UNIVERSITY OF CALIFORNIA,
IRVINE

Synthesis of α-L-Threofuranosyl Nucleosides, Phosphoramidites, and Triphosphates for Synthetic Biology

DISSERTATION

Submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Pharmaceutical Sciences

by

Saikat Bala

Dissertation Committee:
Professor John C. Chaput, Chair
Professor Robert Spitale
Professor Jennifer A. Prescher

2019

DEDICATION

To my parents, Mrinal Bala and Rita Bala, who are my first teachers and taught me the value of education and critical thought.

To all my teachers, the reason of what I have become today.

We not only want to know how nature is (and how her transactions are carried through), but we also want to reach, if possible, a goal which may seem utopian and presumptuous, namely, to know why nature is such and not otherwise.

Albert Einstein

Take up one idea. Make that one idea your life - think of it, dream of it, live on that idea. Let the brain, muscles, nerves, every part of your body, be full of that idea, and just leave every other idea alone. This is the way to success.

Swami Vivekananda

TABLE OF CONTENTS

Page

LIST OF ABBREVIATIONS

AcOH	acetic acid
Ac_2O	acetic anhydride
anh.	anhydrous
Aq	aqueous
atm	atmosphere
Bn	benzyl
B^{PG}	protected base
br s	broad singlet
BSA	*N*,*O*-Bis(trimethylsilyl)acetamide
^oC	degrees Celsius
13C NMR	carbon nuclear magnetic resonance spectroscopy
cat	catalytic
$CDCl_3$	deuterated chloroform
cm	centimeter
conc	concentrated
δ	chemical shift (ppm)
d	doublet
dd	doublet of doublet
DCM	dichloromethane
DIPEA	diisopropylethylamine
DMAP	dimethylaminopyridine
DMF	dimethylformamide

DMSO	dimethylsulfoxide
DMT-Cl	dimethoxytrityl chloride
DNA	deoxyribonucleic Acid
EtOAc	ethyl acetate
Et_3B	triethylborane
Et_3N	triethylamine
g	gram(s)
GCMS	gas chromatography mass spectrometry
1H NMR	proton nuclear magnetic resonance spectroscopy
HPLC	high-performance liquid chromatography
Hz	hertz
J	coupling constant
L	liter
m	multiplet
M	molar
MALDI-TOF	matrix assisted laser desorption ionization time of flight
mg	milligram(s)
µm	microgram(s)
MHz	mega Hertz x
min	Minutes
mL	milliliter
mM	millimolar
mmol	millimole(s)

μmol	micromole(s)
N	normal
NBS	*N*-bromosuccinimide
NCS	*N*-chlorosuccinimide
nm	nanometer
NMR	nuclear magnetic resonance
Pd/C	palladium-on-carbon
ppm	parts per million
*p*TsCl	*p*-toluenesulfonyl chloride
q	quartet
quin	quintet
*R*f	ratio of fronts
RNA	ribonucleic acid
s	singlet
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
t	triplet
TBAF	tetrabutylammoniumfluoride
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TLC	thin Layer Chromatography
TMSOTf	Trimethylsilyl trifluoromethanesulfonate

LIST OF FIGURES

Page

LIST OF SCHEMES

Page

LIST OF TABLES

ACKNOWLEDGMENTS

- I express my sincere gratitude to **Prof. John Chaput** for including me in his group. I thank him for his valuable suggestions, fruitful discussion, constant encouragement and faith in me and my work. His enthusiasm in research, perpetual energy, motivation and commitment to the highest standard inspired and motivate me in my research. He was always accessible and willing to help me throughout my time in my Ph. D., which made my research experience smooth and rewarding. Thank you once again for mentoring me John.

- I am very thankful to my Ph.D. committee members **Prof. Robert Spitale** and **Prof. Jennifer Prescher** for their insightful comments for my research, development and guidance to the project's executions.

- I would like to thank our collaborators, **Dr. Jerry Joyce** (Salk Institute, San Diego**), Dr. Jack Szostak** (Harvard Med. School), **Dr. Martin Egli** (Vanderbilt University), **Dr. Piet Herdewijn** (KU Leuven) , **Dr. Muthiah Manoharan** (Alnylam Pharmaceutical ,Boston) for allowing me to a part of various cool projects.

- I would also like to thank **Prof. Richard Chamberlin** (the big boss) for all those critical and funny organic chemistry discussion in our famous "Chamberlin Meeting". That was the place where I learned lots of chemistry and applied those in my research. You are such an inspiration. I would like to extend my thank to **Mike** (Malibu) and **Peter** for showing us how to be dedicated to science. Thank you all.

- I would like to thank Dr. Sujay Sau ,Dr. Hui Mei and Dr. Jen-Yu Liao for their instruction and knowledge in chemistry. My greatest thanks also goes to my lab mates Cailen, Yajun, Eric, Li, Ali, Allan, Sunit and the undergrads for many years of companionship in research, fun, frustration and celebratory moments in academic life. My deepest thank to Arlene, our caring lab manager for helping me out in lots of different ways, especially in my writing skill. Besides, I would like to thank all the other current and previous lab members

for their help and time. It was really delighted to work in this big and diverse scientific family.

- A special thanks to Spitale lab members especially Sam, Chao, Miles, Mike, Sara, Dallen, Ying for all the discussion and fun throughout these years.

- In addition, a thank to Felix and Ben (UCI Mass facility), Dr. Dennison (NMR facility) and Dr. Ziller (X-Ray crystallography facility) for their endless help. Without them I would notbe able to finish my research project easily in this time period.

- I would like to thank my previous supervisors Prof. Sandeep Verma (M.Sc thesis supervisor, IIT Kanpur) and Prof. Subhendu Sekhar Bag (Summer Internship, IIT Guwahati) for taking me in their lab and help me to connect with the practical application of organic chemistry.

- I owe a great debt to Kaustuv da, Kanika di and Suman da for all the support and fun, mainly those Friday night funs, foods and whiskeys. I do not think I could survive my graduate school without you guys. You guys have become my family here.

- My deepest thank to my seniors from Arizona too. Manas da, Ghonta da, Poulami di, Sudipta da, Tufan da, Chandrani di, Raja da, thank you guys for your love and friendship.

- I would like to thank my non-UCI friends for helping me in lots of ways. Ghora, Indrani, Manali, Gagan, Priyanka, Soumi, Wrickban, and Jash. You people are awesome! Thanks for being my friends for such a long time.

- I would like to thank my family especially my father, Mrinal Kanti Bala, my mother, Rita Bala for their love, support and all sacrifices. It would be impossible for me to here without you. A special thanks to my beloved brother Diganta for his support and questions. You guys are awesome.

- My thanks extended to the people from “Siliguri” too. Thanks for supporting me and making me a part of your life.

- I would also like to convey my thank to my fiancé Deborpita for supporting me in my difficult days. You were always there with me when I needed help. You never got bored whenever I talked about my research with you. You have a masters in history but I am sure you have lots of ideas about chemistry, mainly the Nucleic Acids world by now. Thanks!

CURRICULUM VITAE

Saikat Bala **Fall 2019**

Ph.d. Candidate in pharmaceutical Sciences (949) 346-6457
University of California, Irvine saikatb@uci.edu
Irvine, CA 92697 www.linkedin.com/in/saikat-bala/

Education

University of California-Irvine, Irvine, CA 09/2015-10/2019
Arizona State University, Tempe, AZ 08/2014-09/2015
Ph.D. in pharmaceutical Sciences

Indian Institute of Technology, Kanpur, India 06/2012-05/2014
Masters in Chemistry

University of Calcutta, India 06/2009-05/2012
Bachelors in Chemistry

Research Experience

Doctoral Research, University of California, Irvine 09/2015-06/2018
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. Synthesized 2'-deoxy TNA0nucleoside triphosphate for better understanding of polymerase mechanism of action by X-Ray crystallography. 4. Optimized the oligonucleotide (TNA, GNA, L-RNA, FANA) synthesis by solid phase and Involved the TNA oligonucleotides synthesis by chemical and enzymatic ligations.5.

