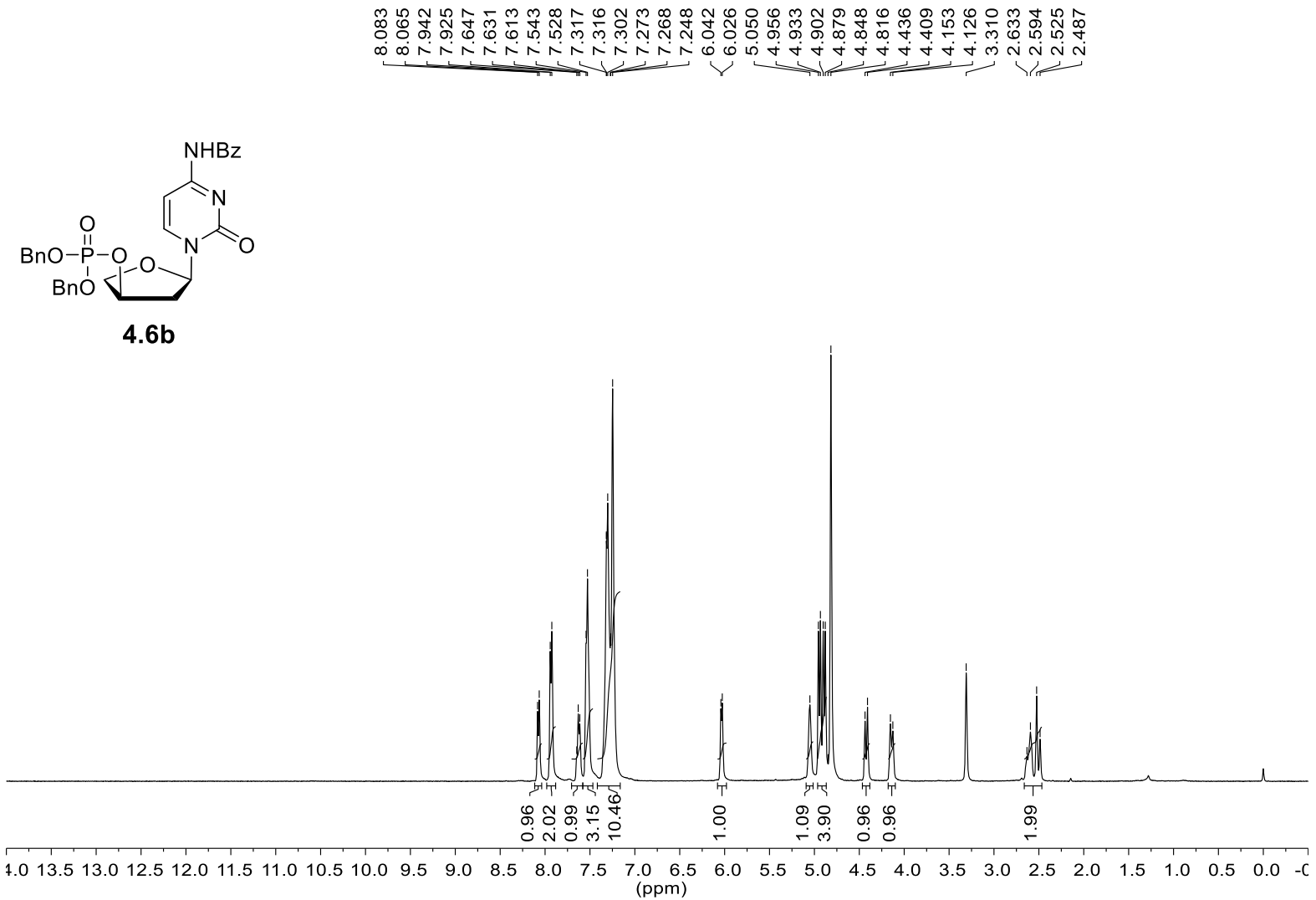


^{13}C NMR spectrum of compound **4.6a** (125.8 MHz, $CDCl_3$)

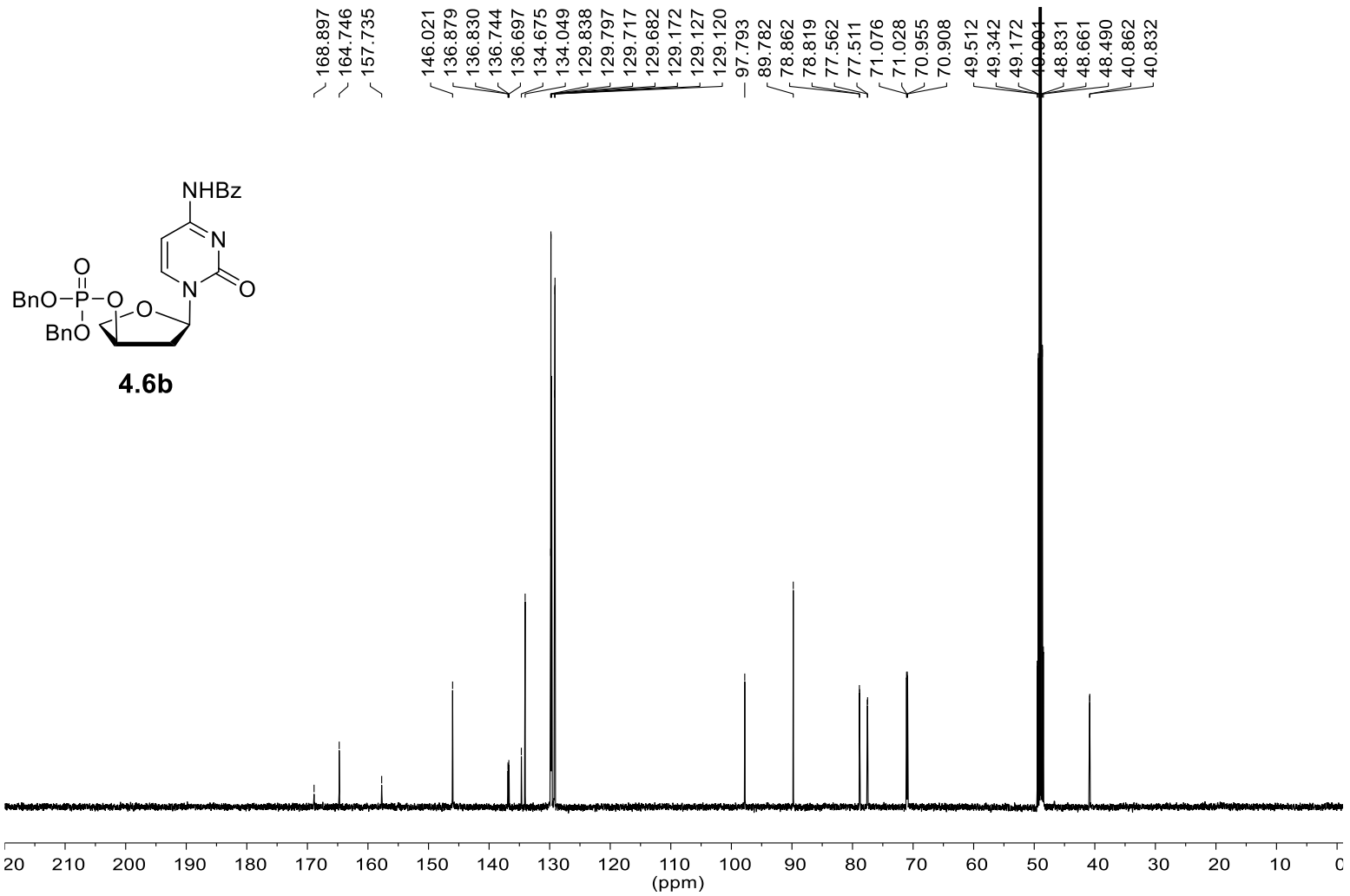


4.6a

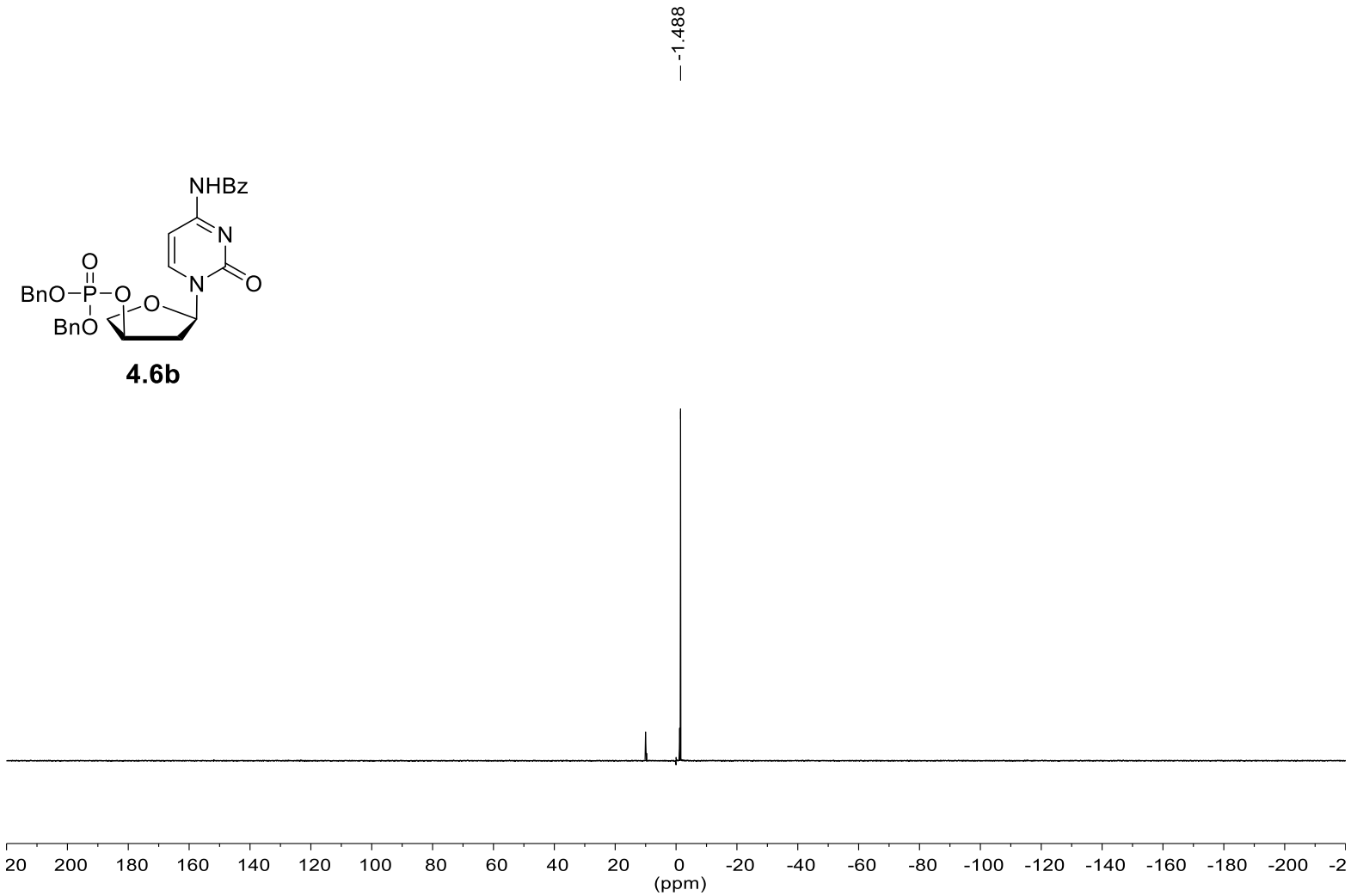
^{31}P NMR spectrum of compound **4.6a** (162 MHz, $CDCl_3$)



^{1}H NMR spectrum of compound **4.6b** (400 MHz, CD_3OD)

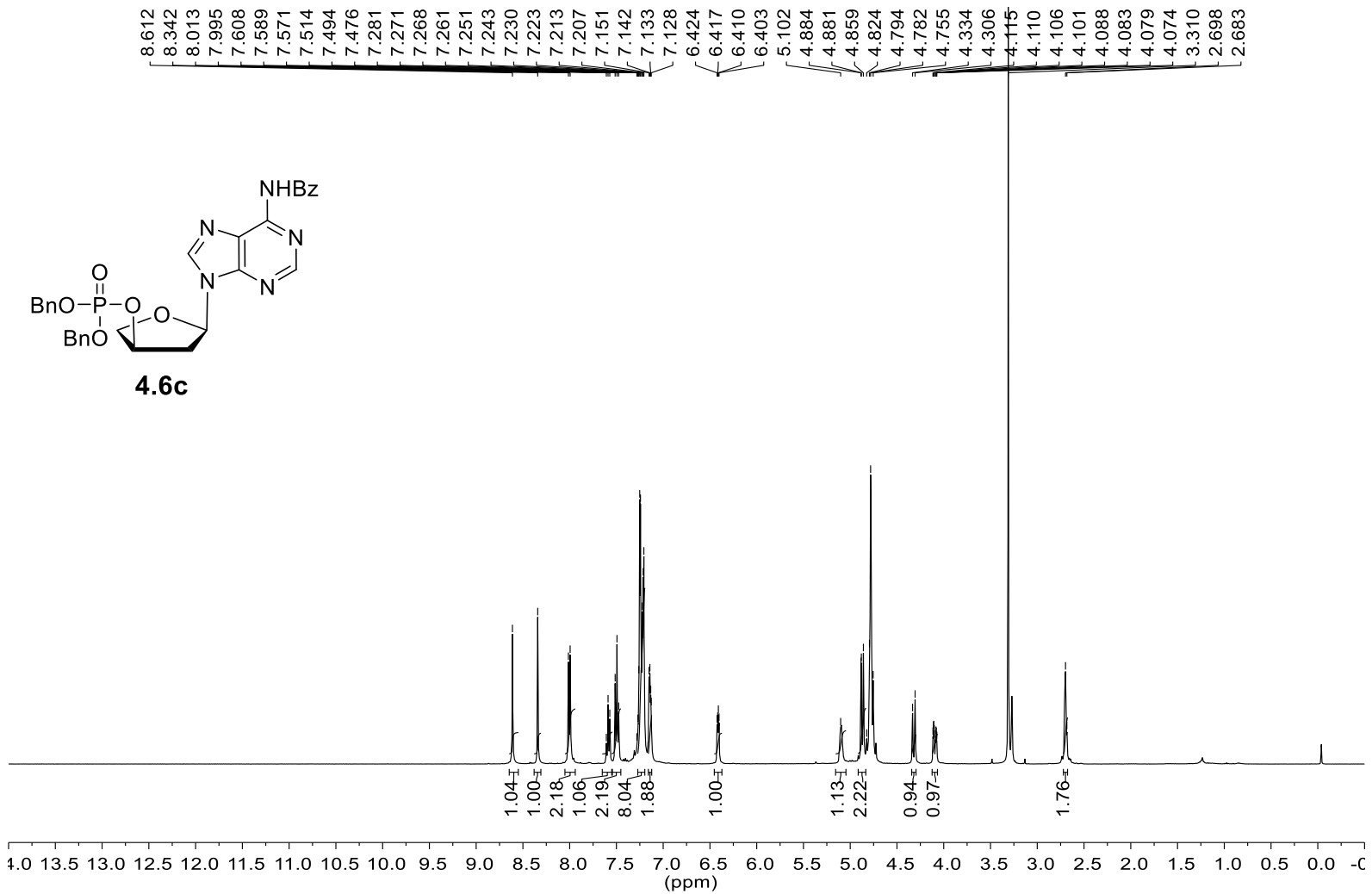


^{13}C NMR spectrum of compound **4.6b** (125.8 MHz, CD_3OD)

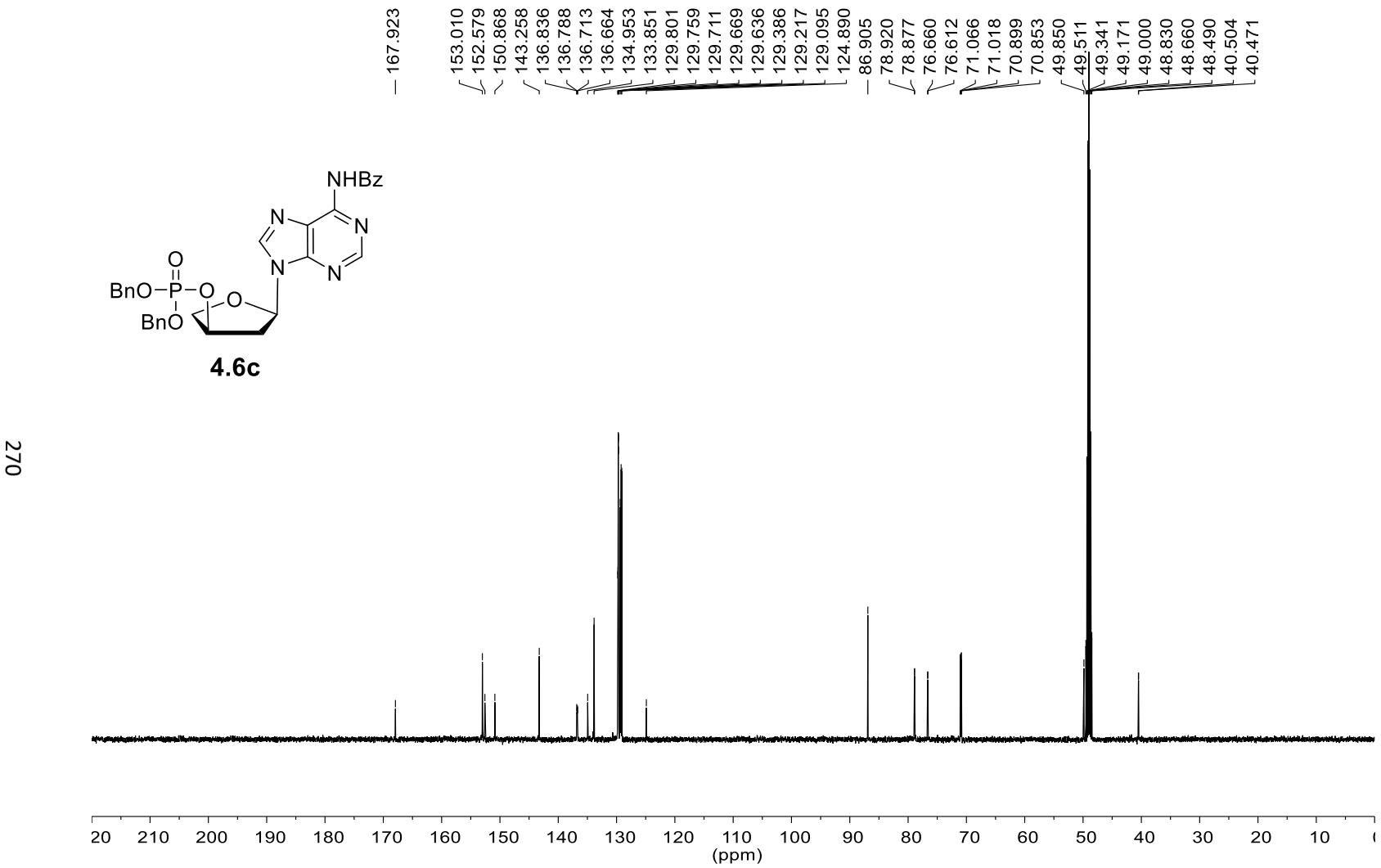


4.6b

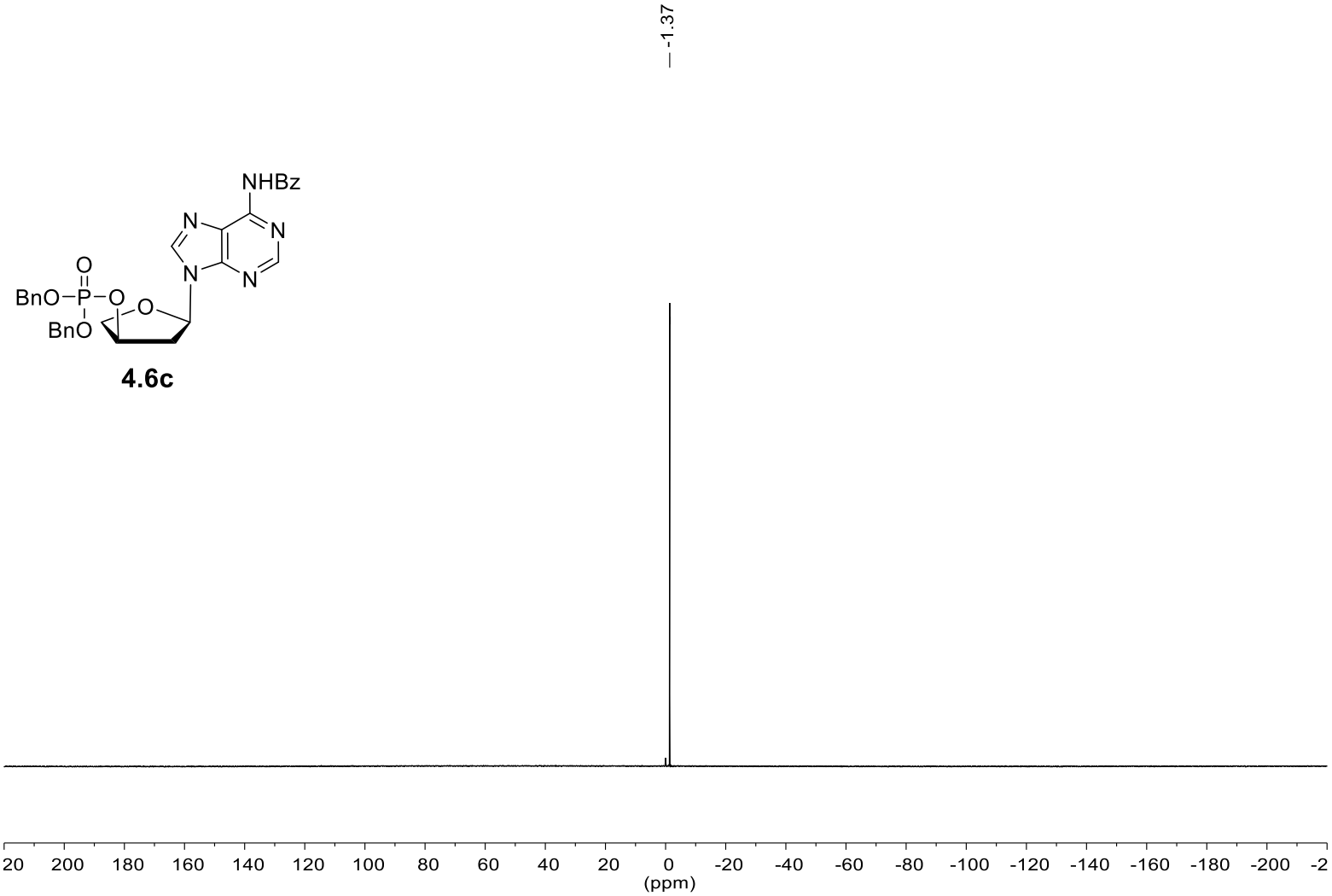
^{31}P NMR spectrum of compound **4.6b** (162 MHz, CD_3OD)