Research Intern, IIT Kanpur, India 08/2013-04/2014
Designed a liquid phase synthesis of a peptide sequence (YIGSD and YIGSR) found in laminine and studied their conformational activities using different microscopic techniques (AFM, SEM, OM)

Research Intern, IIT Guwahati, India 03/2013-07/2013
Synthesized a triazolyl donor and acceptor mediated unnatural base-pair system in natural nucleic acid using Cu-catalyzed Click chemistry (CuAAC) that introduce base pair stability with fluorescence activity.

Patent

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

Publications *(† Equal Contribution)*

1. Liao, J. Y.;†, **Bala. S** .;† Ngor, A.; Yik, E. J.; Chaput , J. C. P(V) reagents for the scalable synthesis of Natural and Modified Nucleoside Triphosphates. *Journal of the American Chemical Society* , **2019**, *141*, 13286- 13289.

2. Horning , D.; **Bala, S.;** Chaput, J. C .; Joyce, G. RNA-Catalyzed Polymerization of Deoxyribose, Threose, and Arabinose Nucleic Acids. *ACS synthetic biology* . **2019**, *8*, 955-961

3. McCloskey, C.; Liao, J. Y,.; **Bala, S**.; Chaput, J. C. Ligase mediated TNA-synthesis on TNA template. *ACS synthetic biology***, 2019 ,** *8,* 282-286

4. **Bala, S**†.; Liao, J. Y†.; Zhang, L.; Tran, C.; Chim, N.; Chaput, J. C. Synthesis of 2'-deoxy- α -L threofuranosyl Nucleoside triphosphate. *Journal of Organic Chemistry***, 2018 ,** *17*, 83(16), 8840-8850

5. Mei, H.; Liao, J, Y.; Jimenez, R. M.; Wang, Y.; **Bala, S.;** McCloskey, C.; Switzer, C.; Chaput, J. C. Synthesis and Evolution of a TNA-aptamer Bearing 7-deaza-7-substituted Guanosine residues. *Journal of the American Chemical Society***, 2018 ,** 140, 5706-5713.

6. **Bala, S**†.; Liao, J. Y†.; Mei, H.; Chaput, J. C. Synthesis of α -L threofuranosyl Nucleoside monophosphate, 3'-Phospho(2-methyl) imidazolides and 3'-triphosphates. *Journal of Organic Chemistry*, **2017**, *82*, 5910-5916.

7. Sau, S. P.; Fahmi, N. E.; Liao, J. Y.; **Bala, S.;** Chaput, J. C. A Scalable Synthesis of α-L-Threose Nucleic Acid Monomers. *Journal of Organic Chemistry*, **2016,** *81*, 2302-2307

8. Anosova, I.; Kowal, E. A.; Sisco, N. J.; Sau, S. P.; Liao, J. Y.; **Bala, S.;** Rozners, E.; Egli, M.; Chaput, J. C .; Van Horn, W. D. Structural Insights into Conformational Differences between DNA/TNA and RNA/TNA Chimeric Duplexes. *ChemBioChem*, **2016**, *17*, 1705-1708,

9. 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). *Biopolymers,* **2016,** *107*,1-7.

10. Culbertson, M. C.; Temburnikar, K.; Sau, S. P.; Liao, J. Y.; **Bala, S.;** Chaput, J. C. Evaluating TNA Stability Under Simulated Physiological Conditions. *Bioorganic & Medicinal Chemistry Letters*, **2016**, *26*, 2418-2421.

Presentation and workshops

1. XXIII International Roundtable of Nucleosides, Nucleotides and Nucleic Acids (IRT), UC San Diego, 2018 ; Poster Presentation: " Synthesis of chain terminating threoufuranosyl neucleoside triphosphates"
2. American Chemical Society (ACS) National Meeting, San Francisco, 2017; Poster Presentation: "A pipeline for generating TNA-triphosphates for the in vitro selection"
3. Graduate Visitation Day, UC Irvine, 2016; Poster Presentation: "Synthesis of α-L-Threofuranosyl Nucleoside 3'-Monophosphates, 3'-Phosphoro(2-Methyl)imidazolides, and 3'-Triphosphates"

Teaching Experience

Teaching Assistant, Department of Pharmaceutical Sciences, University of California, Irvine
Medicinal Chemistry Laboratory and Discussion (Winter 2019 and 2017)
Biopharmaceutics and nanomedicine Laboratory and Discussion (Fall 2017)

Teaching Assistant, Department of Chemistry, Arizona State university
Organic Chemistry Laboratory CHM 238 (Fall 2014)
Organic Chemistry Laboratory CHM 238 (Spring 2015)

Mentorship

Phuong Tran —— Undergraduate Student (2017-2018)

References

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

Jennifer A. Preshcer
Department of Chemistry
UC Irvine
Tel: (949) 824-1706
Email: jpresche@uci.edu

Robert Spitale
Department of Pharmaceutical Sciences
UC Irvine
Tel: (949) 824-7655
Email: rspitale@uci.edu

ABSTRACT OF THE DISSERTATION

Synthesis of α-L-Threofuranosyl Nucleosides, Phosphoramidites, and Triphosphates for Synthetic Biology

By

Saikat Bala

Doctor of Philosophy in Pharmaceutical Sciences

University of California, Irvine, 2019

Professor John C Chaput, Chair

A major goal of synthetic biology is to recapitulate emergent properties of life. Despite a significant body of work, a longstanding question that remains to be answered is how such a complex system arose? In this dissertation, synthetic nucleic acid molecules with alternative sugar-phosphate backbones were investigated as potential ancestors of DNA and RNA. Threose nucleic acid (TNA) is capable of forming stable duplex structure with complementary strands of itself and RNA. This provides a plausible mechanism for genetic information transfer between TNA and RNA. Therefore TNA has been proposed as a potential RNA progenitor. Using molecular evolution, functional sequences were isolated from a pool of random TNA molecules. This implicates a possible chemical framework capable of crosstalk between TNA and RNA. Further, this shows that heredity and evolution are not limited to the natural genetic system based on ribofuranosyl nucleic acids. Structurally different from DNA and RNA, TNA was further explored as a source of nuclease

resistant affinity reagents in therapeutic application. Since TNA molecules are naturally unavailable, synthetic approach has been developed previously for initial research explorations. A detailed synthetic optimization have been investigated in this dissertation to provide quantitative amount of TNA substrates for the downstream research applications in Synthetic biology.

Chapter 2 describes a highly optimized synthetic approach for large scale preparation of α-L-threose nucleoside and phosphoramidites by ten reaction steps form L-ascorbic acid. This new strategy overcomes several shortcomings from original method presented by Dr. Eschenmoser and co-workers and allow us to quantitatively synthesize several key intermediate precursors for the α-L-threose nucleoside phosphoramidites 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 3 discusses the stepwise approach of 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 4 discusses 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 5 discusses the synthesis of a novel organic pyrophosphate reagent which was used to develop a highly optimized and HPLC free approach for gram scale synthesis of both natural (DNA, RNA) and modified (TNA, L-DNA etc.) nucleoside triphosphates (NTPs). Here a new pyrene substituted pyrophosphate reagent was developed which introduced hydrophobicity in nucleoside triphosphate analogies and made it easier to purify by silica gel column chromatography and we can scale it up to a gram synthesis of triphosphates. This is a very cost effective approach to obtain high quality triphosphates. The synthesized products are comparable to commercial products in terms of purity and biological function.

Chapter 1

Background

1.1 Introduction

DNA and RNA are two well-known biopolymers that can store and transfer genetic information.[1-2] RNA has the natural ability to store genetic information, as seen with RNA-based virus genomes,[3] and to perform enzymatic and regulatory functions, observed with riboswitches,[4] ribosomes,[5] molecular recognitions,[6] therapeutics,[7] self-replicating introns,[8] and ribozymes.[9] Since RNA can perform these two essential activities to support life, it may be hypothesized that life originated from RNA, commonly referred to as the RNA world hypothesis, in which RNA is a progenitor in the origins of life on Earth.[10]

This hypothesis has motivated many scientists to investigate RNA as the original genetic polymer of life in the evolutionary path of life. Researchers have systematically studied the chemical structure of RNA using artificial genetic polymers related to RNA. Such research is based on several aspects that include predecessors that: (i) are able to base-pair with natural nucleic acid polymers (RNA or DNA) and itself for the transfer of genetic information between two stages of life, (ii) may be synthesized easily under plausible pre-biotic conditions and in which its chemical structure is simple, and (iii) have chemical structures that are chemically derived from RNA.

The aims of synthetic genetics, a sub-discipline of synthetic biology is to develop artificial genetic polymers, termed XNAs (xeno nucleic acids), in which the letter X refers to the sugars or its substitutes. These XNAs can replicate *in vitro* and eventually in a model protocell.[11] On the basis of backbone moiety, there are different types of XNAs (Figure 1.1).

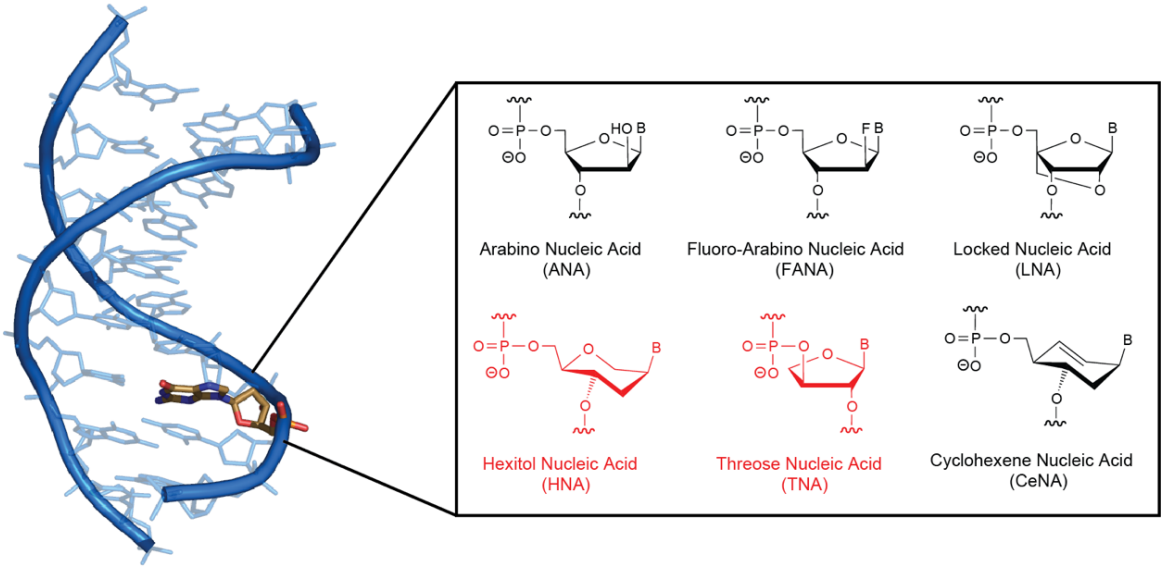


Figure 1.1 XNAs are classified by their sugar compositions and may be categorized according to the number of atoms, which include the 4-carbon sugar (Threose in TNA, the "locked" ribose analog in LNA), the 5-carbon sugar (Aabinose in ANA, 2'-Fluoroarabinose in FANA), and 6-membered sugar (cyclohexene in CeNA, anhydohexitol in HNA)

Among these XNAs, HNA and TNA aptamers were first discovered in 2012. FANA and ANA are likely candidates for aptamer validation due to their structural similarity to RNA, whereas LNA and CeNA are not, due to their structurally restrictive conformation and the difficulty in synthesizing the latter.[12,13] Thus, four XNAs (TNA, HNA, FANA and ANA) are viewed as suitable candidates capable of undergoing evolution. Most XNAs are amenable to directed evolution via a process called systematic evolution of ligands by exponential enrichment (SELEX), but therapeutic value varies based on the commercial availability of their monomers and biostability against nucleases, FANA and ANA are commercially available and are relatively resistant against high nuclease degradation. However, TNA and HNA have shown comparatively better nuclease resistance ability than FANA and ANA, but they must be synthesized chemically in house from the starting nucleosides.

In our lab, we have focused our efforts on TNA synthesis and evaluation. TNA is an artificial genetic polymer that was developed by Dr. Albert Eschenmoser as part of his search for nucleic acid structures that were related to DNA/RNA.[14] TNA is composed of a repeating four-carbon threose backbone (Figure 1.2), which is structurally different from the natural five-carbon ribose backbone in DNA and RNA. As a result, the backbone linkage is one atom shorter than that observed in DNA and RNA.

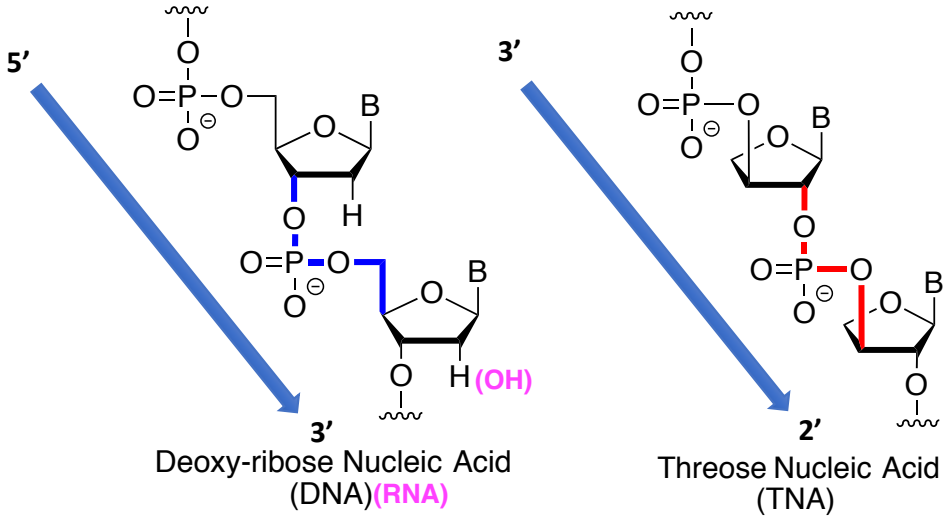


Figure 1.2 Structural differences between the linearized backbone of DNA (RNA) with TNA.