^{1}H NMR spectrum of compound **4.6c** (400 MHz, CD_3OD)

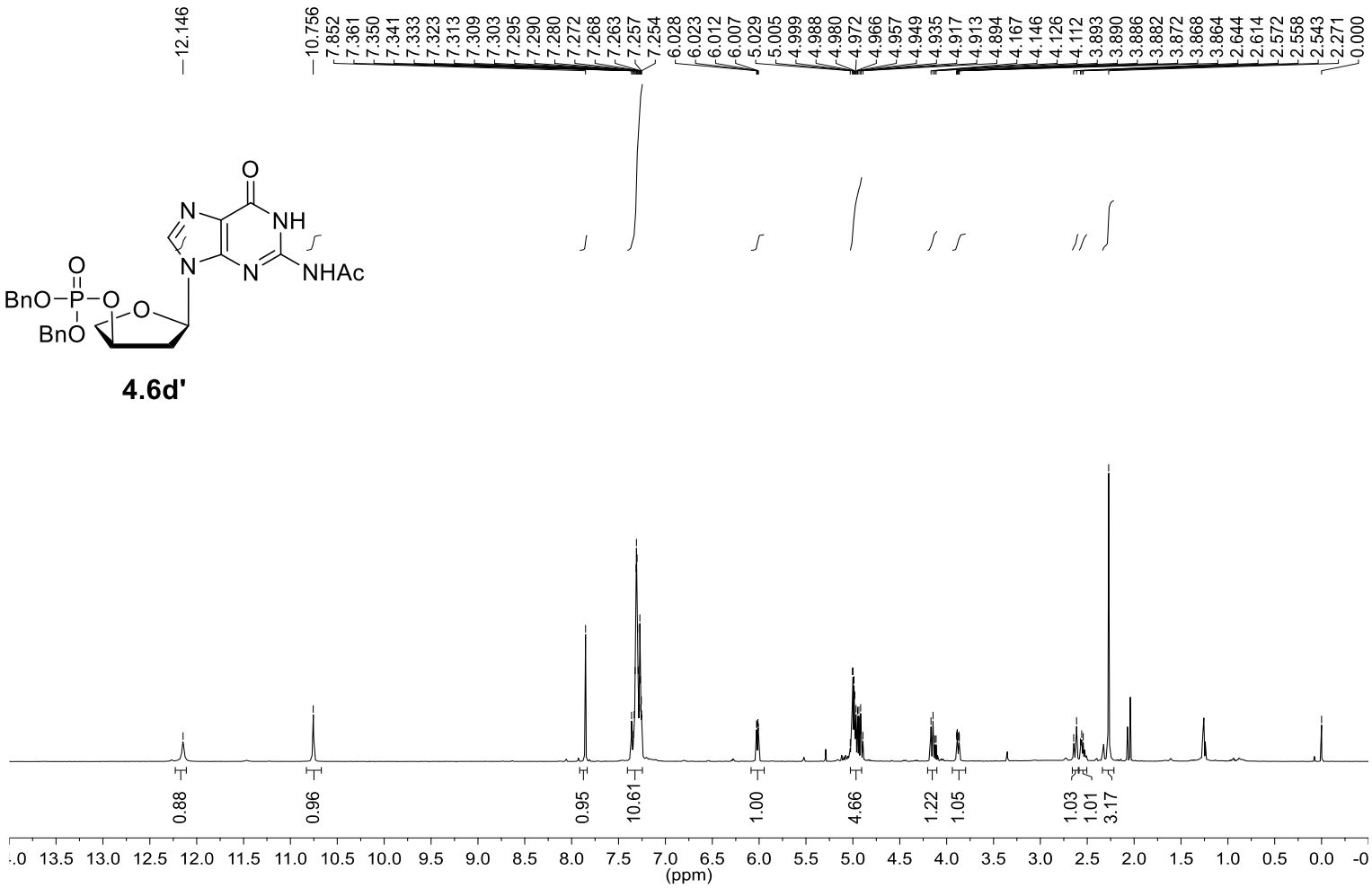


^{13}C NMR spectrum of compound **4.6c** (125.8 MHz, CD_3OD)



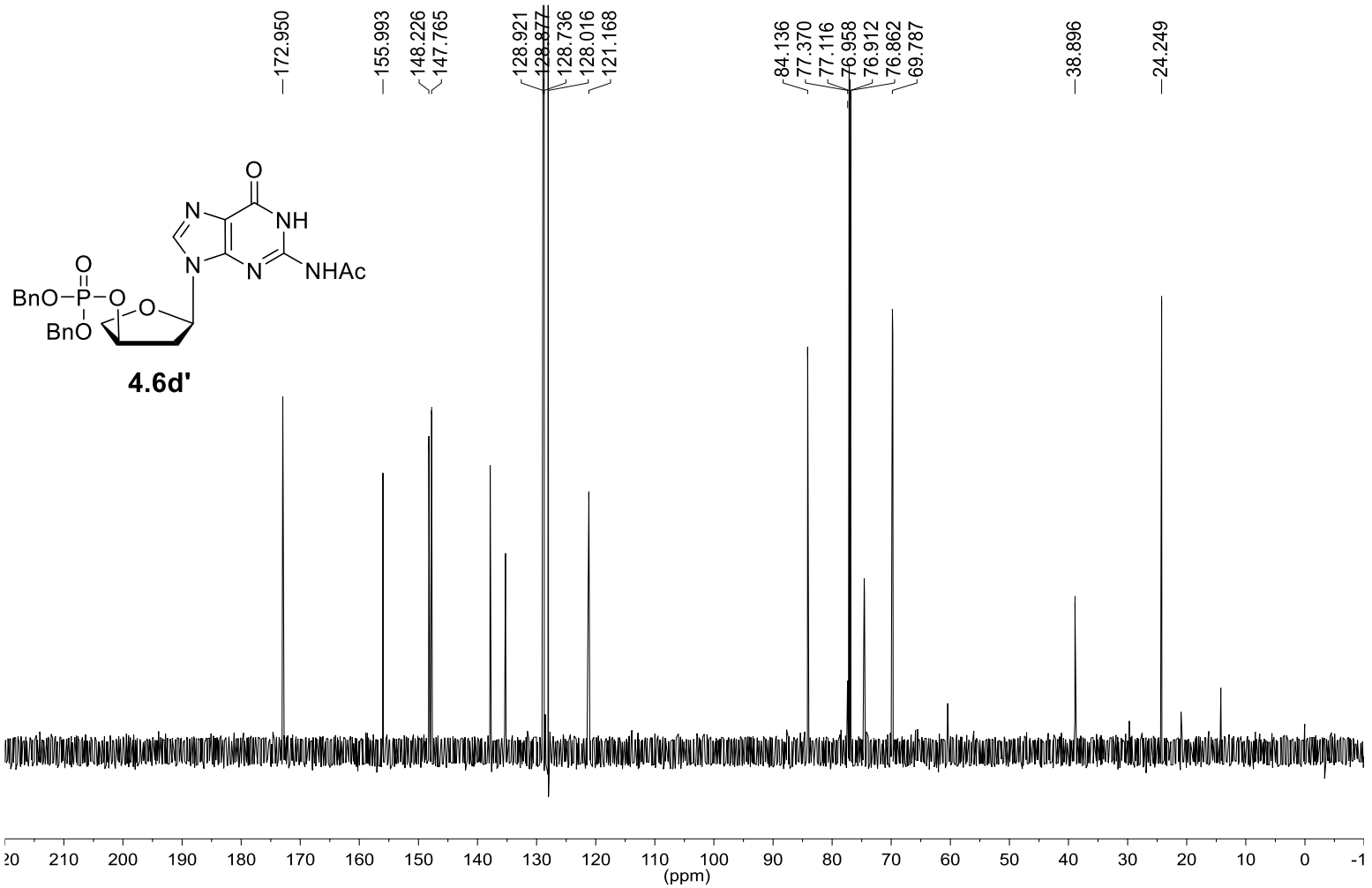
4.6c

^{31}P NMR spectrum of compound **4.6c** (162 MHz, CD_3OD)

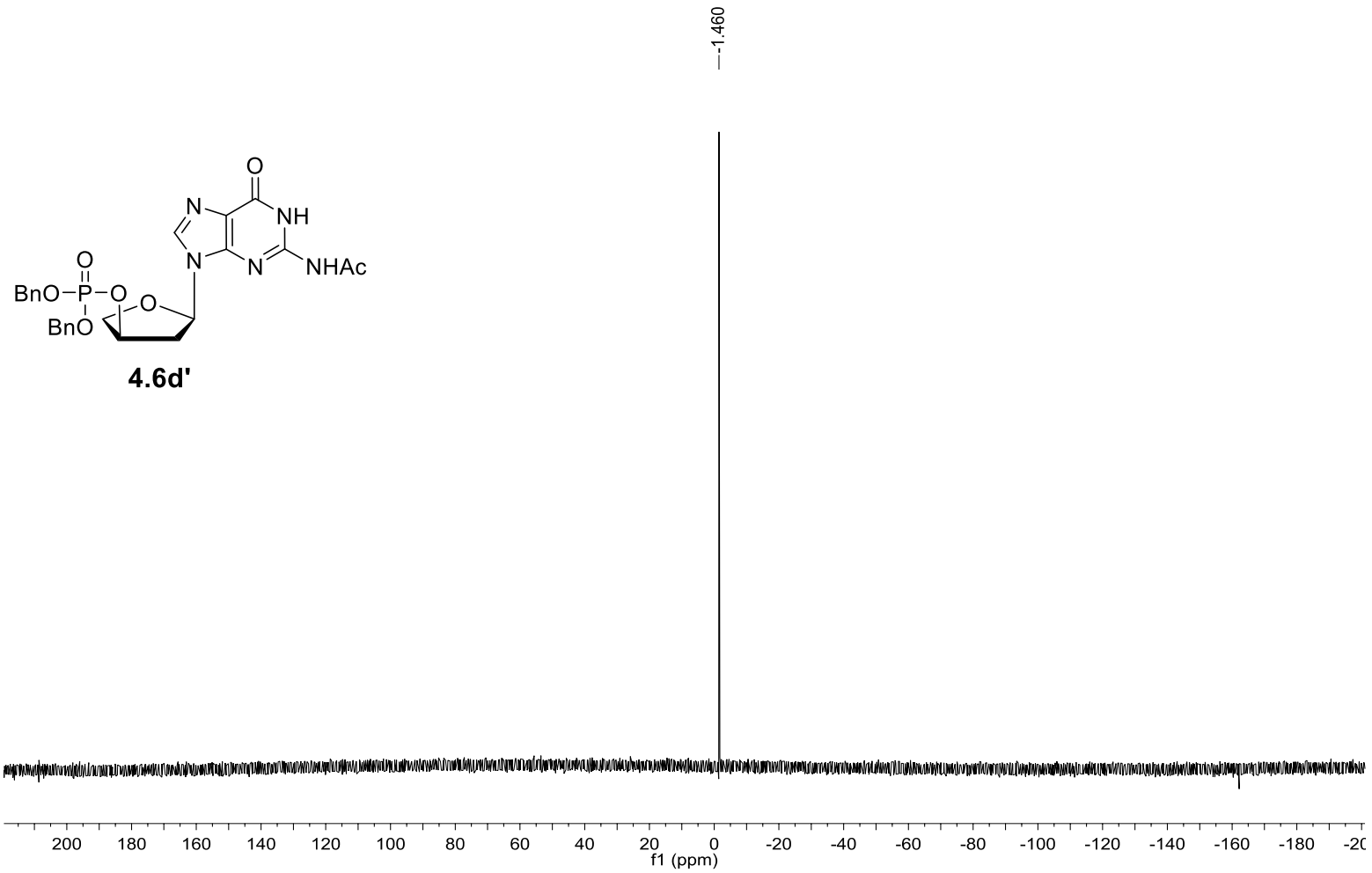


4.6d'

^{1}H NMR spectrum of compound **4.6d'** (400 MHz, $CDCl_3$)

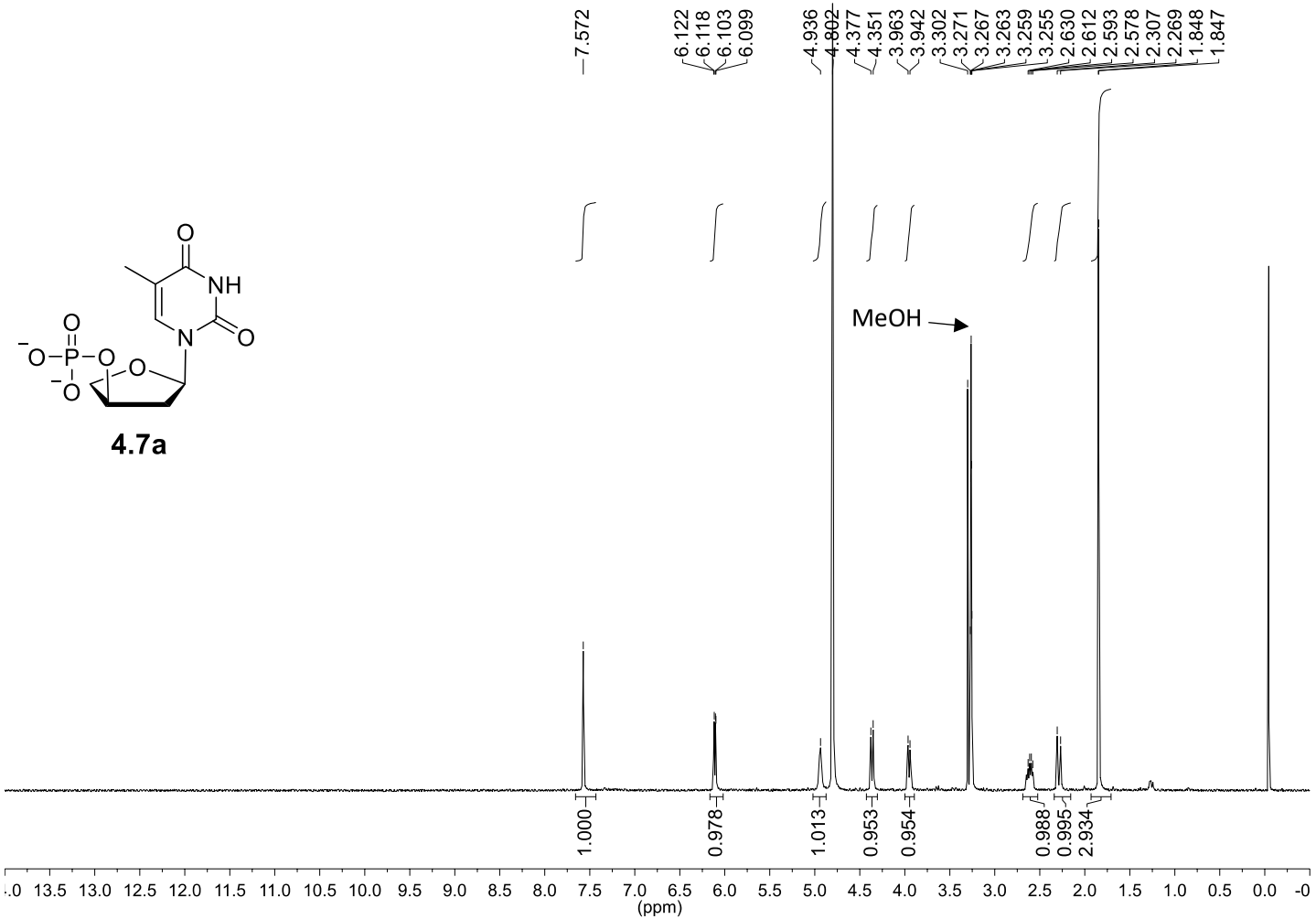


^{13}C NMR spectrum of compound **4.6d'** (125.8 MHz, $CDCl_3$)



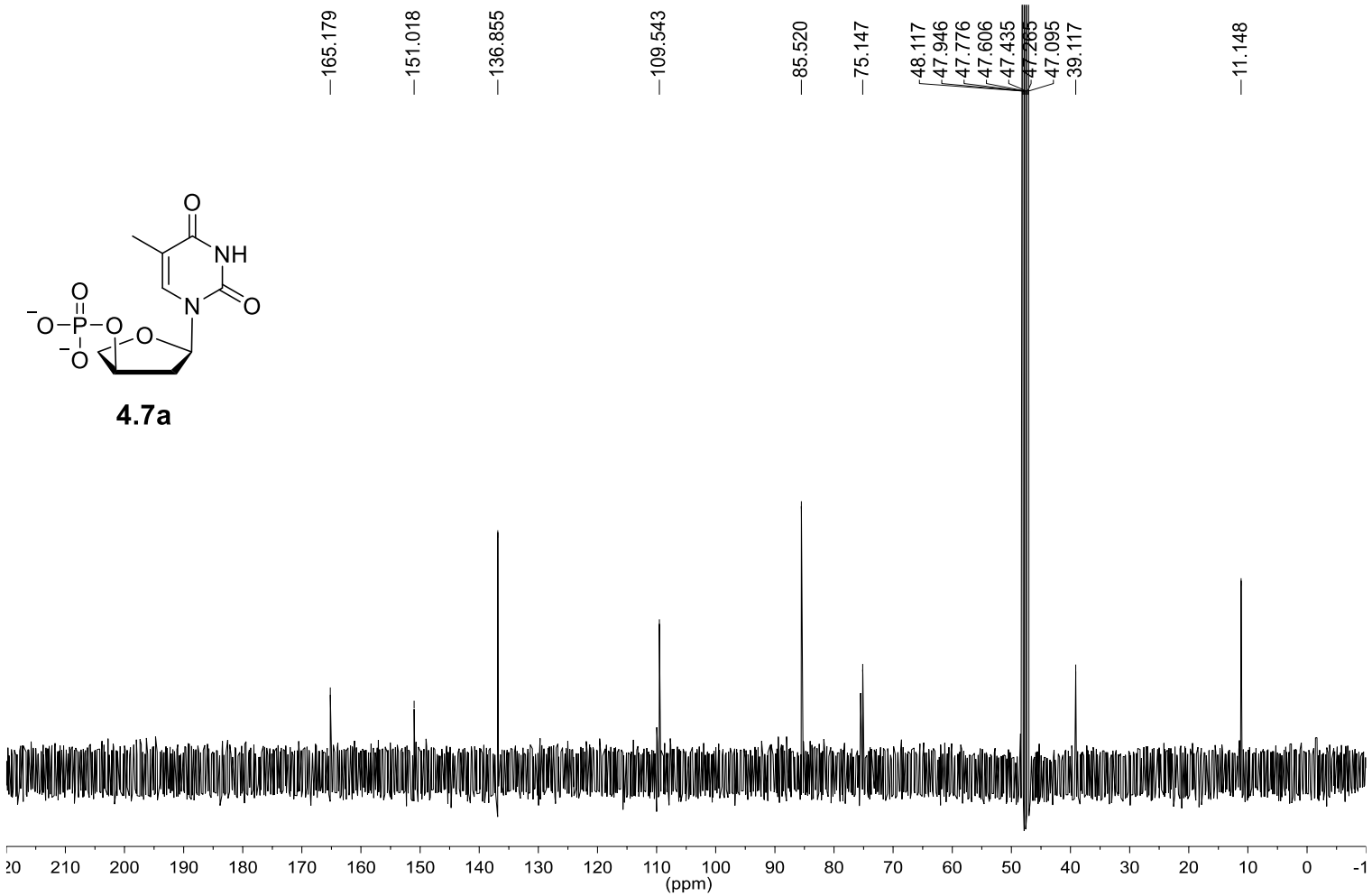
4.6d'

^{31}P NMR spectrum of compound **4.6d'** (162 MHz, $CDCl_3$)