Why is TNA important? TNA adopts an A-type helix when it forms Watson-Crick base pairs with complimentary strands of DNA, RNA, and itself (Figure 1.3).[15] This implies that TNA is able to transfer genetic information between itself and with the natural genetic polymers. TNA is also very stable to biological nucleases.[16] We have investigated the stability of two different TNA oligonucleotides in human liver microsome (HLM) and 50% human serum (HS), which are biologically relevant matrices that are rich in nucleases and are

commonly used to test nucleic acid stability and nuclease resistance. TNA shows high nuclease resistance ability in both HLM and HS (Figure 1.4). Therefore, TNA can form strong homo and hetero duplexes and is very stable to biological nucleases.[17]

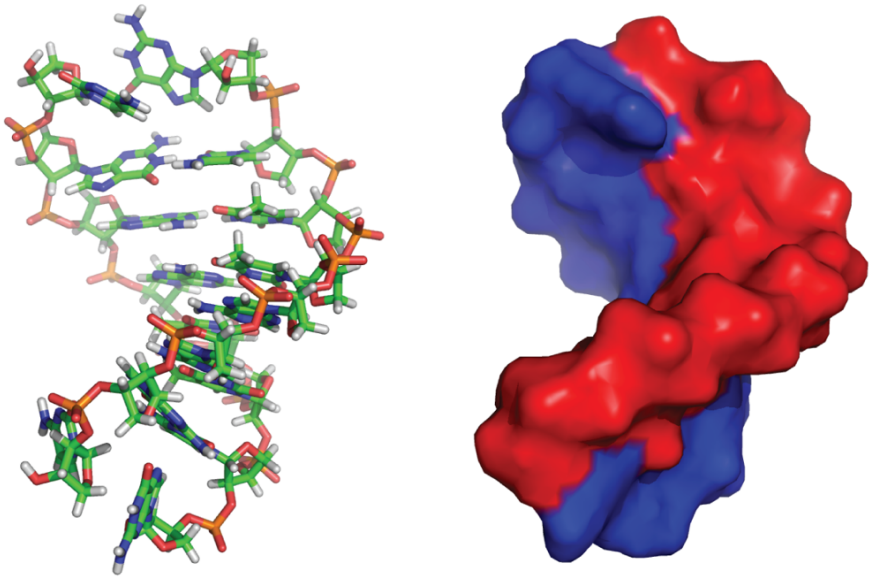


Figure 1.3 Solution NMR spectral analysis shows that TNA forms A-type of helix when it forms a duplex with a complimentary strand of TNA.

(Figure was used with permission from *Nature Chemistry*, **2012**,*4*,183)

TNA is also capable of *in vitro* selection. Engineered polymerases can transcribe functional TNA aptamers selected from a large number of sequences. Moreover, TNA overcomes the problem of post-selection modifications that plagues the field of RNA, as the TNA backbone is completely resistant to nuclease degradation.[18] Unmodified nucleic acids are very susceptible to nuclease attack and are therefore generally unsuitable as therapeutic agents. RNA aptamers need different modifications to avoid nuclease attack, but TNA does

not need such extra modifications. For these reasons, we are extremely interested in TNA and its potential in biological applications.

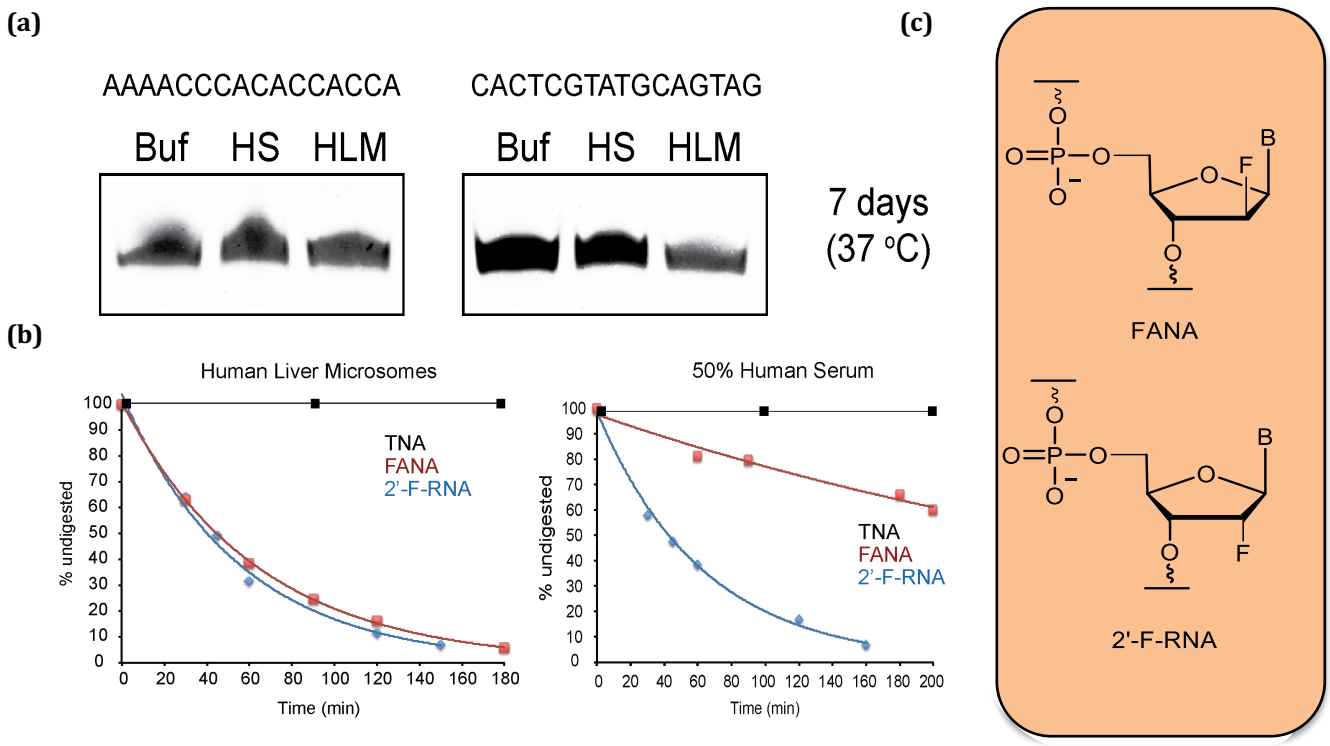


Figure 1.4 Comparative nuclease stability assays performed in human serum (HS) and human liver microsomes (HLM). **(a)** TNA polymers remain undigested after 7 days of incubation at 37°C in 50% human serum or 0.5 mg/mL human liver microsomes. Buf denotes buffer only. **(b)** FANA and 2'-F-RNA both rapidly degrade under simulated biological conditions. **(c)** Chemical structure of FANA and 2'-F-RNA.

(Figure was used with permission from *Bioorg. Med. Chem. Lett*, **2016**,*26*, 2418-2421)

TNA is an unnatural nucleic acid that is not commercially available. Thus, we synthesize TNA monomers and polymers chemically in lab. All the monomers shown in figure 4 are used for all the projects in our lab. The top row corresponds to the TNA-2-phosphoramidite with 3' –O-DMT protection (tN-amidites, N = A,T,G,C). These listed

compounds (Figure 1.5A) are used for making TNA oligonucleotides on an automated DNA synthesizer via solid-phase synthesis. All the bases are protected with respective protecting groups to avoid side reactions. The compounds (Figure 1.5B) correspond to the TNA triphosphates (tNTPs, N = A,T,G,C), which are very essential for enzymatic TNA oligomer synthesis, enzyme engineering, and aptamer selection.[19]