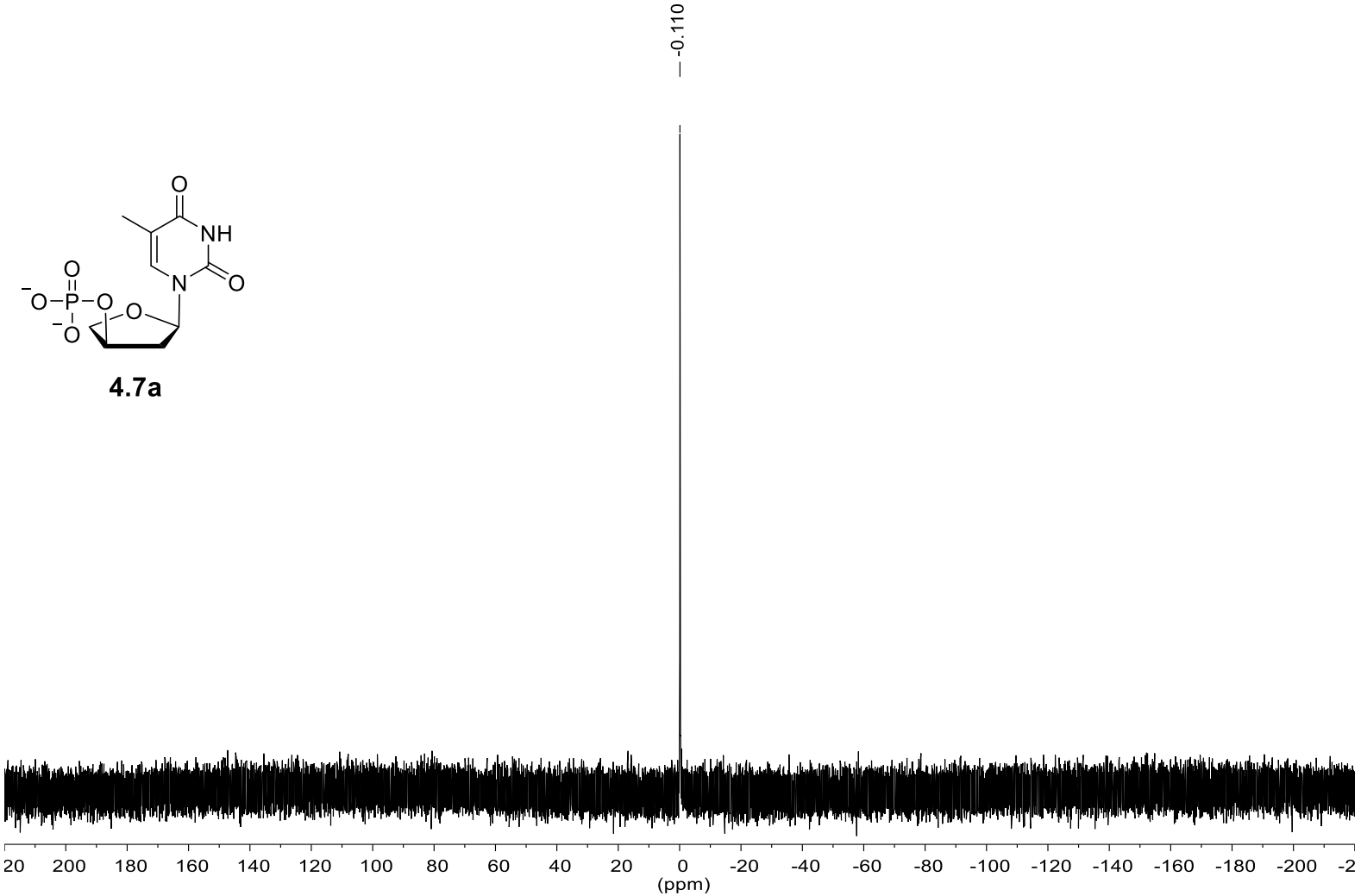


4.7a

^{1}H NMR spectrum of compound **4.7a** (400 MHz, CD_3OD)

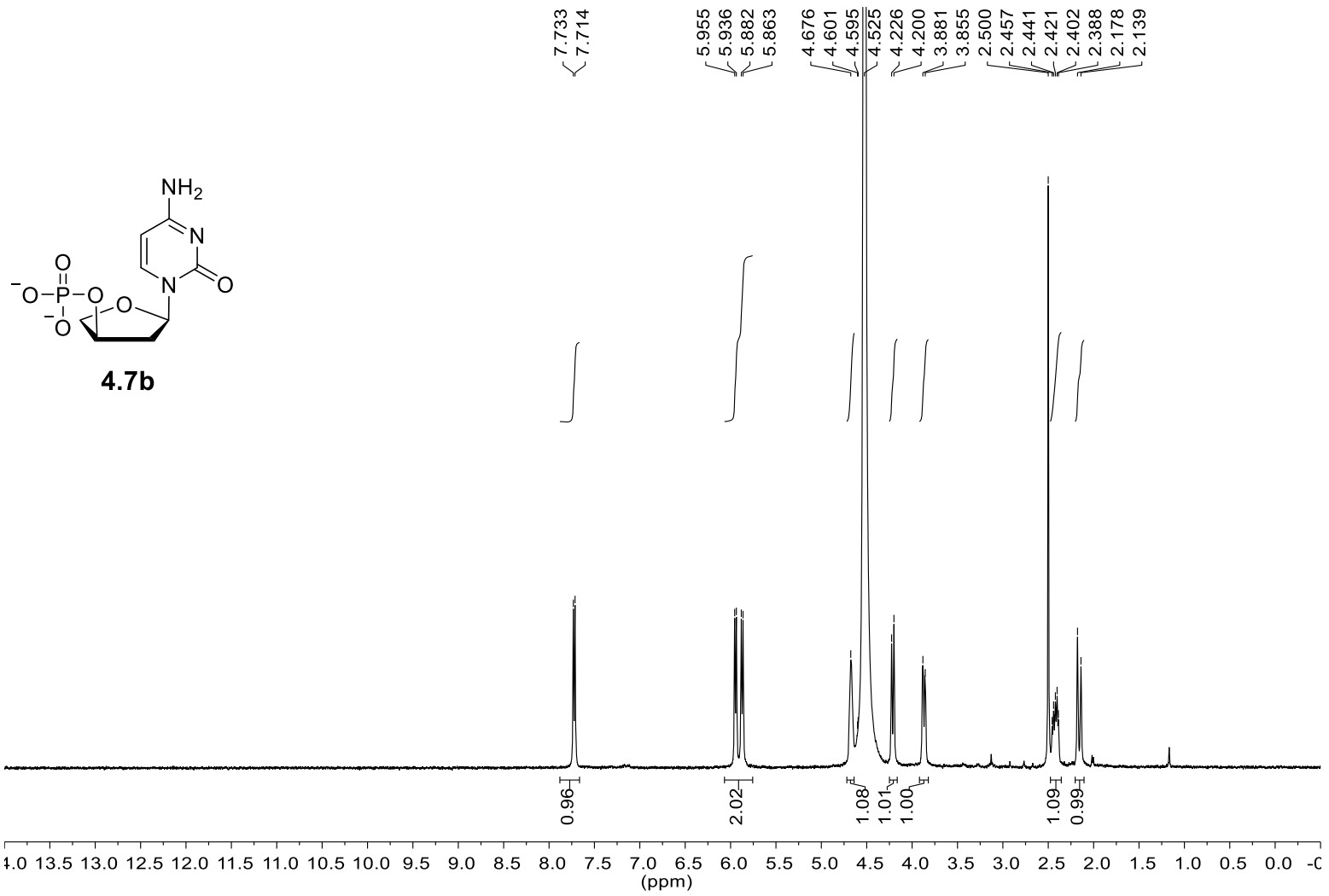


^{13}C NMR spectrum of compound **4.7a** (125.8 MHz, CD_3OD)

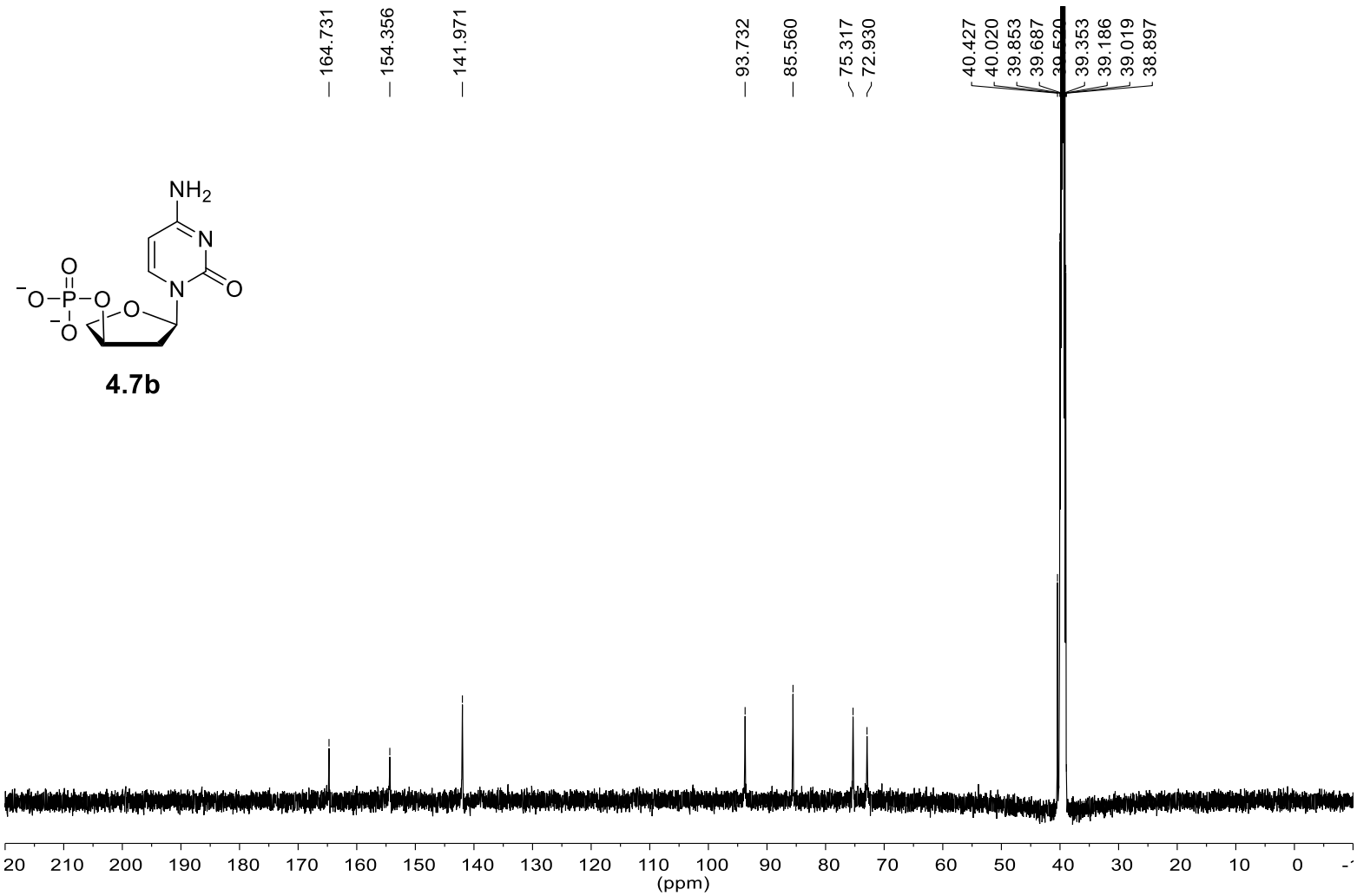


4.7a

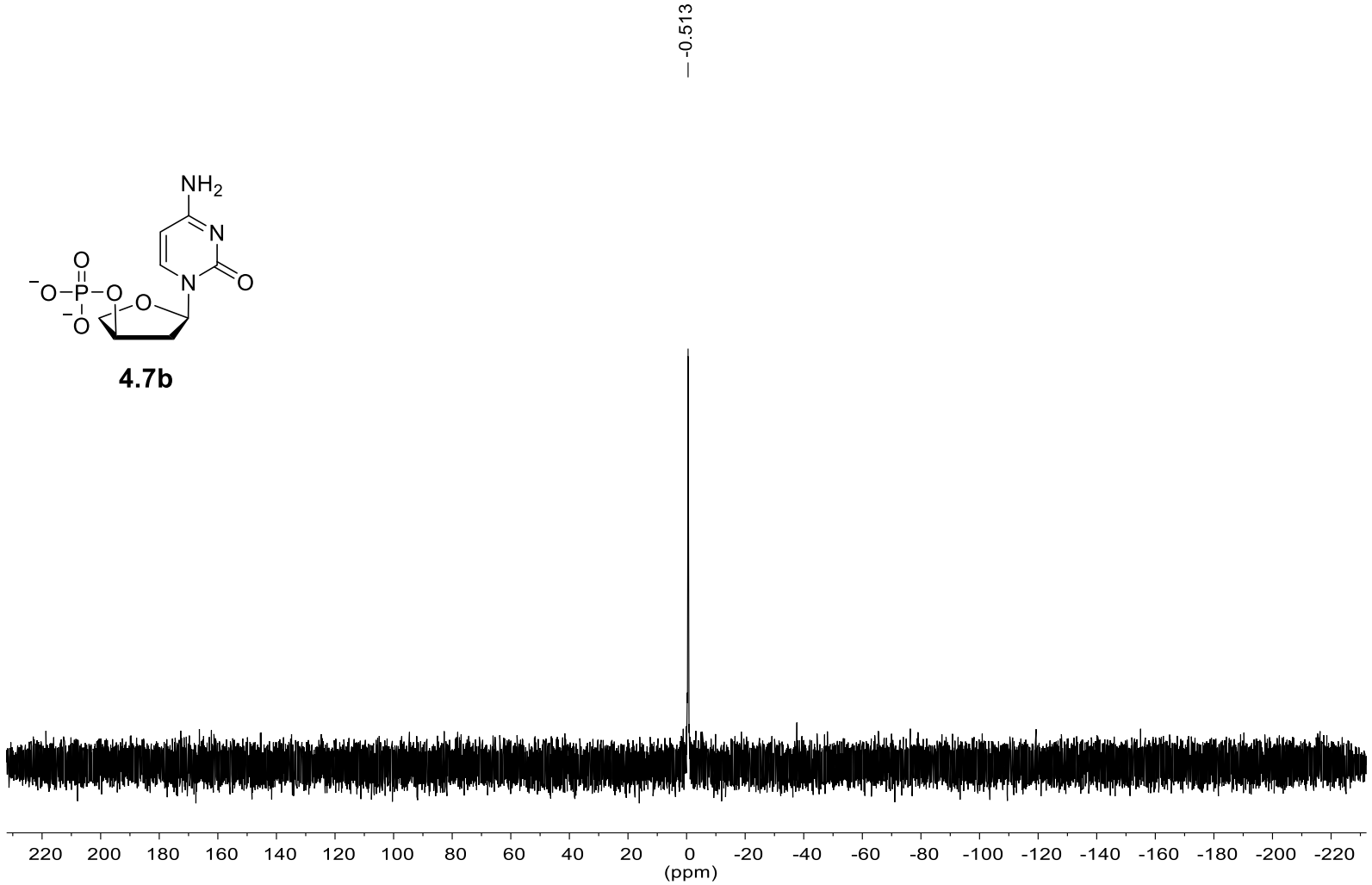
^{31}P NMR spectrum of compound **4.7a** (162 MHz, CD_3OD)



^{1}H NMR spectrum of compound **4.7b** (400 MHz, DMSO-d_6)

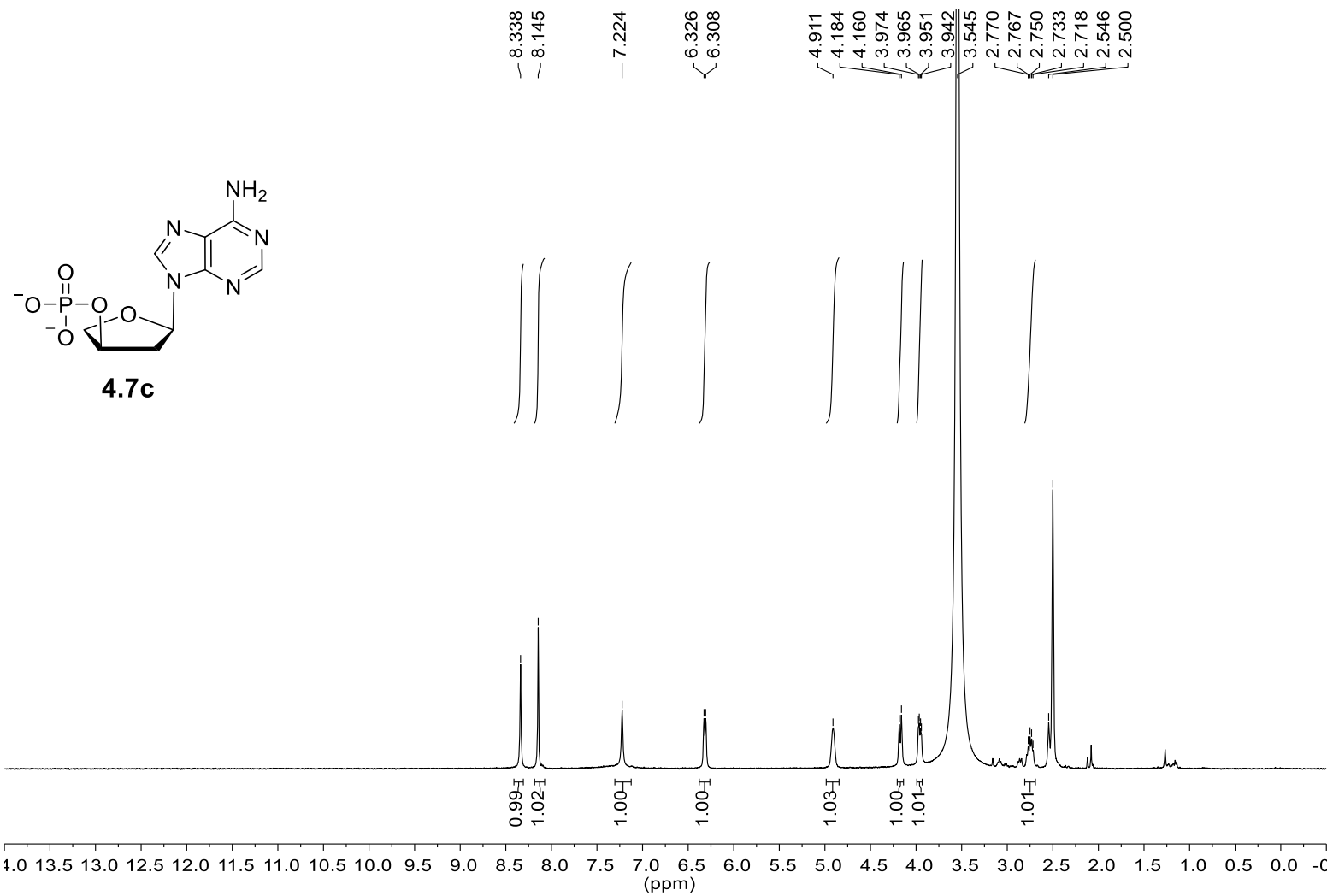


^{13}C NMR spectrum of compound **4.7b** (125.8 MHz, DMSO-d_6)