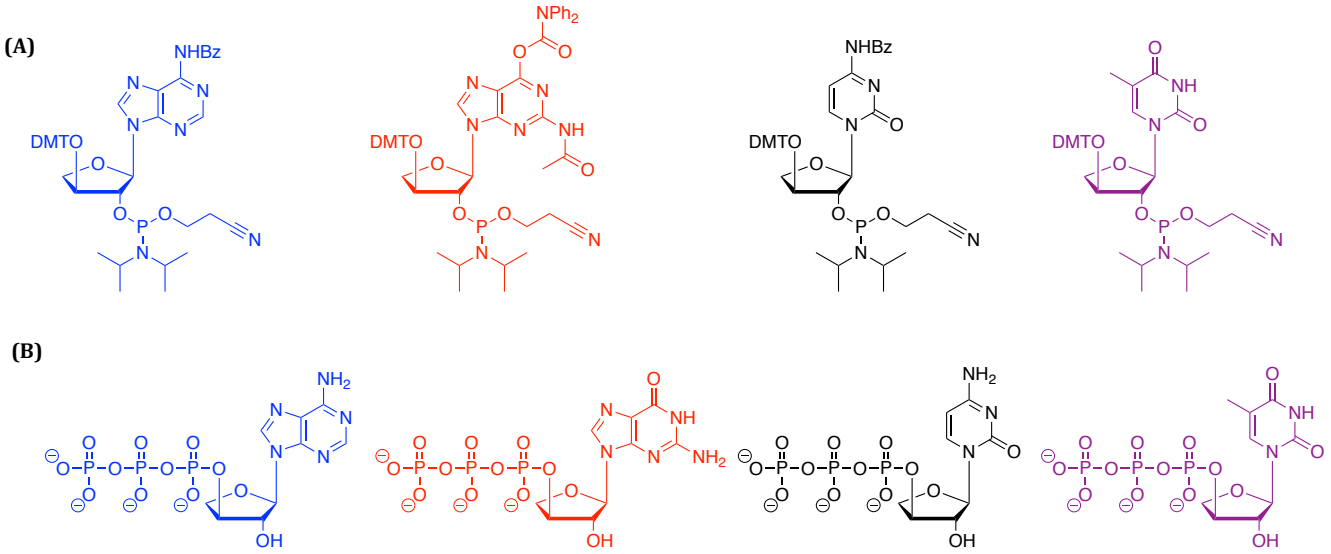


Figure 1.5 Different TNA monomers important for TNA biology (A) TNA-2'-phosphoramidite monomers with the 3'-OH are DMT protected. The nucleobases are also protected with respective protecting groups. (B) TNA-3'-triphosphates monomers (tNTPs)

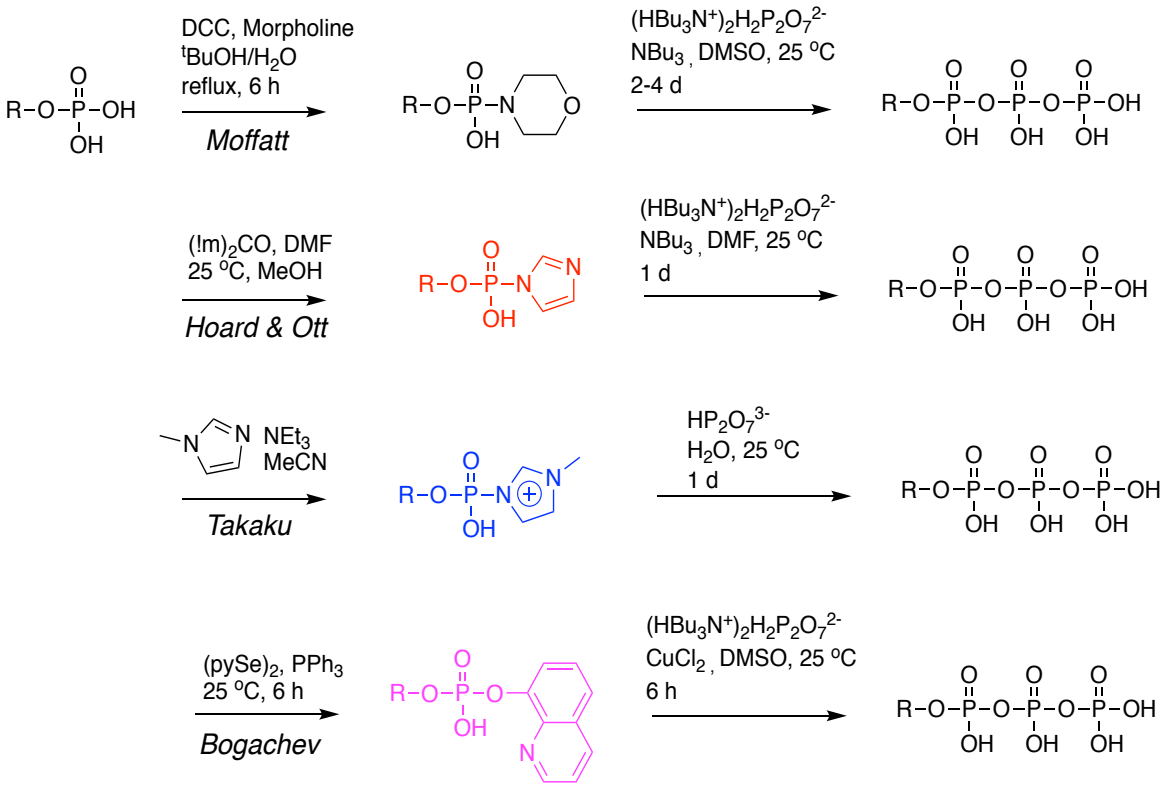
In my thesis work, I have developed a generalizable approach to the large scale synthesis of nucleoside triphosphates (NTPs). Natural and modified nucleic acid triphosphates impact nearly every major aspect of healthcare research from DNA sequencing to drug discovery. Various research groups in this field and literatures have attempted to demonstrate a strategy to nucleoside triphosphate synthesis that is

generalizable, high-yielding and cost-effective. However, this remains an unsolved problem, as some methods discovered have proven to be effective for only some substrates and to be only moderately effective for other substrates. Thus far, not one synthesis strategy has been found to be universally satisfactory for all substrates.

The synthesis, isolation, and storage of nucleoside triphosphates (NTPs) are difficult for several reasons[24]. (1) Most methods for the synthesis of NTPs involve the combination of charged ionic reagents (e.g., pyrophosphates) with more lipophilic substrates (e.g., protected nucleosides). As a result, it is difficult to find appropriate solvents. Moreover, purification involves the isolation of charged water-soluble products from a mixture of hydrophobic and hydrophilic impurities. (2) Nucleoside triphosphates (NTPs) are not robust, and it is difficult to make quantitative generalizations of their rates of decomposition. Usually the rate of hydrolysis of the product is accelerated in the presence of acids or bases, even if stored in lower temperature. But most recently, we have found a protocol that can prevent the degradation of NTPs for a long period of time.[25] Here we demonstrated we can store triphosphate products in Tris buffer (pH= 8.2) in -80 ^{o}C for more than 6-8 months without any degradation.