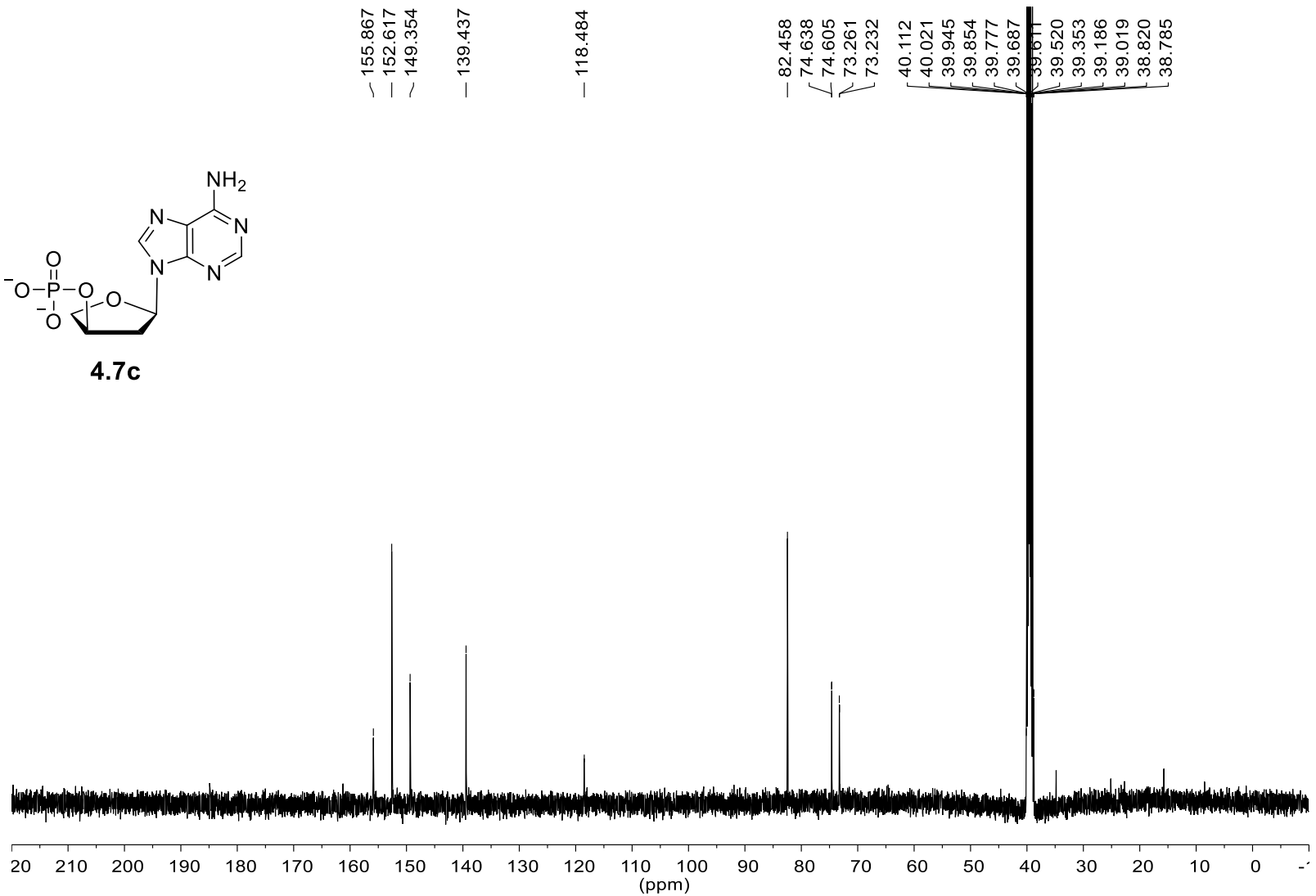


4.7b

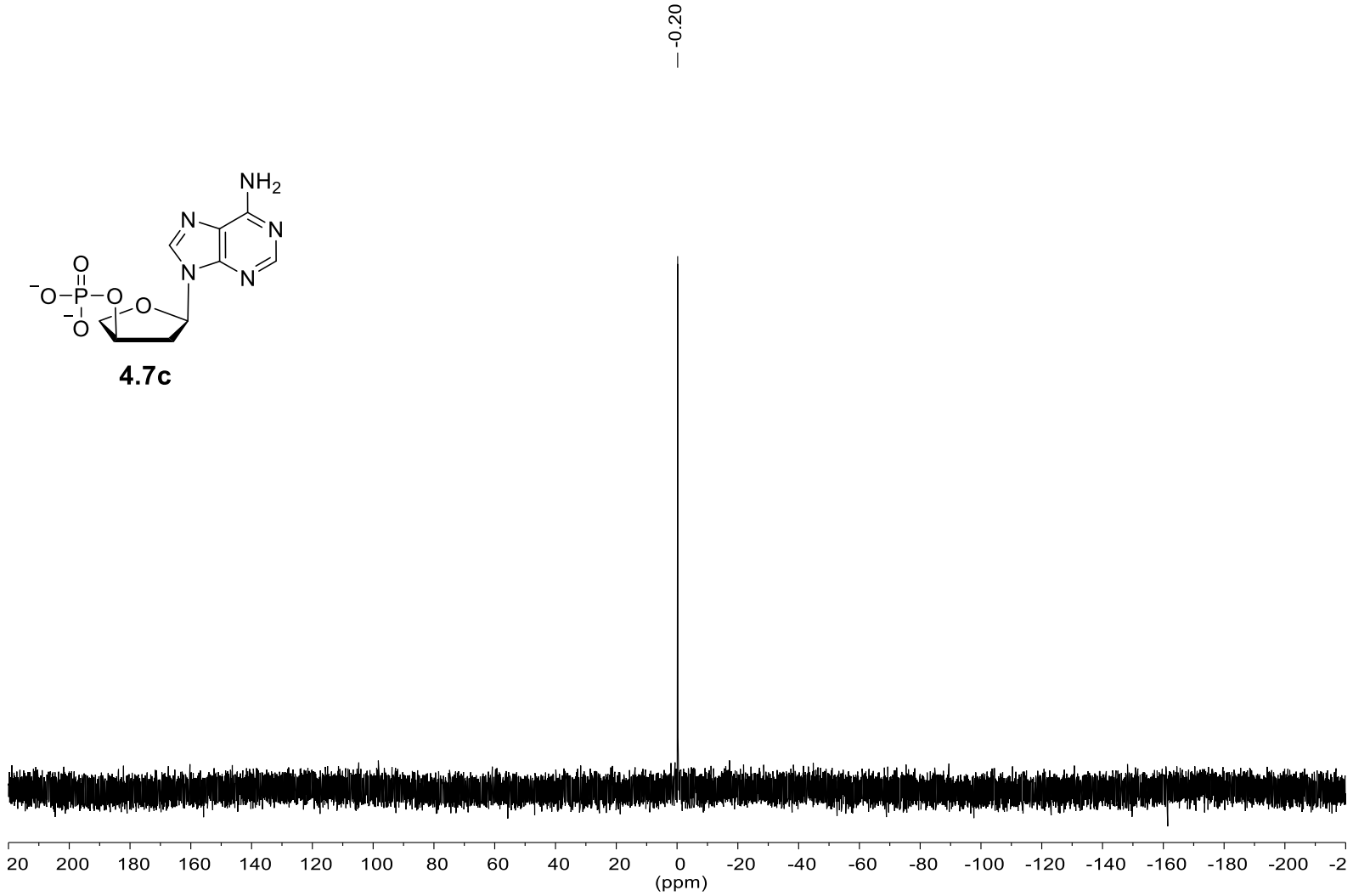
^{31}P NMR spectrum of compound **4.7b** (162 MHz, DMSO-d_6)



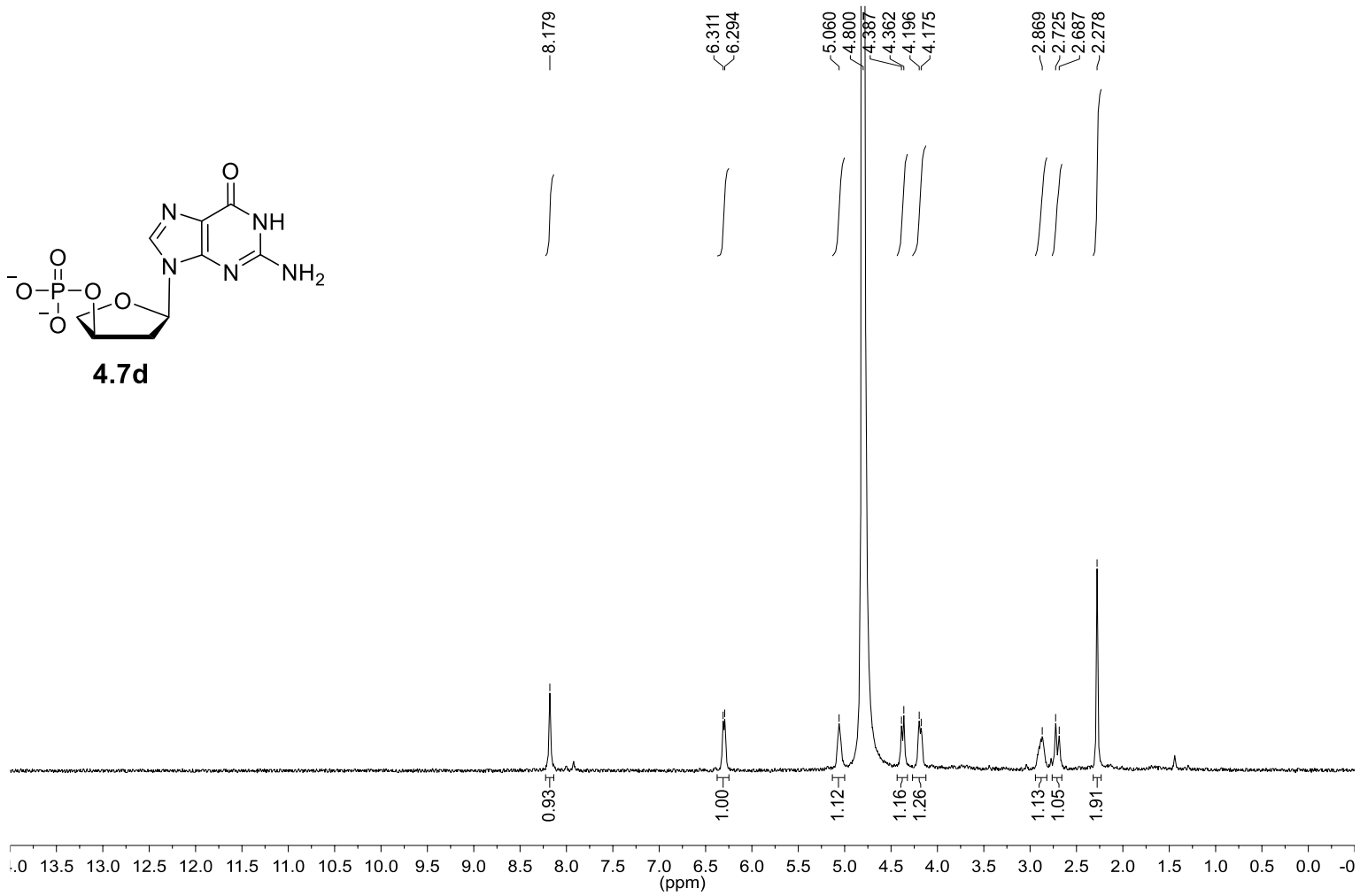
^{1}H NMR spectrum of compound **4.7c** (400 MHz, DMSO-d_6)



^{13}C NMR spectrum of compound **4.7c** (125.8 MHz, DMSO-d_6)

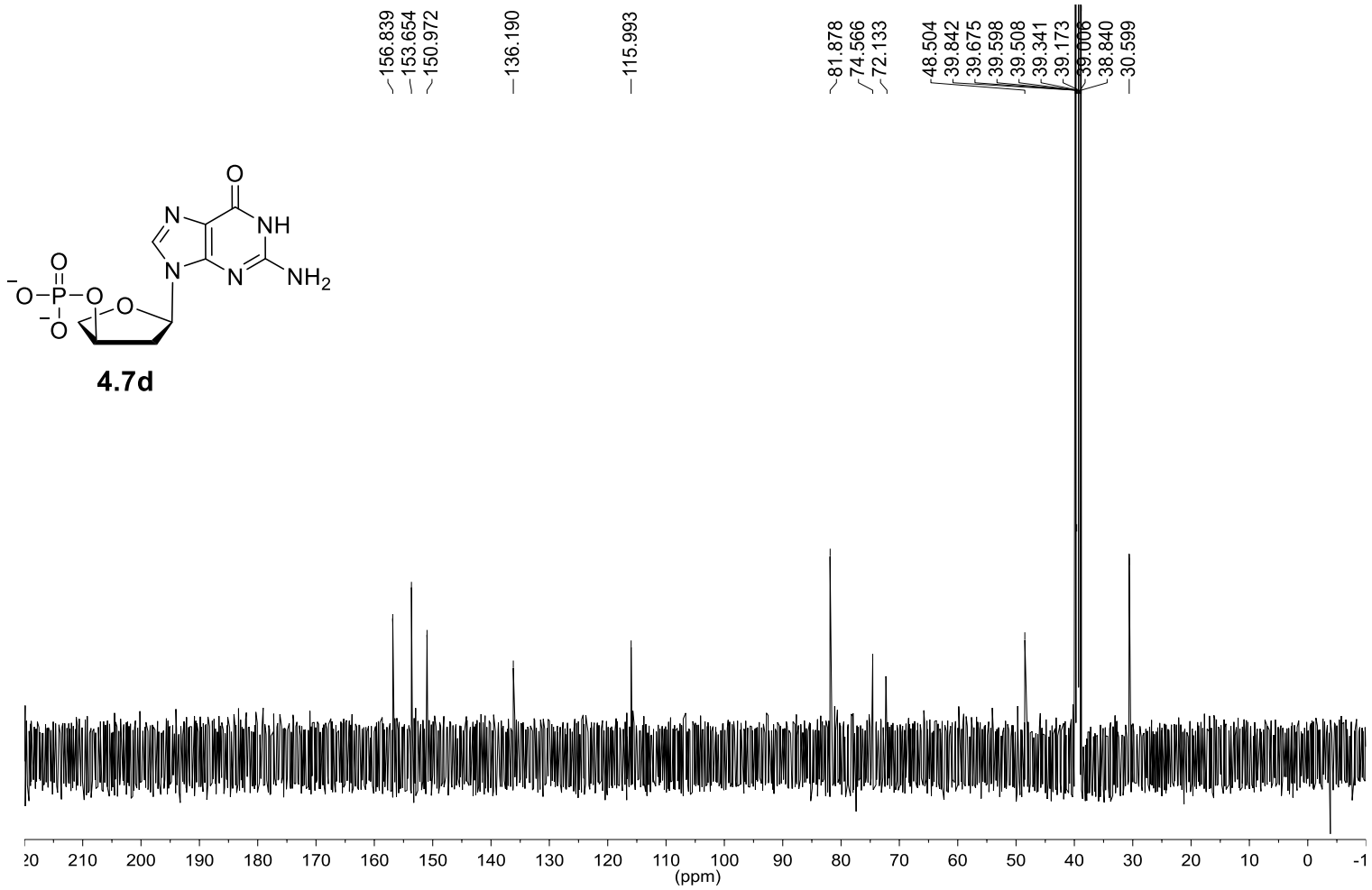


^{31}P NMR spectrum of compound **4.7c** (162 MHz, DMSO-d_6)

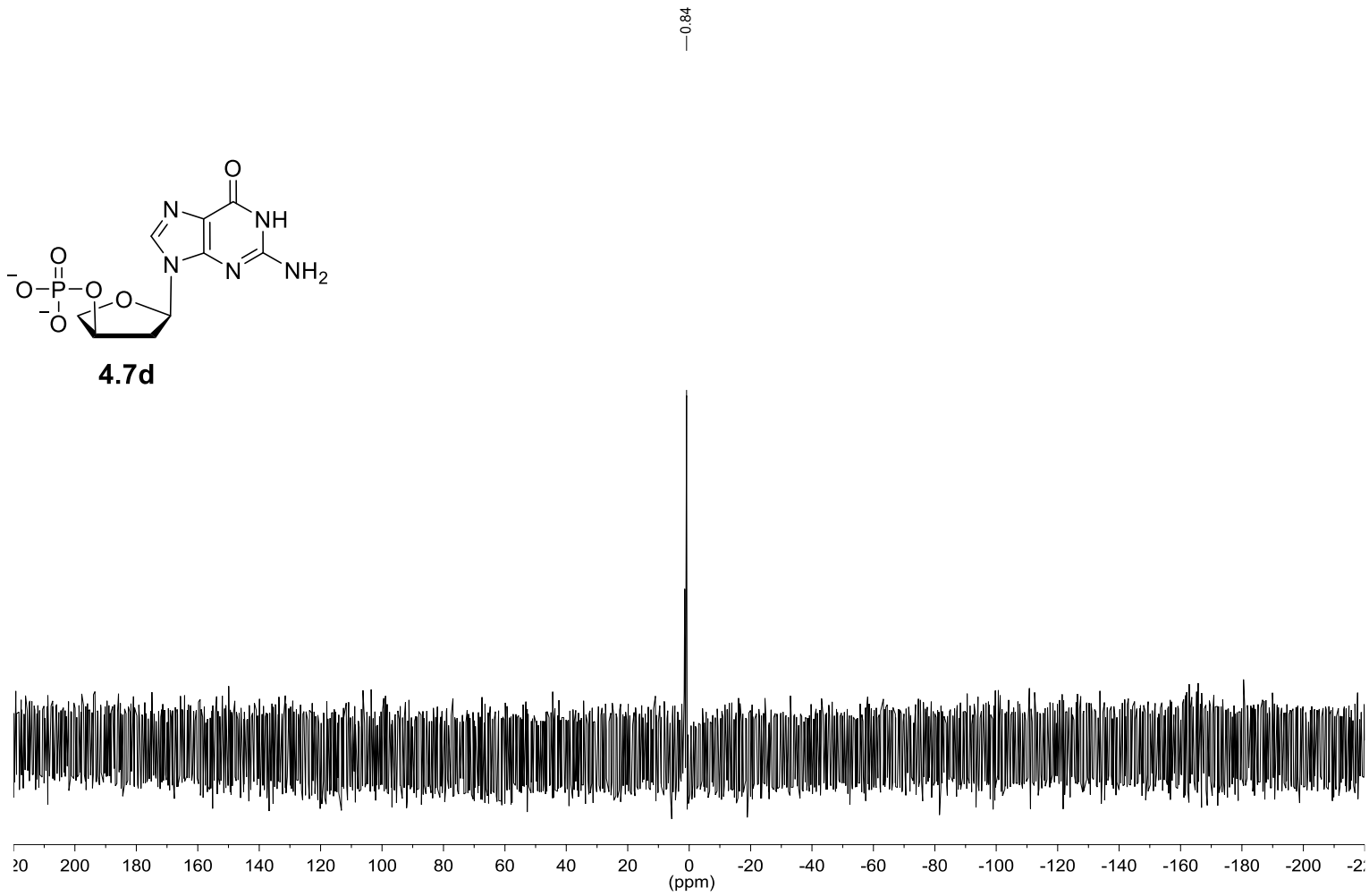


4.7d

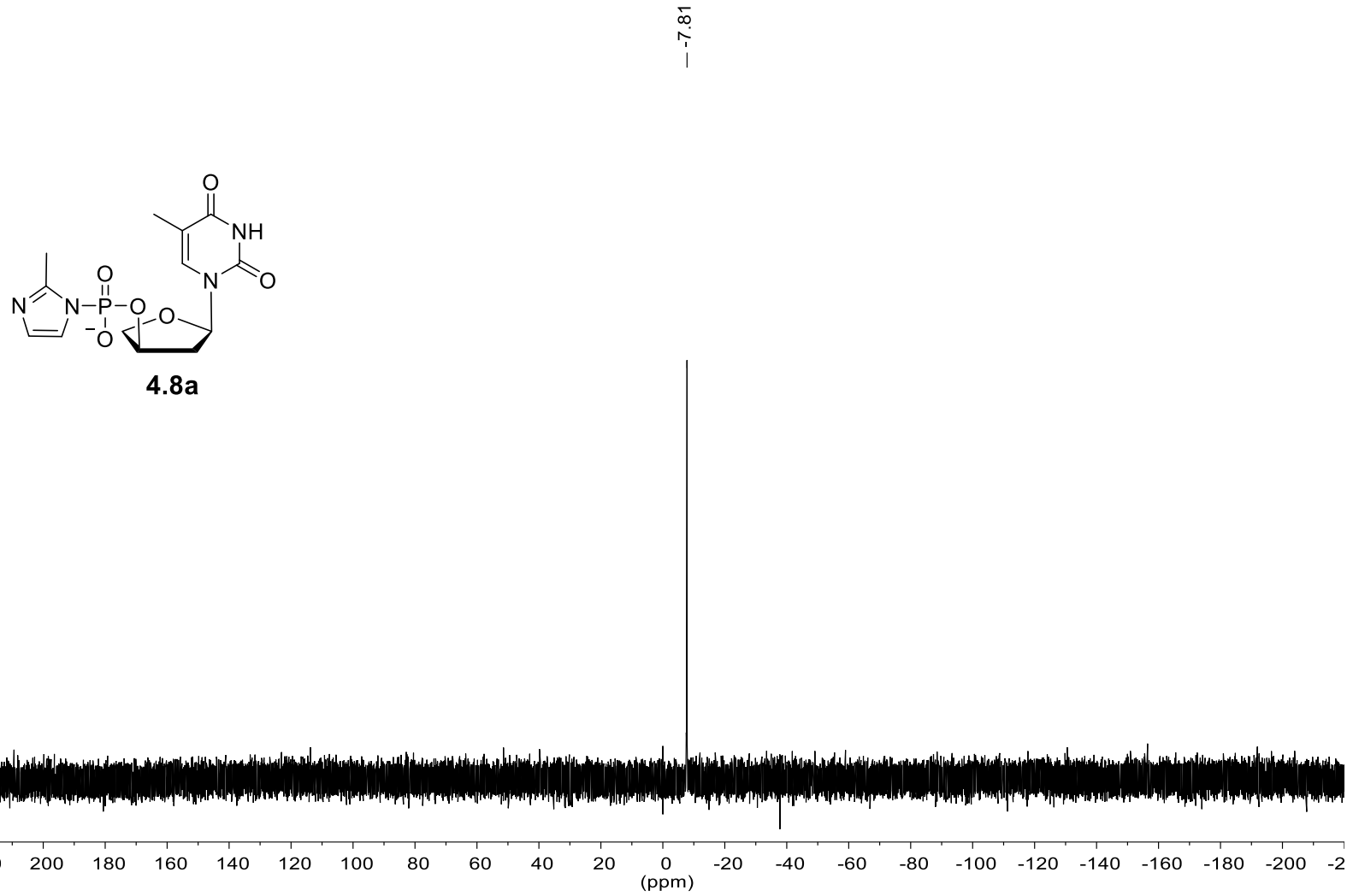
^{1}H NMR spectrum of compound **4.7d** (400 MHz, D_2O)



^{13}C NMR spectrum of compound **4.7d** (125.8 MHz, DMSO-d_6)