The most widely used approach to the synthesis of nucleoside triphosphates (NTPs) is the nucleophilic attack of the pyrophosphates on an activated nucleoside monophosphate (NMP). There are several early works that describe the phosphorylation of nucleosides and their activation techniques (Scheme 1.1).[26-29] It is shown that the treatment of an inorganic pyrophosphate reagent generates the desired product. The use of the nucleoside

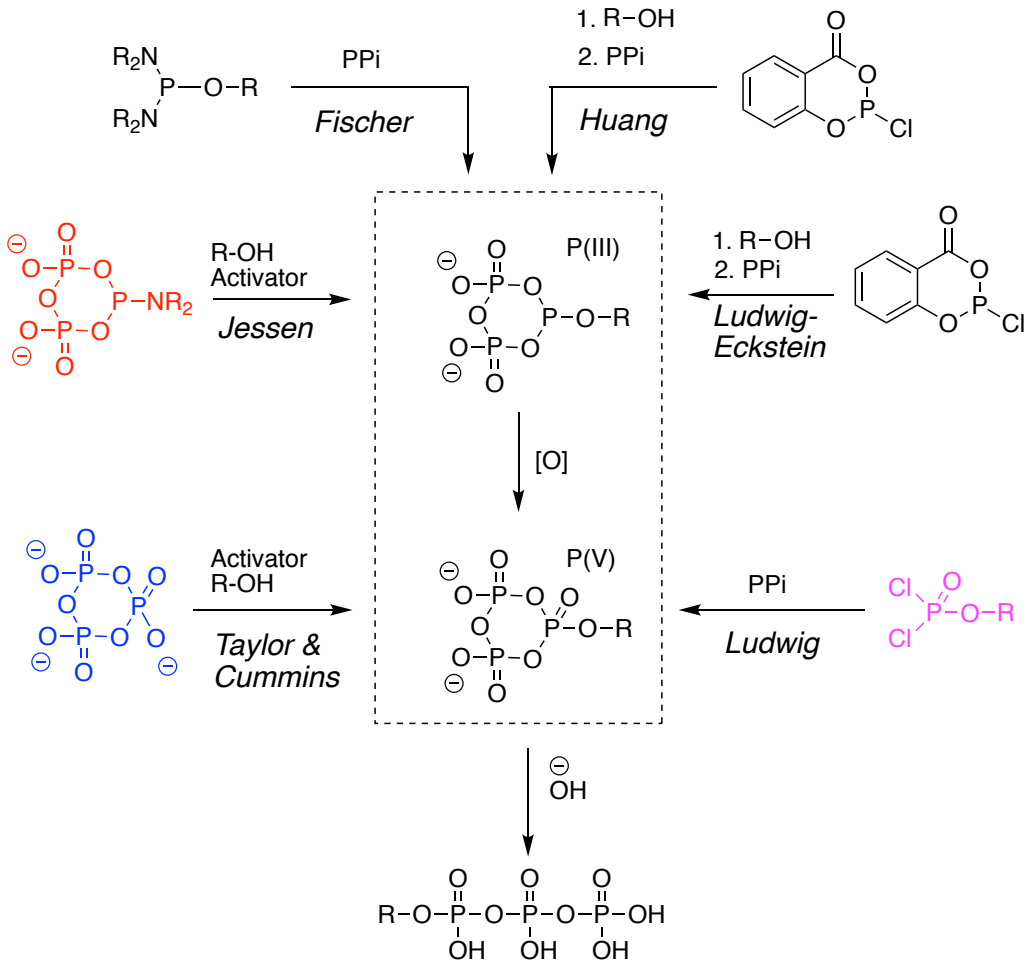
phosphoramidates is one of the most useful strategy for the activation of monophosphates. Here the products are purified by high performance liquid chromatography (HPLC) or strong ion exchange chromatography.



Scheme 1.1 Different chemical methods for the activation of nucleoside monophosphates (NMPs) and the synthesis of nucleoside triphosphates from NMPs. In these methods, dichloro-phosphate is used for the phosphorylation and synthesis of the nucleoside monophosphates.

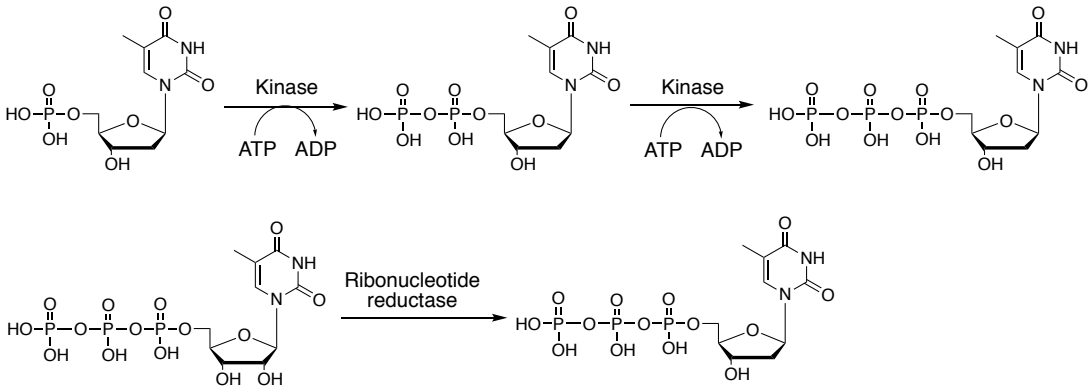
Another common synthetic strategy is the use of *activated phosphites or phosphoramidites* derived from the nucleosides. This type of technique involves an *in situ* ring opening of cyclotriphosphate intermediates (Scheme 1.2).[30-39] This method was found to be most efficient in the phosphorylation of the less reactive hydroxyl (-OH) group of the

nucleoside (as seen with the secondary alcohol of TNA). The P(III) in the reagent can phosphorylate the secondary hydroxyl group in a more facile manner. Experimental facts demonstrate that the attack of the nucleophile at the ring opening step results in a linearized product. Yet, HPLC is the best way to isolate the products further.



Scheme 1.2 Synthesis of nucleoside triphosphates using activated phosphites or phosphoramidite derivatives derived from nucleosides. Here the triphosphates are made via a cyclotriphosphate intermediate.
(Figure was adapted from *J. Am. Chem. Soc.* **2019**, *141*, 15013-15018)

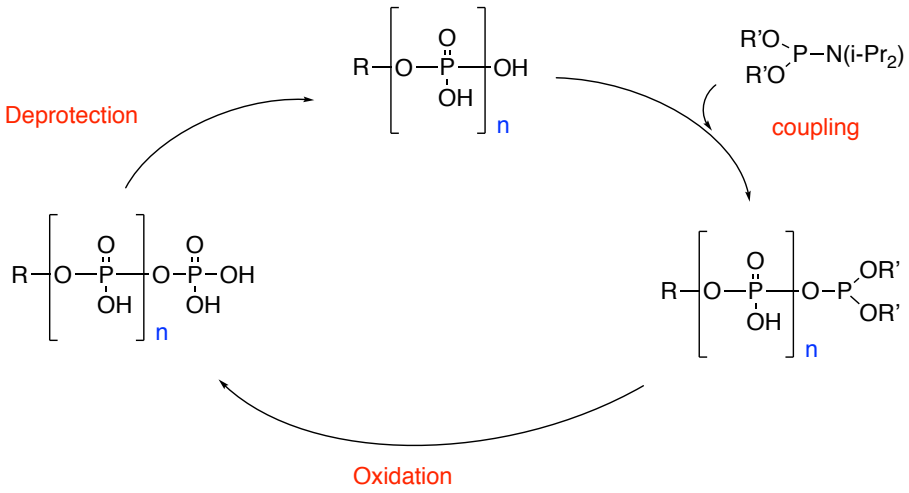
Most of the commercial sources obtain their nucleoside triphosphates (NTPs) via *biocatalytic methods***.** These methods are extremely useful for the synthesis of natural triphosphates (A, C, G, and T). Such is made possible by the enzymatic extraction of the desired product from their natural sources. It is observed that adenosine may be converted into ATP using a mixture of three enzymes: adenosine kinase, adenylate kinase, and acetate kinase.[40] Similarly, deoxyribonucleic triphosphates may be extracted from the corresponding ribonucleotide counterpart with ribonucleotide reductase (Scheme 1.3).



Scheme 1.3 Different biocatalytic syntheses of deoxyribonucleoside triphosphates (dNTPs).

The methods discussed earlier are all beneficial for the synthesis of nucleoside triphosphates, but none are sufficient for the synthesis of both natural and modified nucleoside triphosphates with equal efficiency. These biocatalytic methods all require HPLC or strong ion exchange chromatography for the purification of the desired product. Thus, they are not suitable for the large scalable syntheses of triphosphates. Sau. et al. developed an iterative synthetic approach for the synthesis of nucleoside triphosphates that bypasses

HPLC purification and is scalable (Scheme 1.4)[41]. Here they described an efficient phosphytilation to the nucleoside, followed by oxidation, hydrogenation, and the addition of a phosphate group one at a time. The phosphorylated product can be purified by silica gel column chromatography.

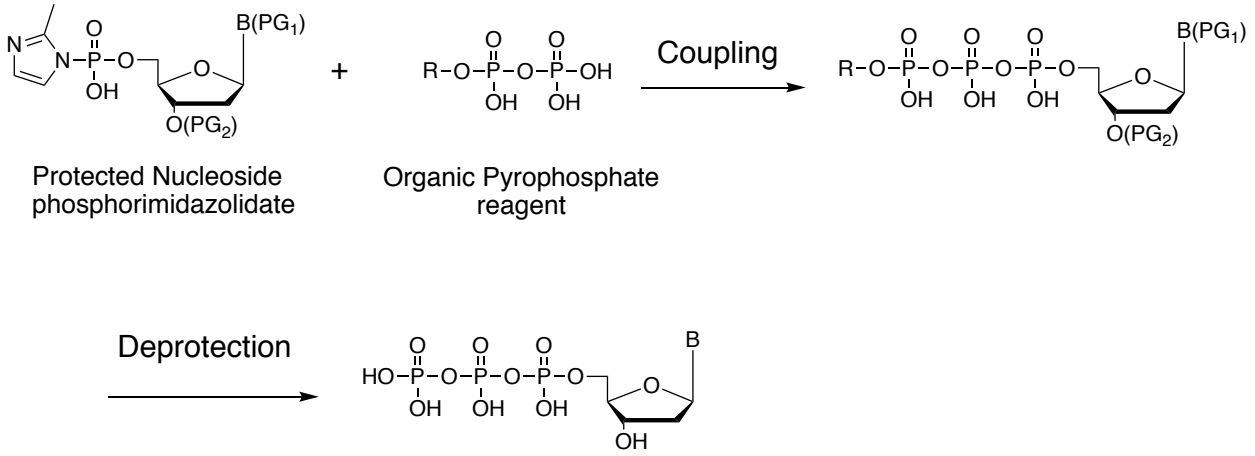


Scheme 1.4 Schematic synthetic diagram for nucleoside triphosphate synthesis using an iterative synthetic method.
(Figure was used with permission from *Org. Lett*, **2018** ,*19*, 4379-4382)

Using this iterative synthetic method, gram scale product may be obtained with greater ease, as noted with the lack of HPLC usage, but the overall synthesis process is heavily prolonged and contains difficult steps that require silica gel purification. Overall, the products were contaminated with heavy metal ions, which nulls the polymerase activity in the enzymatic reaction.

Recently, we developed a novel P(V) pyrophosphate reagent that can synthesize both natural and unnatural nucleoside triphosphates on the gram scale.[42] The idea originates

from use of the activated monophosphate method as mentioned earlier, but here they developed the pyrophosphate reagent with a polycyclic aromatic (PA) moiety, in which the pyrene group attributes hydrophobicity to the triphosphate precursor (Scheme 1.5). This hydrophobic precursor aids purification via silica gel column chromatography.



Scheme 1.5 General synthetic scheme for synthesizing NTPs using the organic pyrophosphate method.

In summary, methods for synthesizing nucleoside triphosphates initially evolved rapidly during the 1950s, though efforts since then have tapered off, with noticeable languish the last two decades. A revival of interest in the synthesis of these important molecules would be appropriate and pivotal to the nucleic acid chemistry field. The thesis described therein is the development of several novel and generalized techniques to obtain nucleosides , phosphoramidites and triphosphates in high quality and large quantity.

References

1. Eschenmoser, A. *Science*, **1999**, *284*, 2118-2124.
2. Schoning, K.; Scholz, P.; Guntha, S .; Wu, S.; Krishnamurthy, R.; Eschenmoser, A . *Science*, **2000**, *290*, 1347-1351.
3. Anosova, I.; Kowal, E. A.; Dunn, M.R.; Chaput, J. C.; Van Horn, W.D.; Egli, M. *Nucl. Acids Res.* **2016**, *44*: 1007-1021.
4. Inoue, T.; Orgel, L. E. *J. Mol. Biol.* **1982**, *162*, 201-217.
5. Chaput, J. C.; Ichida, J. K.; Szostak, J. W. *J. Am. Chem. Soc.*, **2003**, *125*, 856-857.
6. Dunn, M. R.; Chaput, J. C. *Curr. Protoc. Nucleic Acid Chem.,* **2014,** 9.8.1-9.8.19
7. Mondal, M.; Breaker, R. *Nature Reviews Molecular Cell Biology,* **2004,** *5*, 451-463.
8. Larsen, A. C.; Dunn, M. R.; Hatch, A.; Sujay, S. P.; Youngbull, C.; Chaput, J. C. *Nature Communications*, **2016,** 7: 11235.
9. Gilbert, W. *Nature* , **1986,** *319*, 618.
10. Boryski, J.; Manikowski, A. *Nucleosides & Nucleotides*, **1995**, *14***,** 287-290.
11. Yu, H.; Zhang, S.; Chaput, J. C. *Nature Chemistry*, 2012, *4*, 183-187.
12. 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. J.; Herdewijn, P.; Holliger, P. *Science*, **2012**, *336*(6079) 341-344.
13. Ebert, M. O.; Hoan, H. K.; Luther, A.; Krishnamurthy, R.; Eschenmoser, A., Jaun, B. *Helv. Chim. Acta*. **2002**, *85*, 4055-4073.

14. Chaput, J. C.; Yu, H.; Zhang, S. *Chemistry & Biology*, **2012**, *19*(11) 1360-1371

15. Culbertson, M. C; Temburnikar, K.; Sau, S. P.; Liao, J. Y.; Bala. S.; Chaput, J. C. *Bioorganic & Medicinal Chemistry Letters*, **2016,** *26(10)*, 2418–2421.

16. Dunn, M. R.; Otto, C.; Fenton, K. E. ; and Chaput, J. C. *ACS Chem. Biol.* **2016, *11*** (5): 1210-1219.

17. Eugene W. M.; Ng, David T.; Shima, Perry Calias.; Emmett T. Cunningham, Jr.; David R. Guyer.; Anthony P. Adamis. *Nature Reviews Drug Discovery*, **2006**, 5, 123

18. Larsen, A. C.; Dunn, M. R.; Hatch, A.; Sujay, S.P.; Youngbull, C.; Chaput, J. C. *Nature Communications*, **2016**, *7*, 11235.