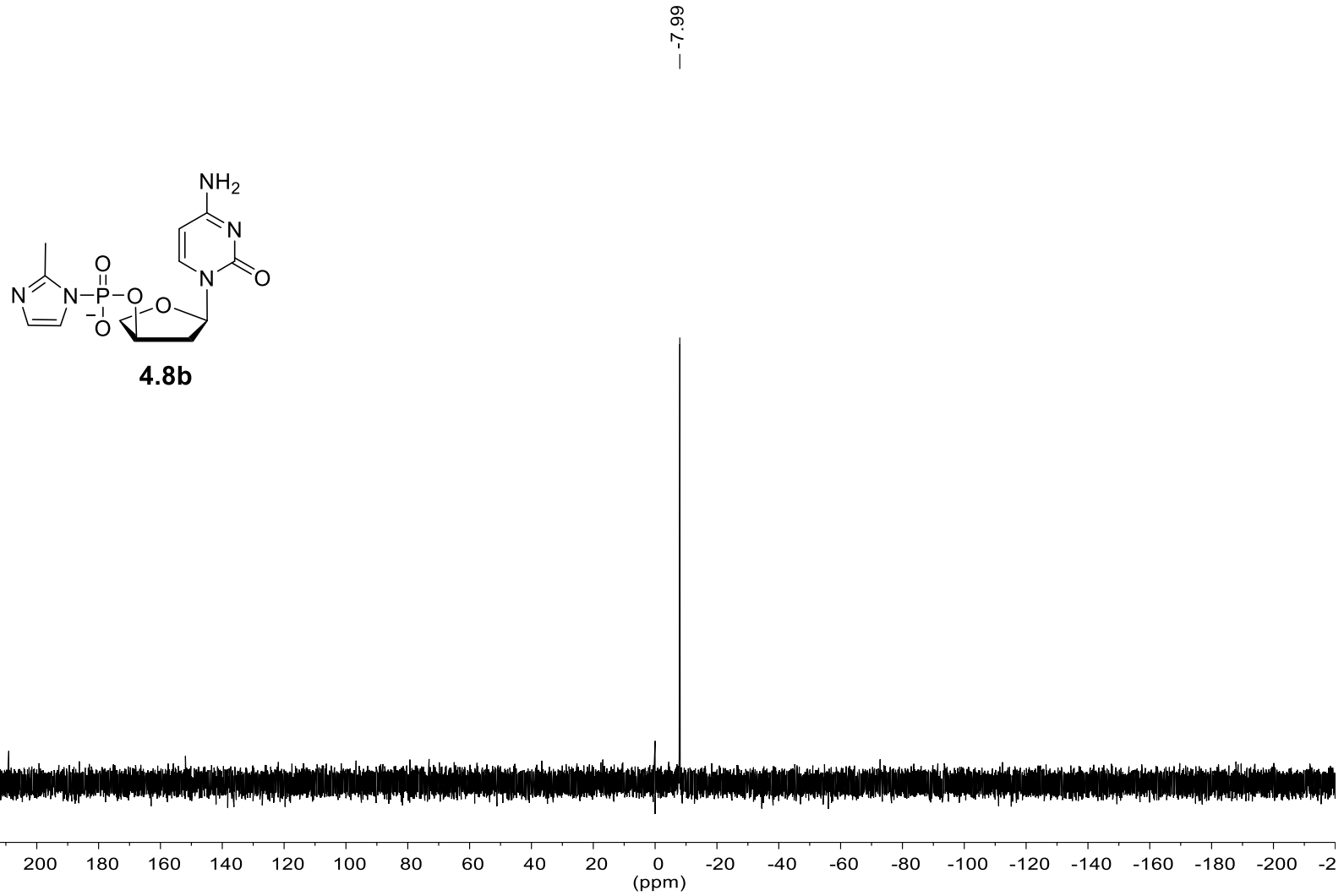


^{31}P NMR spectrum of compound **4.7d** (162 MHz, DMSO-d_6)



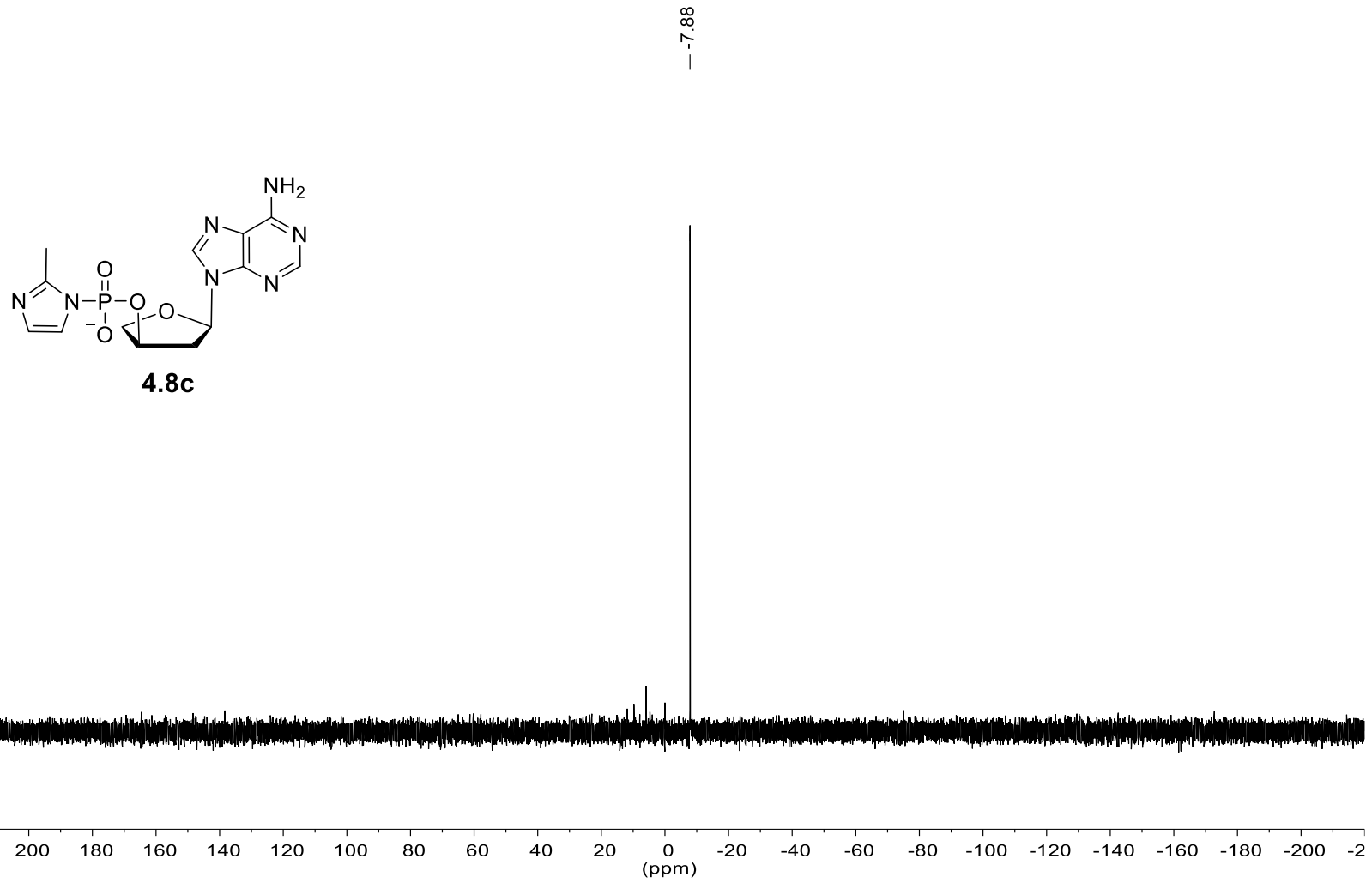
4.8a

^{31}P NMR spectrum of compound **4.8a** (162 MHz, D_2O)



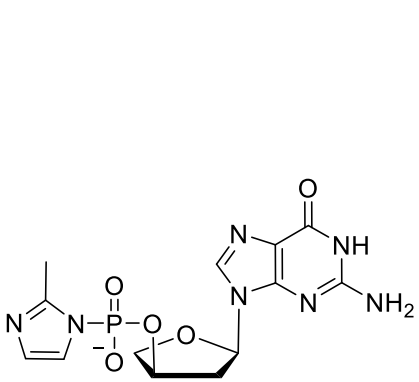
4.8b

^{31}P NMR spectrum of compound **4.8b** (162 MHz, D_2O)



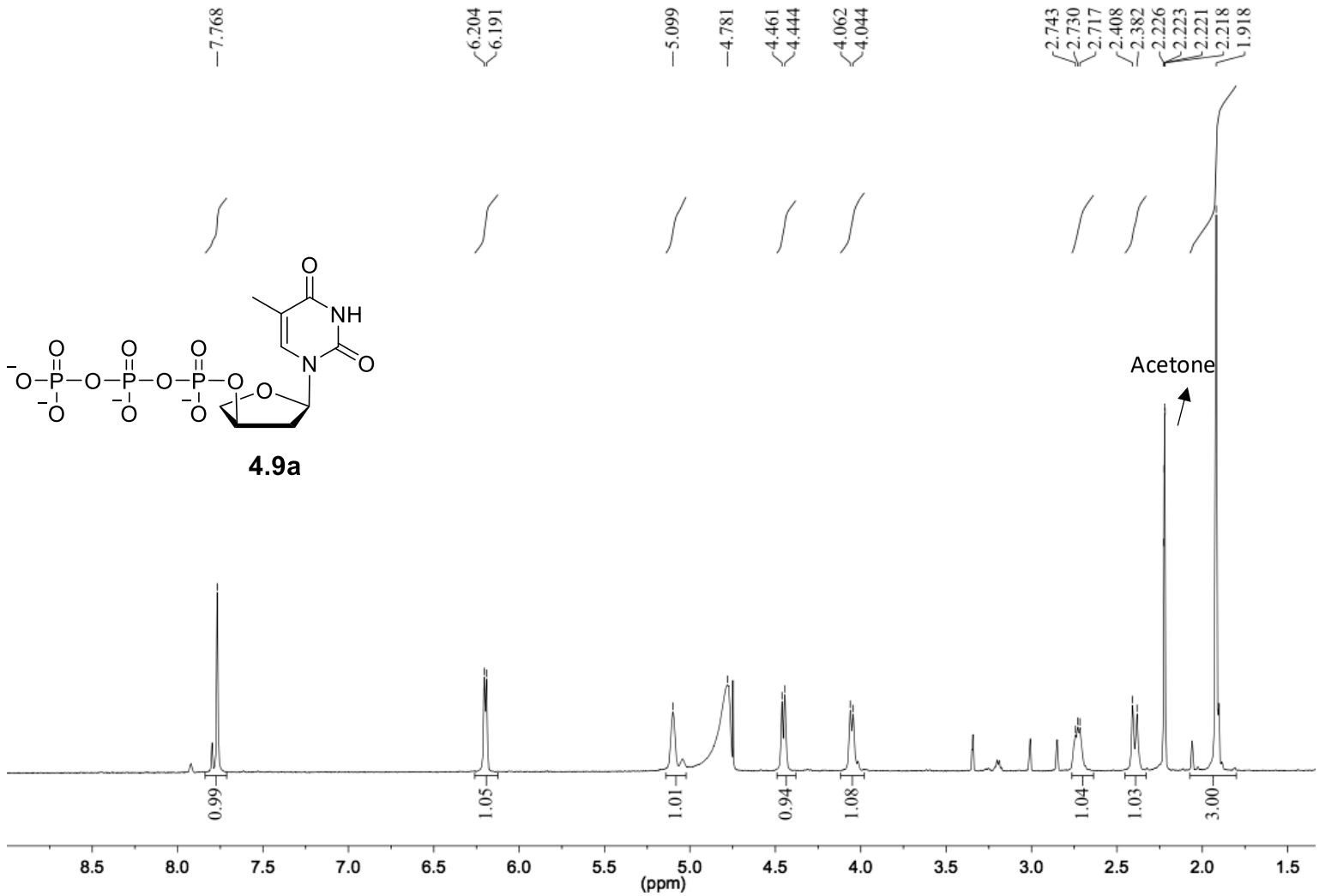
4.8c

^{31}P NMR spectrum of compound **4.8c** (162 MHz, D_2O)



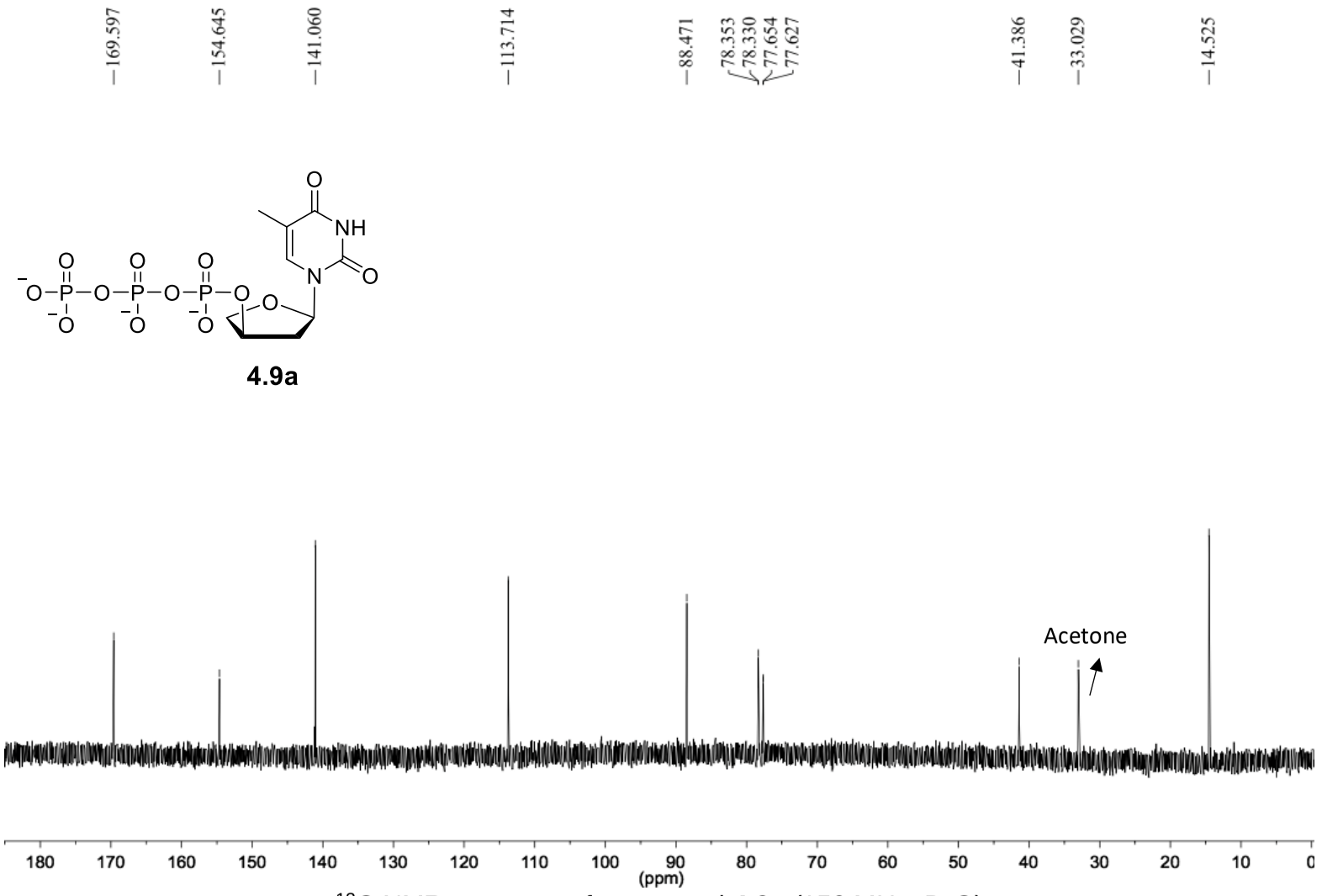
4.8d

^{31}P NMR spectrum of compound **4.8d** (162 MHz, DMSO-d_6)



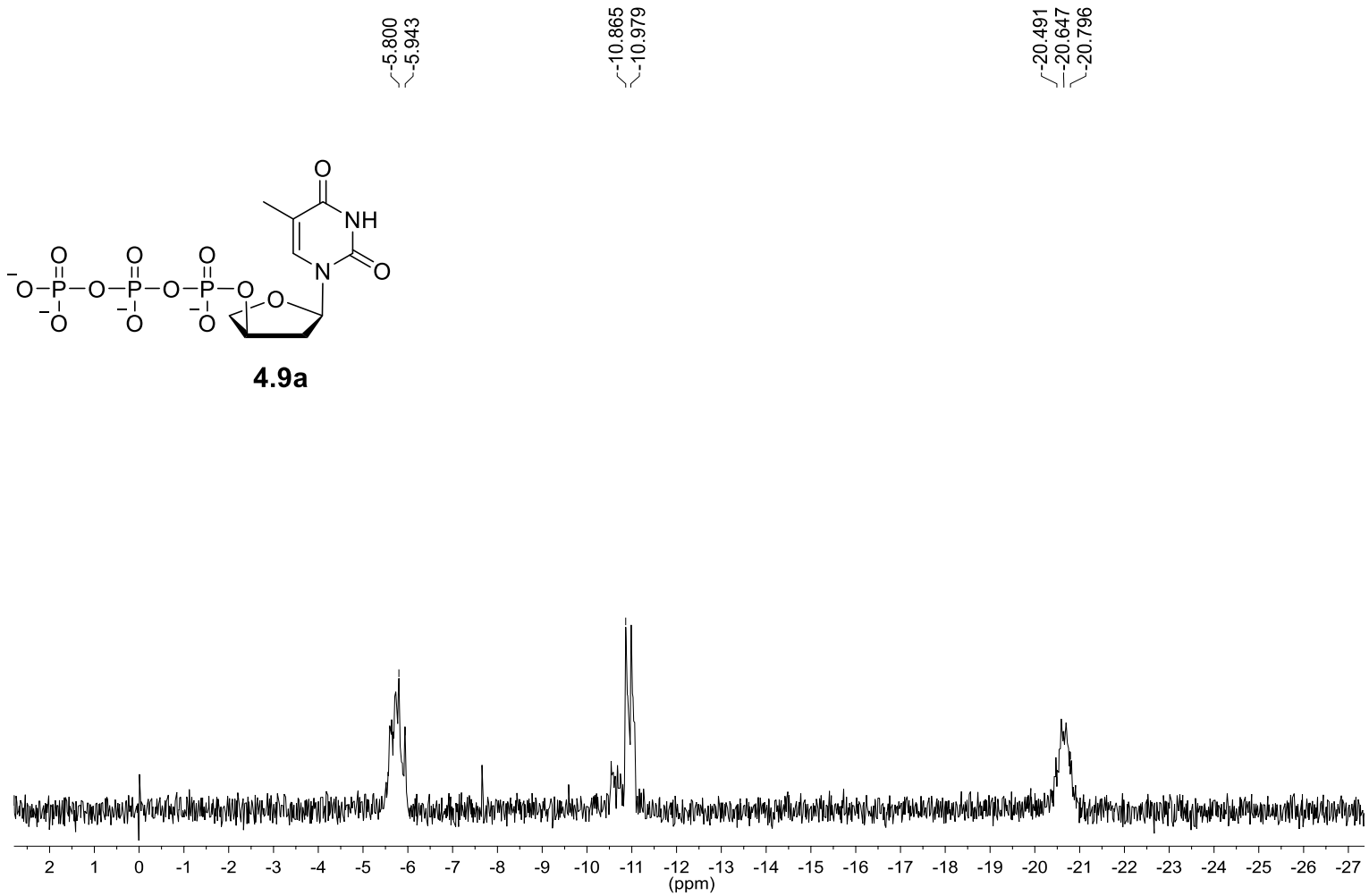
4.9a

^{1}H NMR spectrum of compound **4.9a** (600 MHz, D_2O)



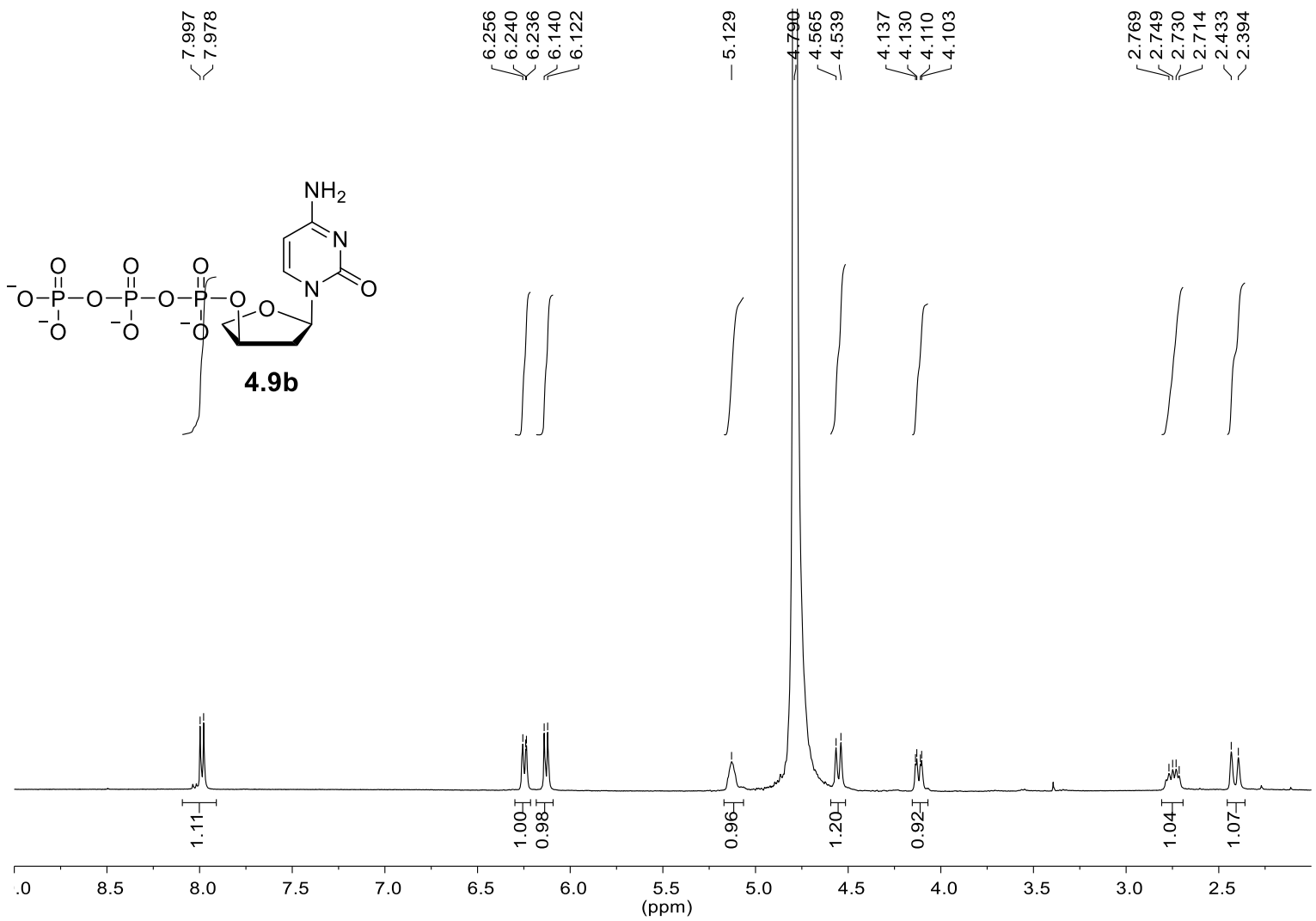
4.9a

^{13}C NMR spectrum of compound **4.9a** (150 MHz, D_2O)

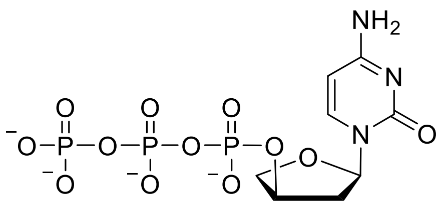


4.9a

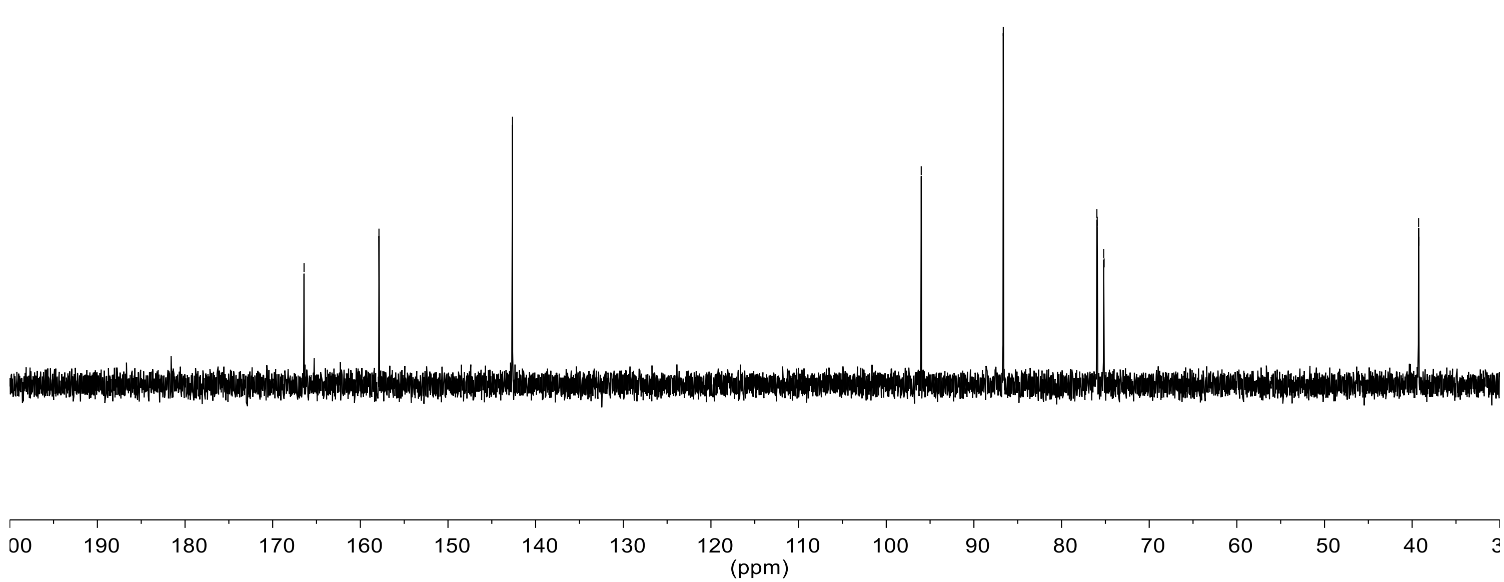
^{31}P NMR spectrum of compound **4.9a** (162 MHz, D_2O)



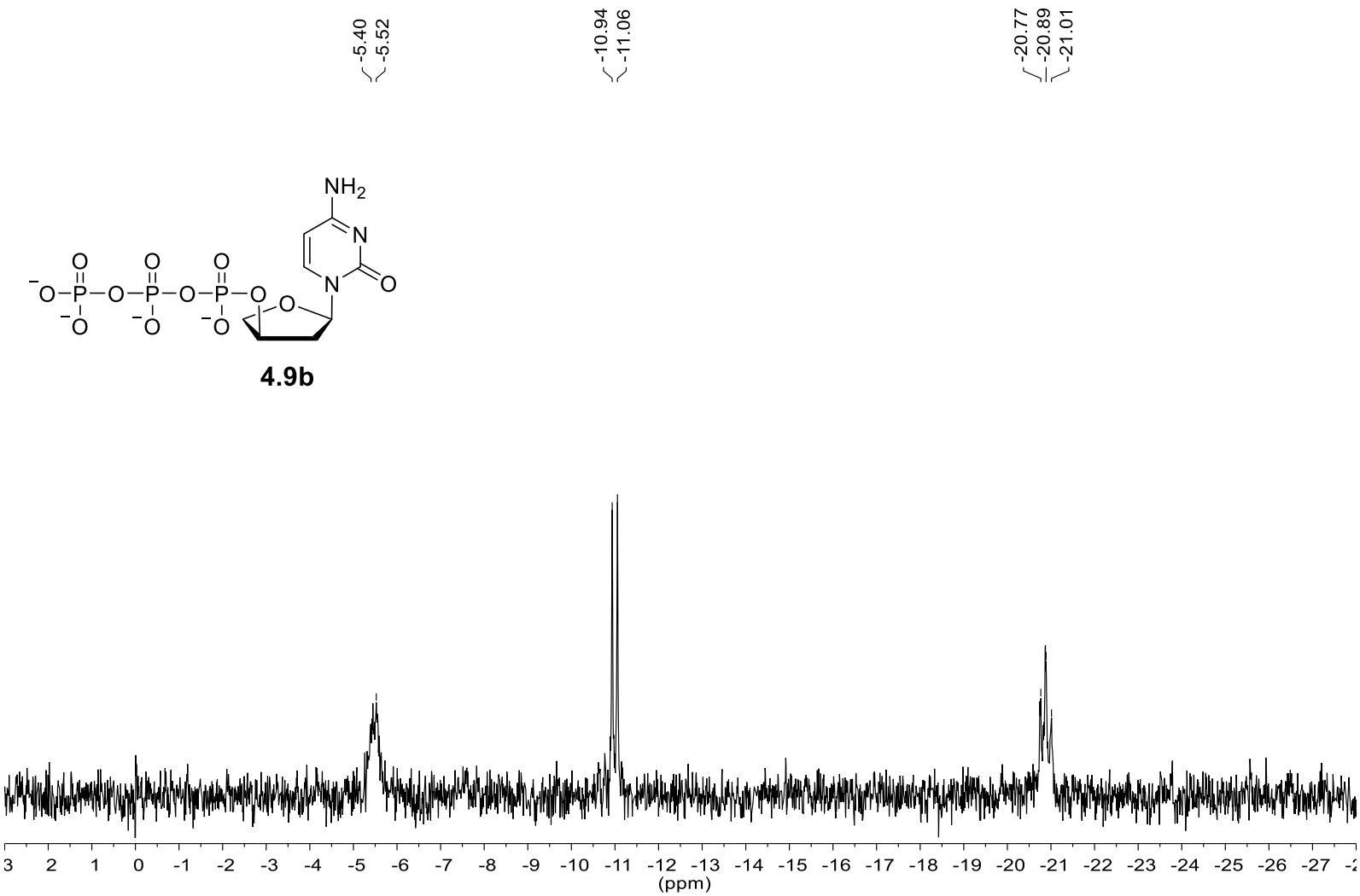
^{1}H NMR spectrum of compound **4.9b** (600 MHz, D_2O)



4.9b

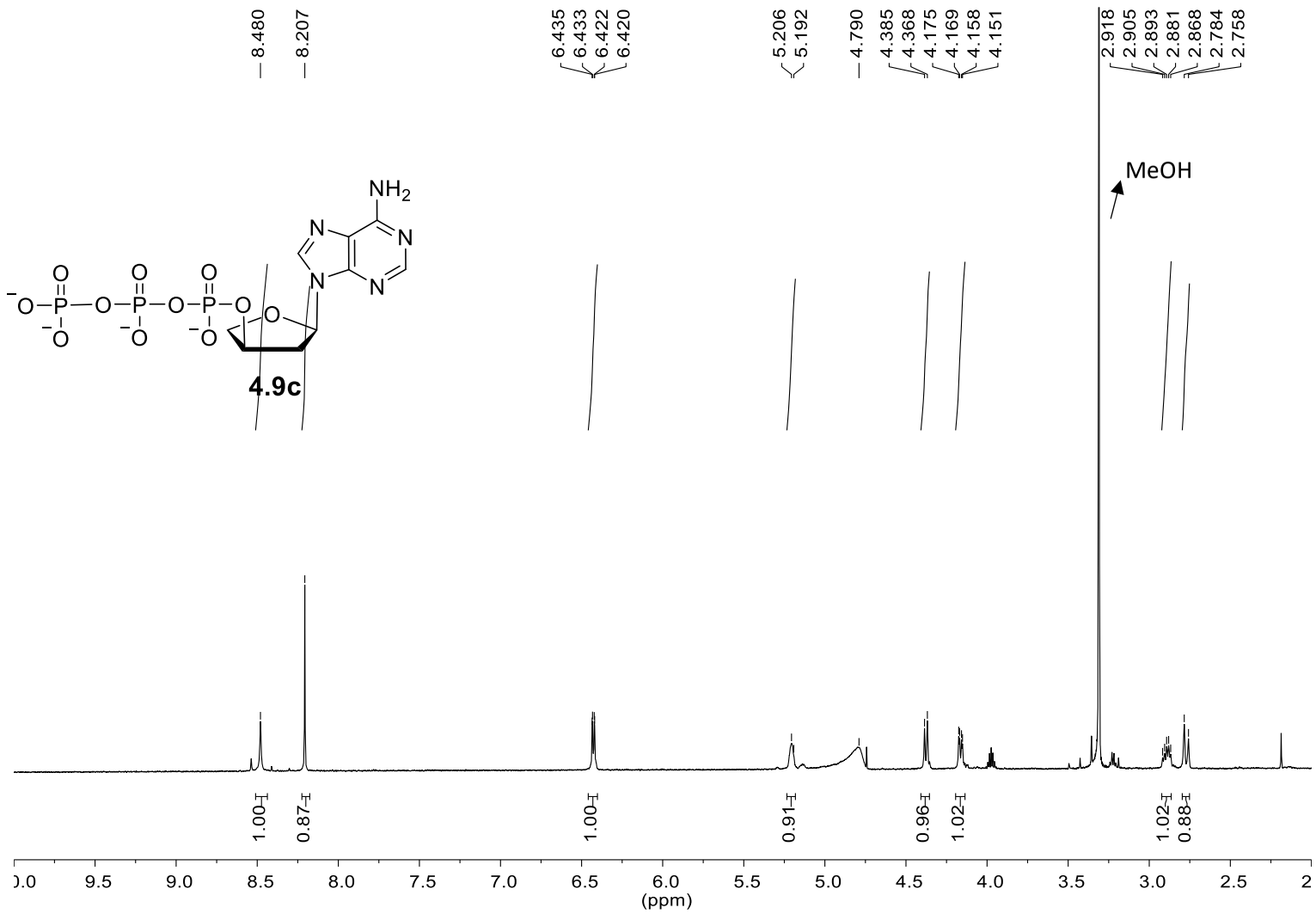


^{13}C NMR spectrum of compound **4.9b** (150 MHz, D_2O)

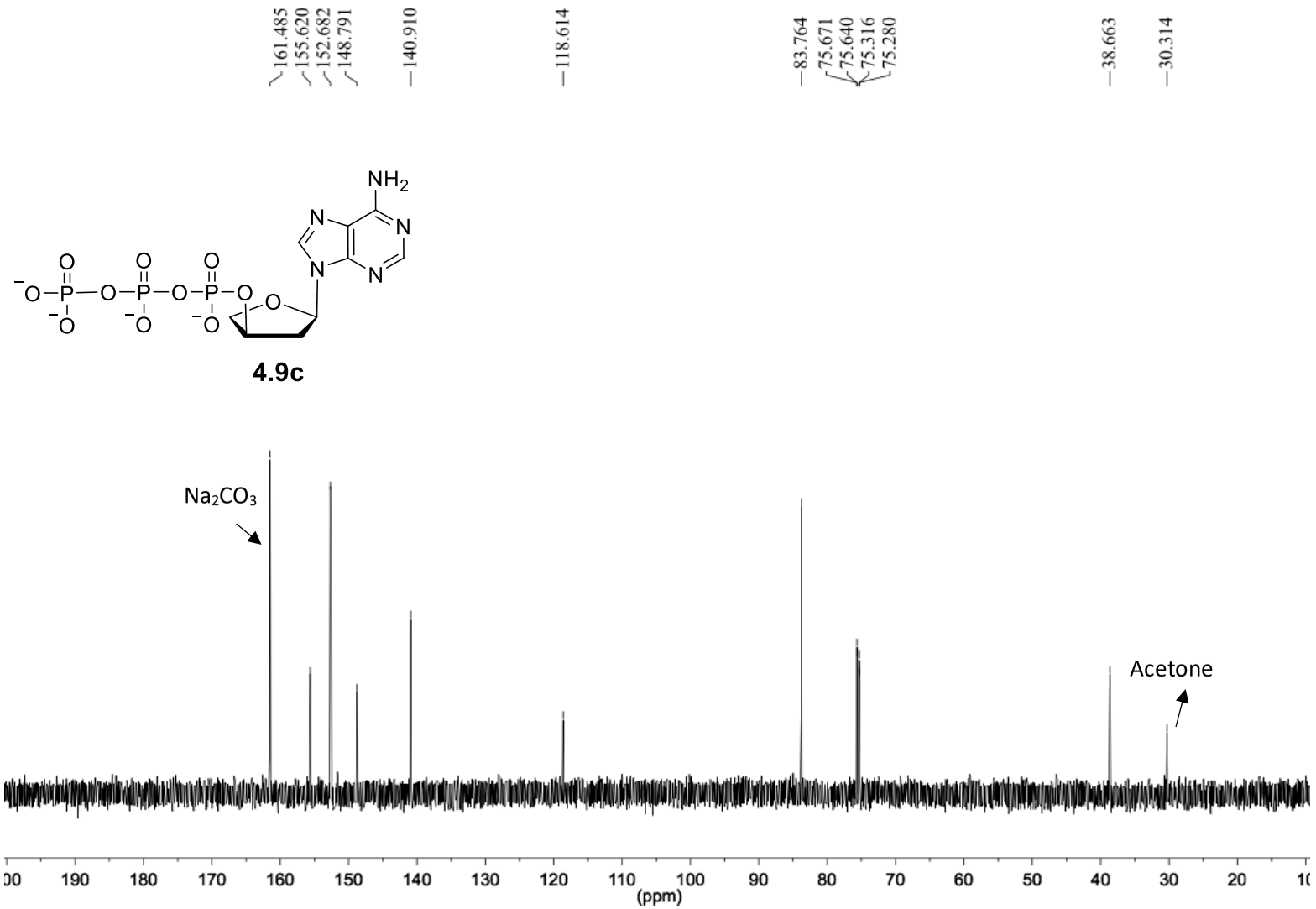


4.9b

^{31}P NMR spectrum of compound **4.9b** (162 MHz, D_2O)

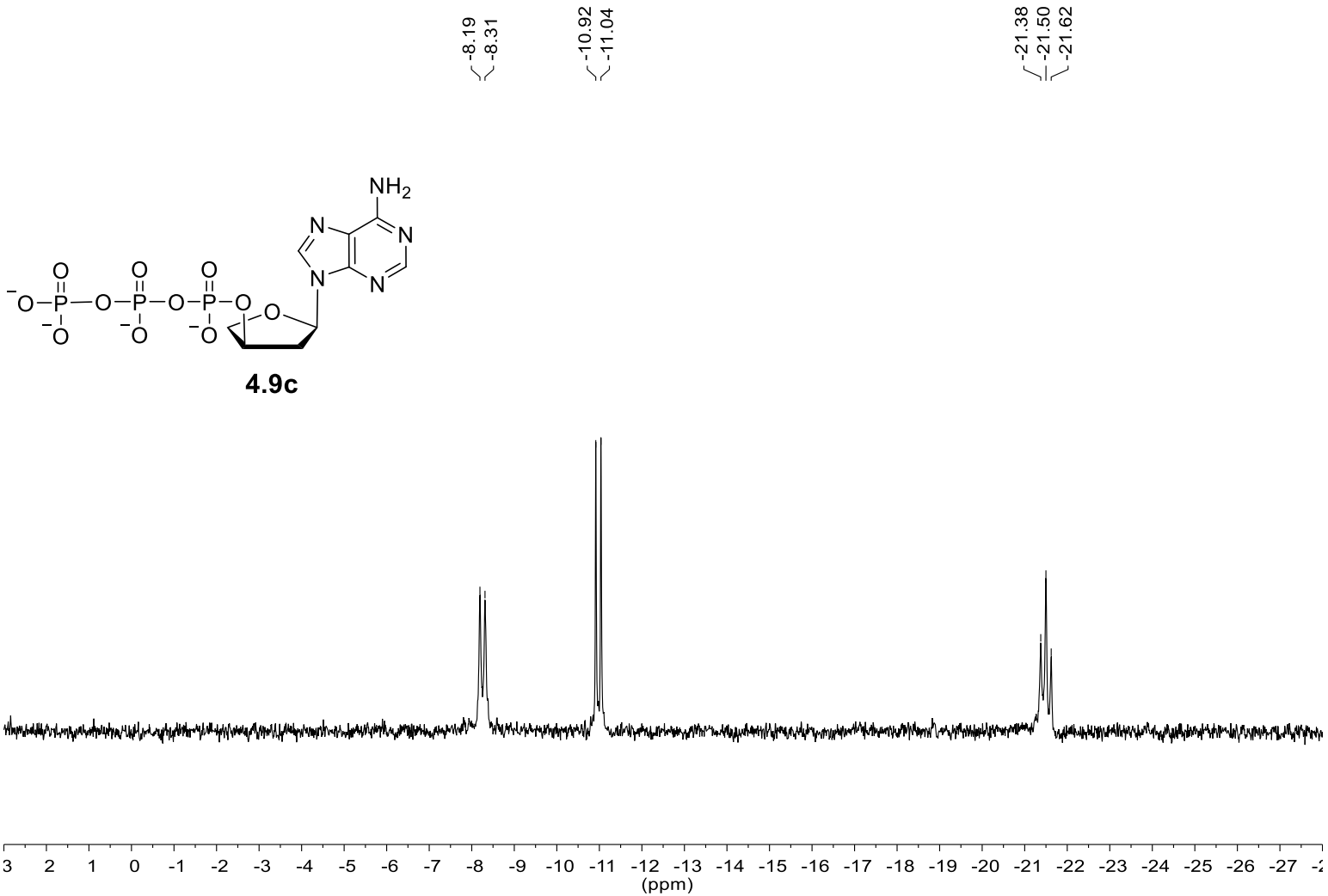


^{1}H NMR spectrum of compound **4.9c** (600 MHz, D_2O)



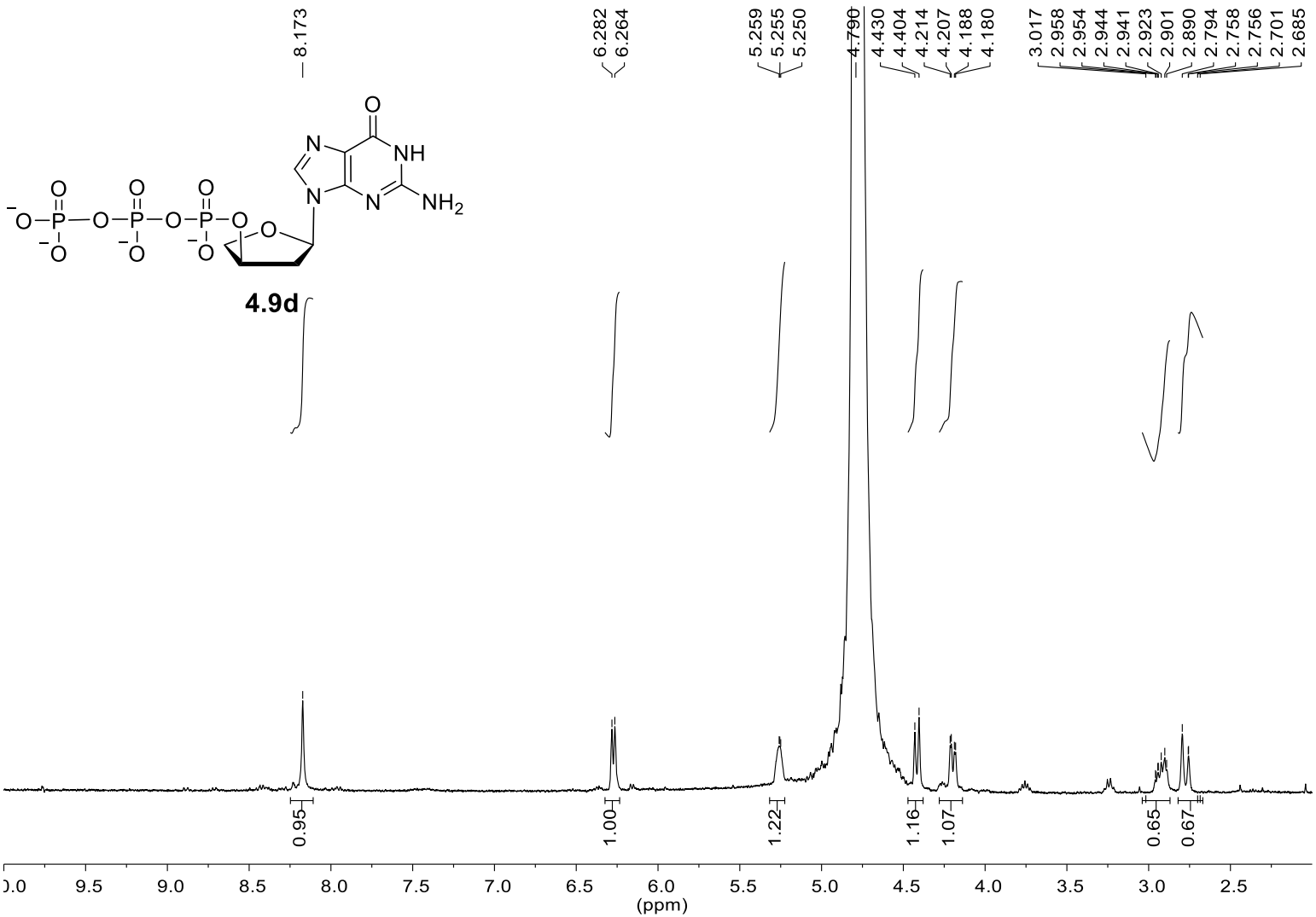
4.9c

^{13}C NMR spectrum of compound **4.9c** (150 MHz, D_2O)

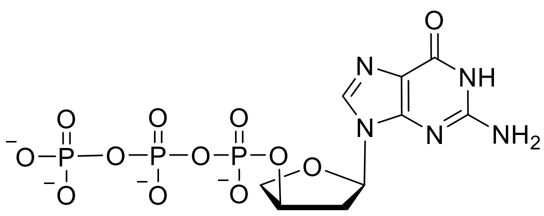


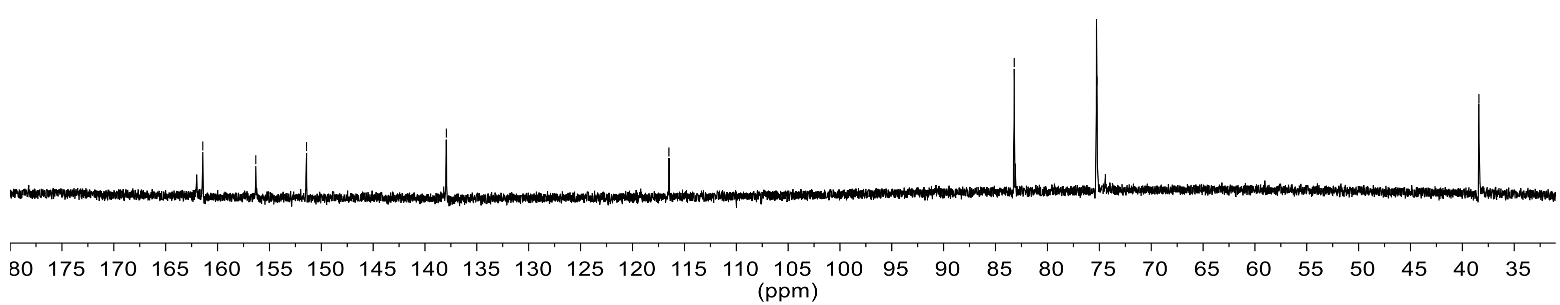
4.9c

^{31}P NMR spectrum of compound **4.9c** (162 MHz, D_2O)

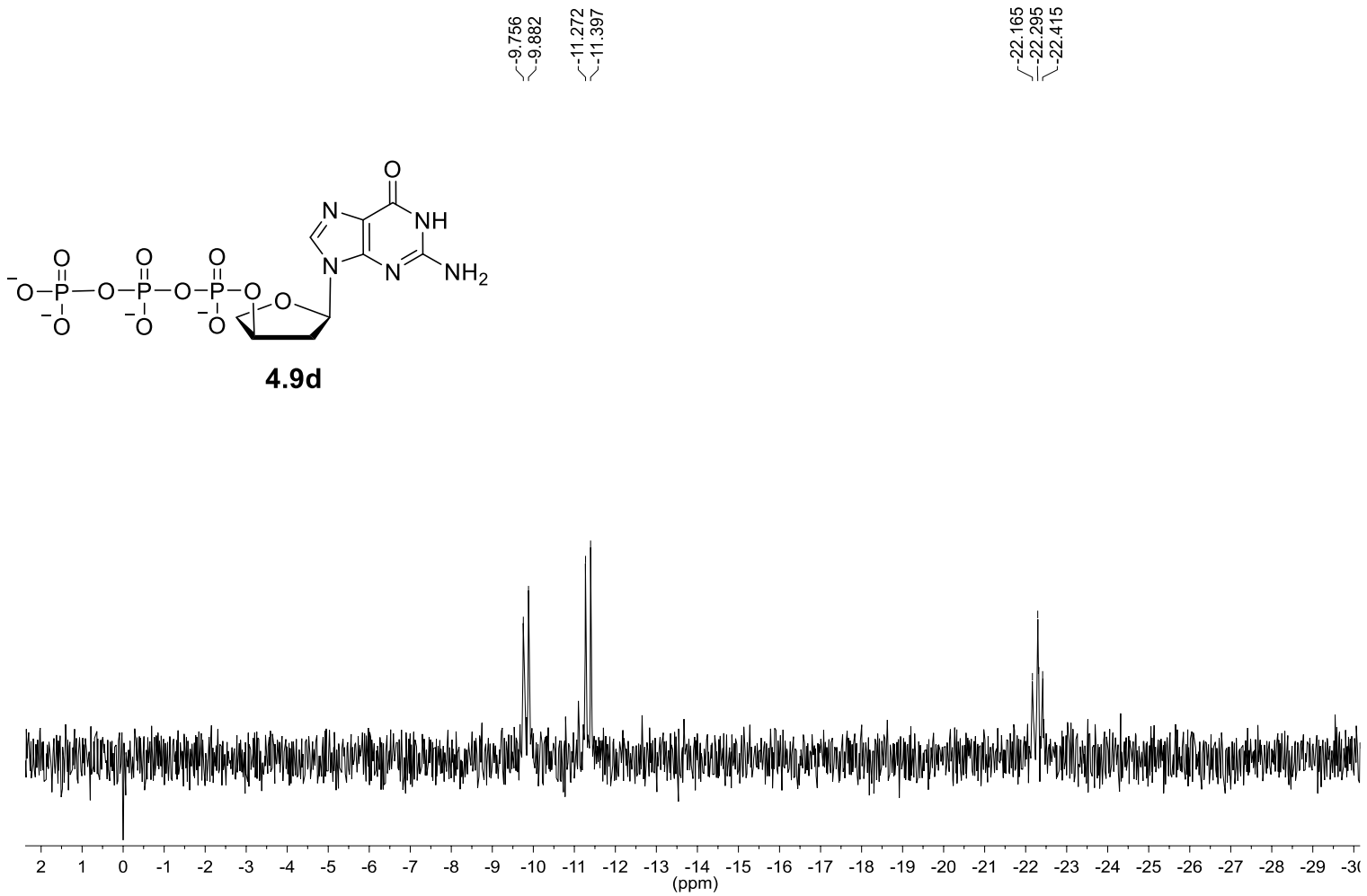


^{1}H NMR spectrum of compound **4.9d** (600 MHz, D_2O)





^{13}C NMR spectrum of compound **4.9d** (150 MHz, D_2O)

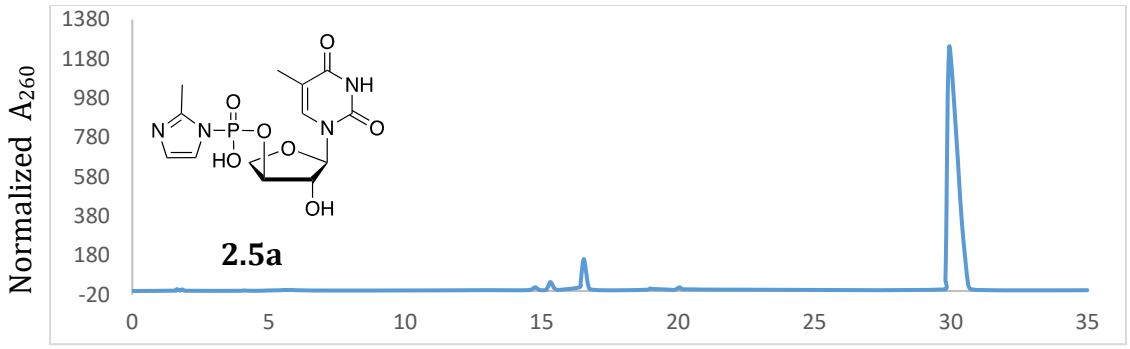


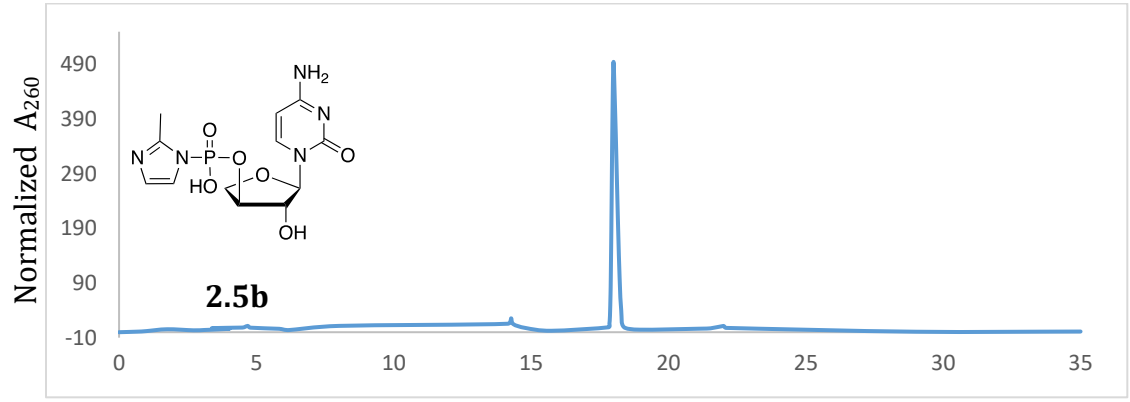
4.9d

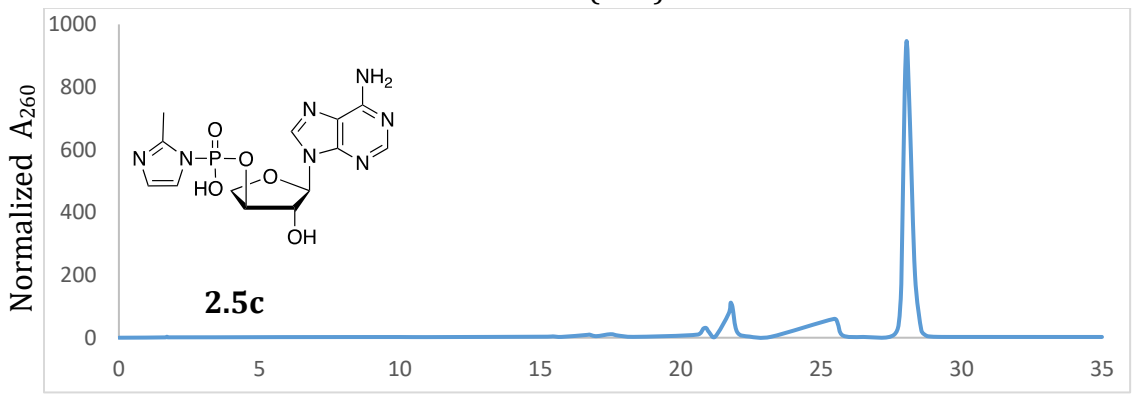
^{31}P NMR spectrum of compound **4.9d** (162 MHz, D_2O)

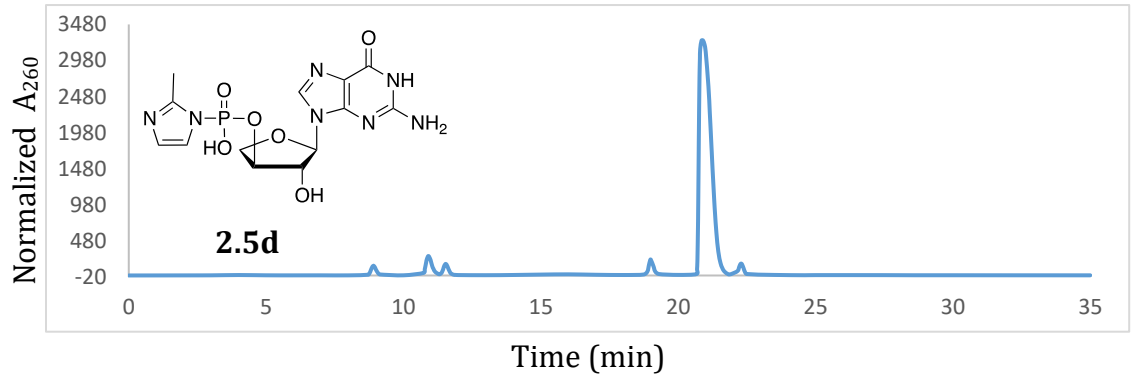
APPENDIX B

HPLC DATA

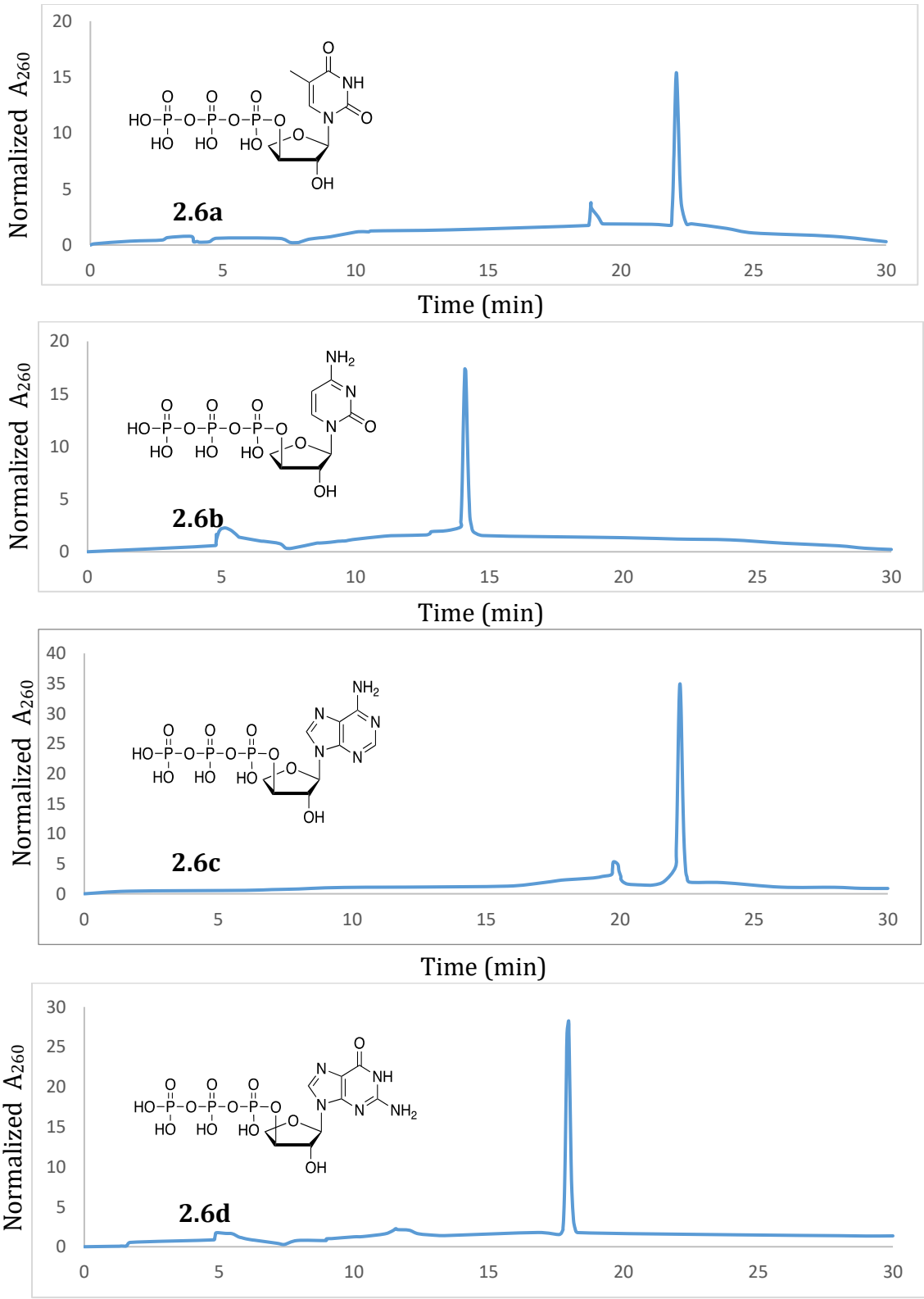




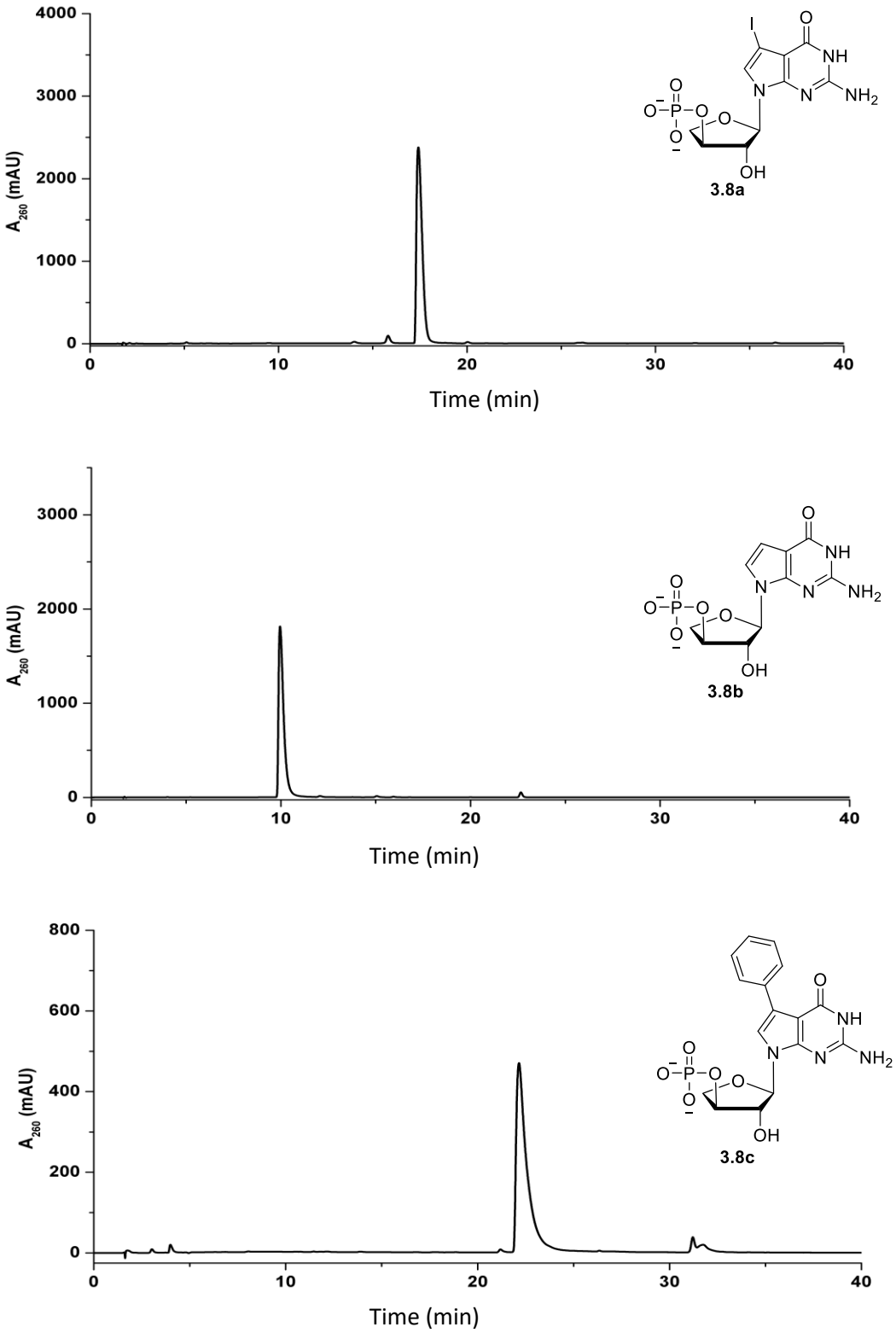




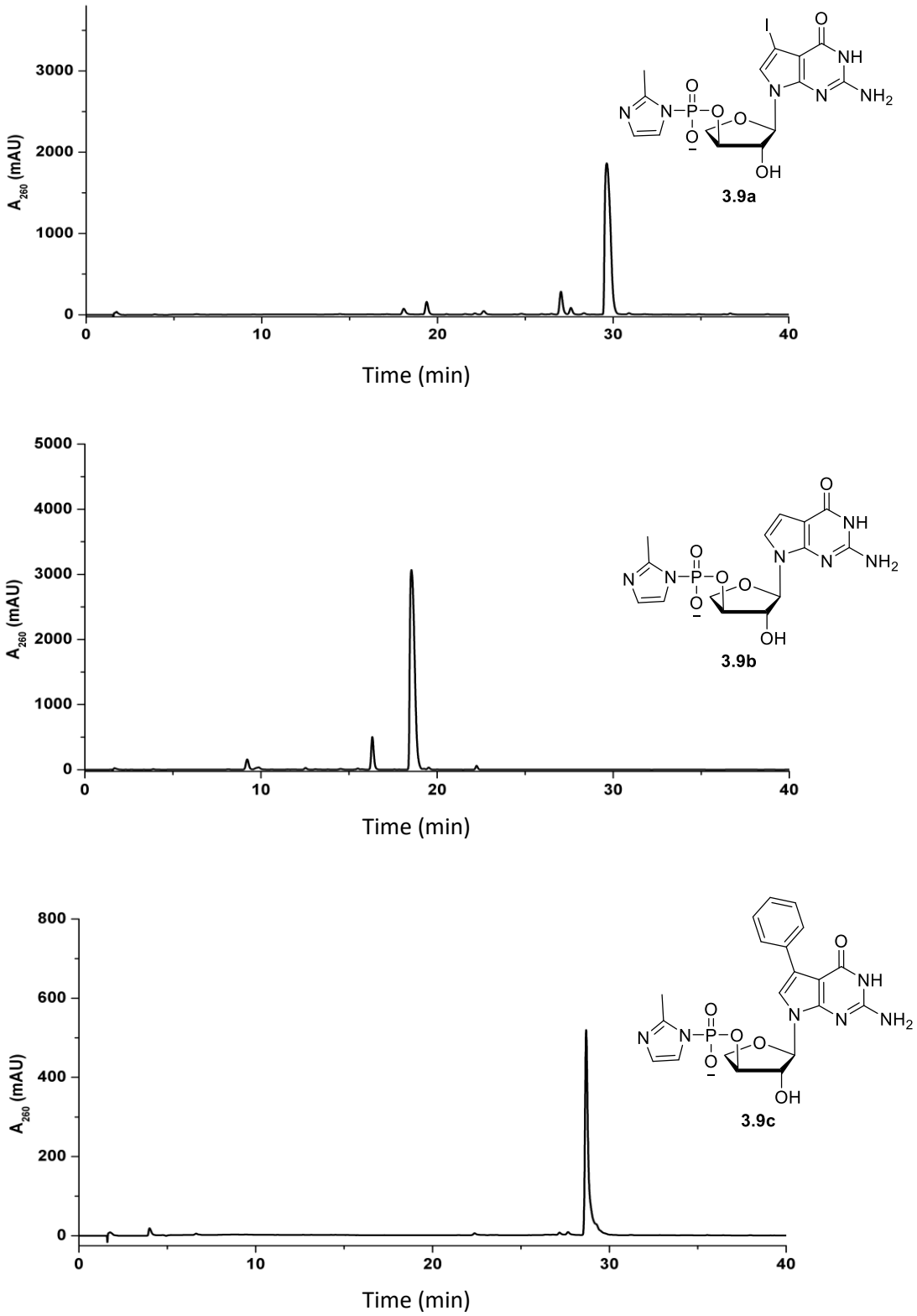
HPLC analysis of **2.5a-d** (MeCN/0.1M TEAA from 0 to 50%).



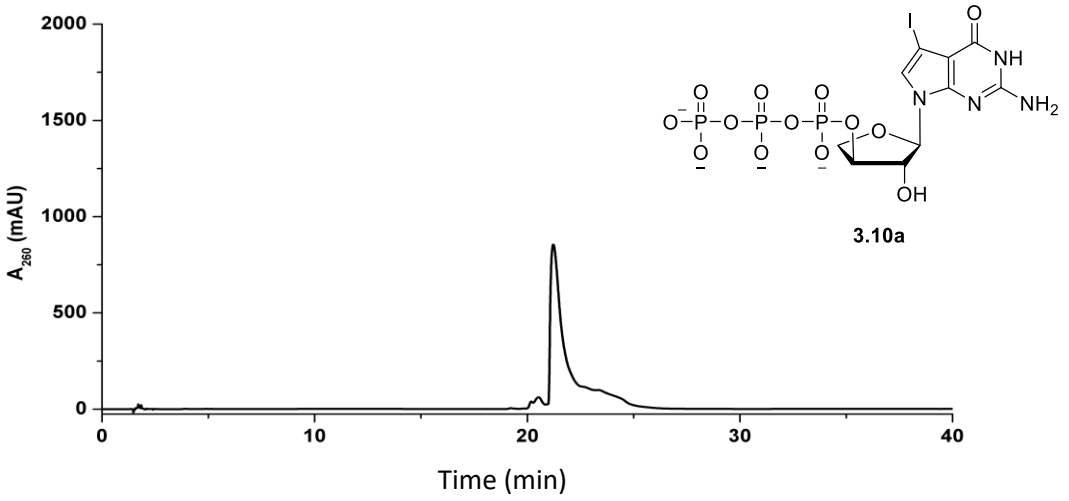
HPLC analysis of **2.6a-d** (MeCN/0.1M TEAA from 0 to 50%).

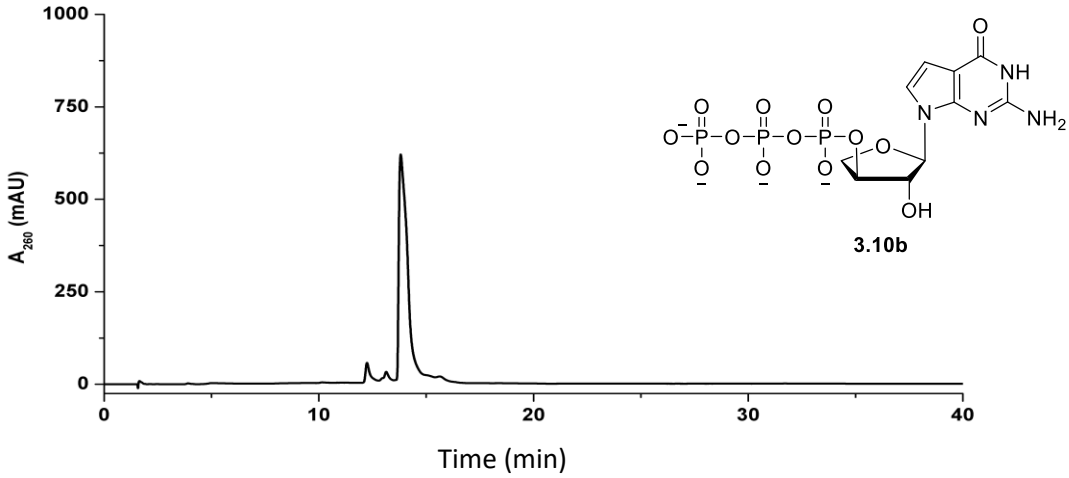


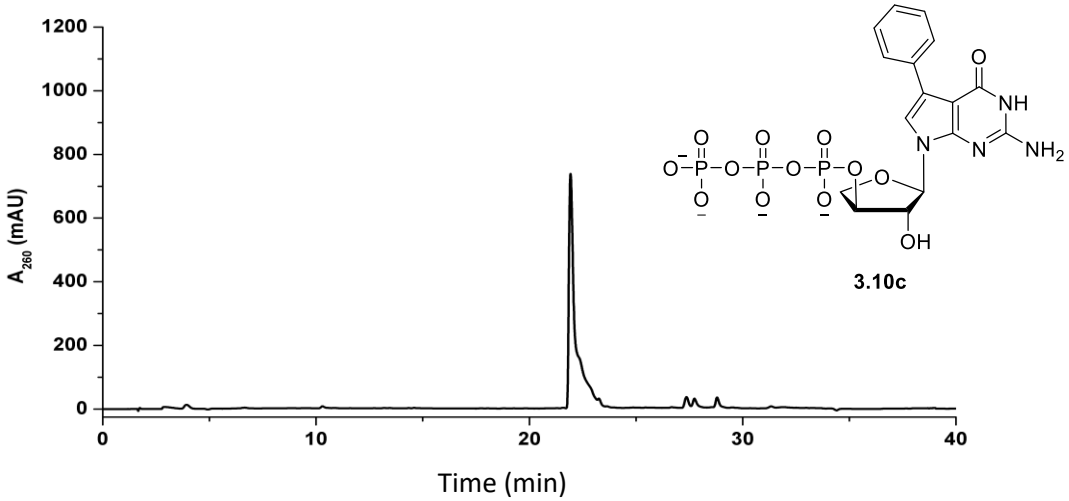
HPLC analysis of the TNA nucleoside 3'-monophosphate (**3.8a-c**).



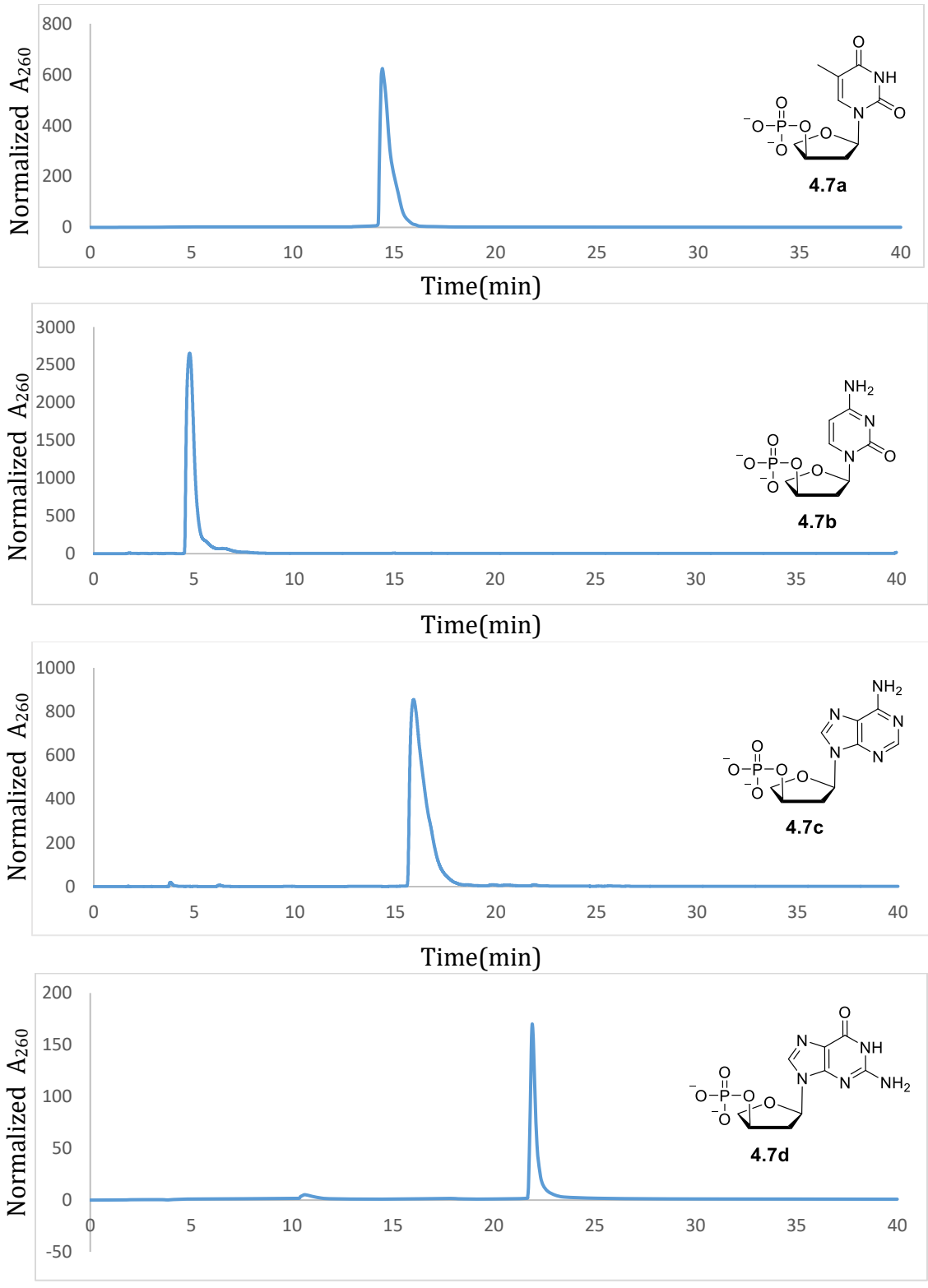
HPLC analysis of the TNA nucleoside 3'-phosphor-(2-methyl)imidazolide (**3.9a-c**).



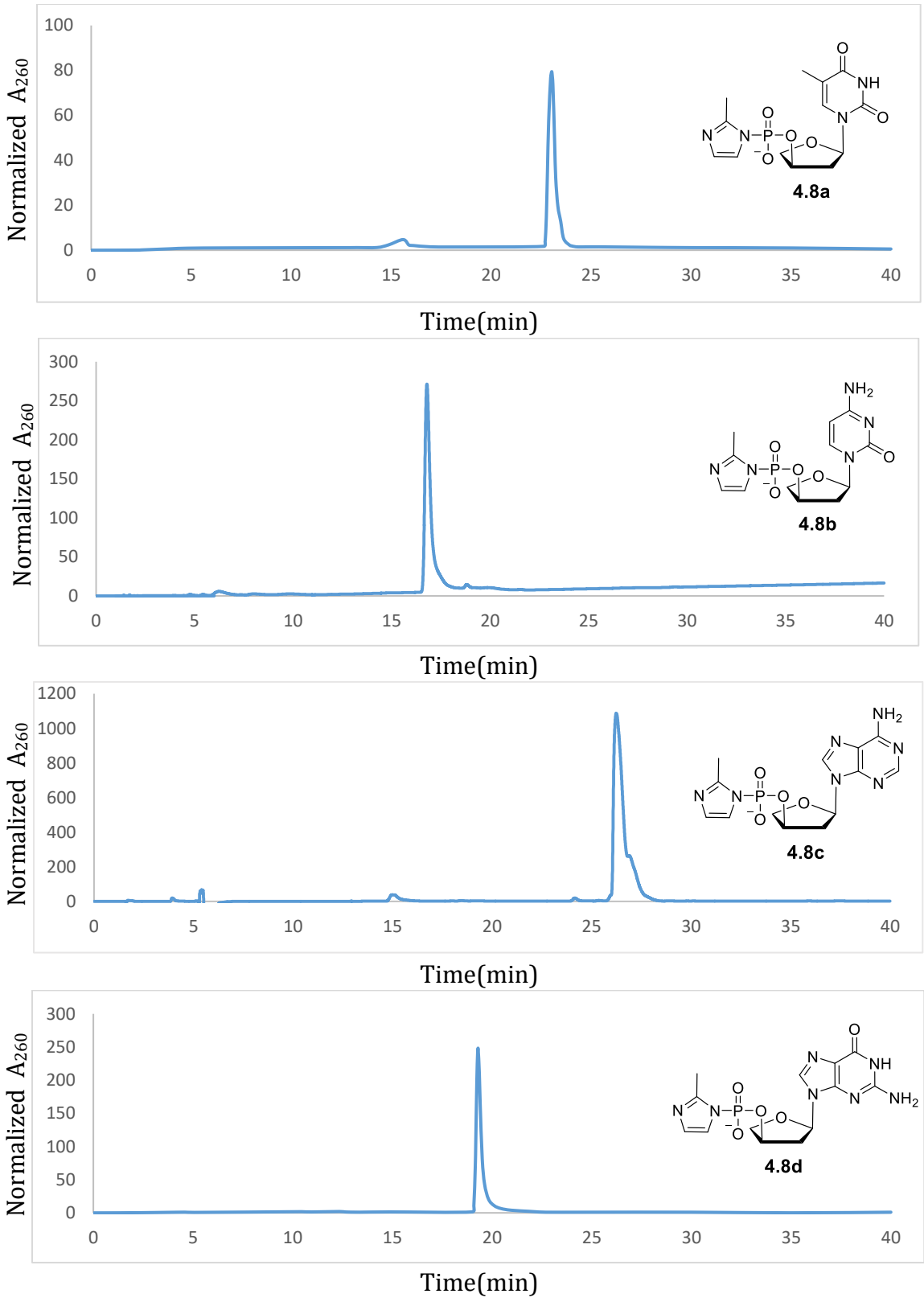




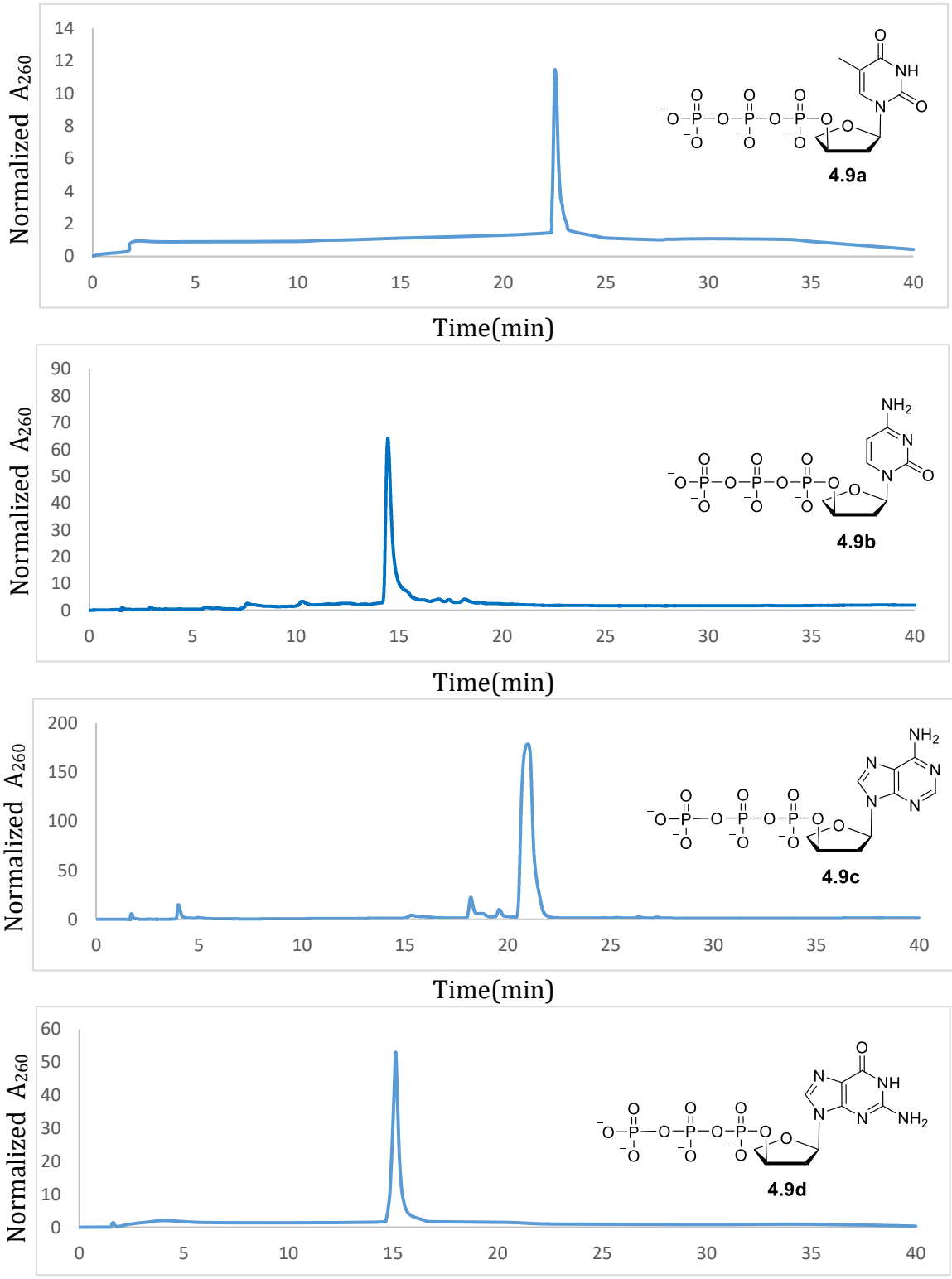
HPLC analysis of the TNA nucleoside 3'-triphosphate (**3.10a-c**).



HPLC analysis of **4.7a-d** (MeCN/0.1M TEAA from 0 to 50%).



HPLC analysis of **4.8a-d** (MeCN/0.1M TEAA from 0 to 50%).



HPLC analysis of **4.9a-d** (MeCN/0.1M TEAA from 0 to 50%).