19. Wu, X.; Delgado, G.; Krishnamurthy, R.; Eschenmoser, A. *Org. Lett*. **2002**, *4*, 1283-1286

20. Sau, S. P.; Fahmi, N. E.; Liao, J. Y.; Bala, S.; Chaput, J. C. *J. Org. Chem*., **2016,** *81,* 2302-2307.

21. Ludwig, J.; Eckstein, F. *J. Org. Chem*. **1989**, *54*, 631-635.

22. Yoshikawa, M.; Kato, T.; Takenishi, T. *Tetrahedron letters*, **1967**, *50*, 5065-5068.

23. Sau, S. P.; Chaput, J. C. *Bioorganic & Medicinal Chemistry Letters*, **2016,** *26 (14)*,3271-3273.

24. Burgess, K.; Cook, D. *Chem. Rev.* **2000**, 100(6), 2047- 2060

25. Zhang, S.; Zhang, N.; Blain, J, C.; Szostak, J. W. *J. Am. Chem. Soc.*, **2013**, *135*, 924-932.

26. Bala, S.; Liao, J. Y.; Mei, H.; Chaput, J. C. *J. Org. Chem*. **2017**, *82*, 5910–5916.

27. Ludwig, J. N.; Eckstein, F. *J. Org. Chem*. **1989**, *54*, 631.

28. Hoard, D. E.; Ott, D. G. *J. Am. Chem. Soc*. **1965**, *87*, 1785.

29. Moffatt, J. G.; Khorana, H. G. *Can. J. Chem*. **1961**, *39*, 649.

30. Moffatt, J. G. *Can. J. Chem.* **1964**, *42*, 599.

31. Takaku, H.; Konishi, T.; Hata, T. *Chem. Lett*. **1977**, 655.

32. Bogachev, V. S.; Russ. *J. Bioorg. Chem*. **1996**, *22*, 599

33. Fischer, B.; Boyer, J. L.; Hoyle, C. H. V.; Ziganshin, A. U.; Brizzolara, A. L.; Knight, G. E.; Zimmet, J.; Burnstock, G.; Harden, T. K.; Jacobson, K. A. *J. Med. Chem*. **1993**, *36*, 3937.

34. Nahum, V.; Zundorf, G.; Levesque, S. A.; Beaudoin, A. R.; Reiser, G.; Fischer, B., *J. Med. Chem.* **2002**, *45* (24), 5384-5396.

35. Caton-Williams, J.; Lin, L.; Smith, M.; Huang, Z. *Chem. Commun*. **2011**, *47* (28),8142-4.

36. Dutta, A. K.; Captain, I.; Jessen, H. J. *Top. Curr. Chem*. **2017**, *375* (3), 51.

37. Singh, J.; Steck, N.; De, D.; Hofer, A.; Ripp, A.; Captain, I.; Keller, M.; Wender, P. A.; Bhandari, R.; Jessen, H. *Angew. Chem. Int. Ed. Engl*. **2019**, *58* (12), 3928-3933.

38. Shepard, S. M.; Cummins, C. *J. Am. Chem. Soc*. **2019**, *141* (5), 1852- 1856.

39. Mohamady, S.; Taylor, S. D. *Org. Lett*. **2016**, *18* (3), 580-3.

40. Baughn, R. L.; Adalsteinsson, O.; Whitesides, G. M. *J. Am. Chem. Soc*. **1978**, *100*, 304.

41. Sau, S. P.; Chaput, J. C. *Org. Lett*. **2017**, *19*, 4379–4382.

42. Liao, J. Y.; Bala, S.; Ngor, A,; Yik, E, J; Chaput, J. C. *J. Am. Chem. Soc*. **2019,** *141*, 13286-13289

Chapter 2

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.

2.1 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.

2.2 Abstract of the Chapter

Recent advances in polymerase engineering have made it possible to copy information back and forth between DNA and artificial genetic polymers composed of TNA (α -L - threofuranosyl- (3′ ,2′) nucleic acid). This property, coupled with enhanced nuclease stability relative to natural DNA and RNA, warrants further investigation 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 constructing multigram quantities of TNA nucleosides that can be readily converted to nucleoside 2′ -

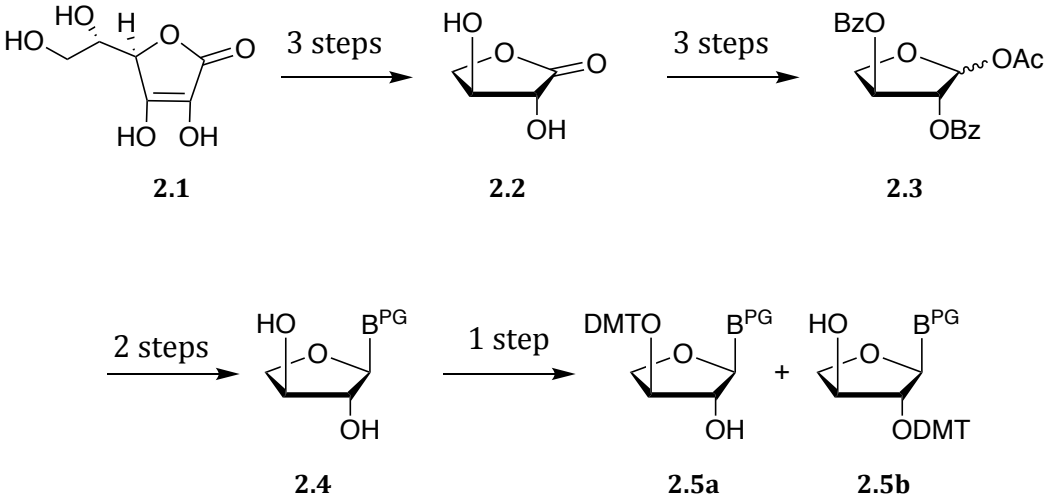
phosphoramidites or 3′ - triphosphates for solid-phase and polymerase-mediated synthesis, respectively. The synthetic protocol involves 10 chemical transformations with three crystallization steps and a single chromatographic purification, which results in an overall yield of 16– 23% depending on the identity of the nucleoside (A, C, G, T).

2.3 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 2-1**) first described by Eschenmoser and colleagues suffered from a number of shortcomings that

Scheme 2.1 Original Strategy Developed to Synthesize α-L-Threose Nucleosides



(Scheme was used with permission from *J.Org. Chem*, **2016**, 81, 2302-2307)

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′) **2.5a** and **2.5b** 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.

2.4 Results and Discussion

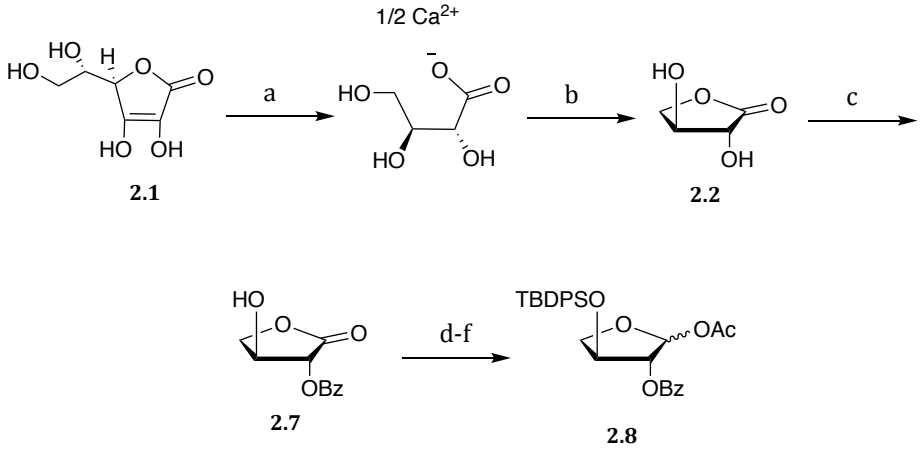
The synthesis of L-threonolactone (**2.2**) from L-ascorbic acid (vitamin C) has been reported previously (**Scheme 2-1**). Accordingly, L-ascorbic acid undergoes oxidative cleavage in the presence of hydrogen peroxide and $CaCO_3$ to produce the calcium salt of L-threonate (**2.6**). Compound **2.6** is then converted to the desired L-threonolactone (**2.2**) using a Dowex cation exchange resin to promote acidification and intramolecular lactonization. However, due to the limited solubility of **2.6**, both steps require to remove the large volumes of water prior to cyclization and crystallization.

We began by examining the crystallization of **2.6** 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 **2.6** to precipitate as a white crystalline material in 85% yield. This subtle change to the

protocol resulted in a yield that favorably comparable against previous reports (65–79%) and allowed large-scale reactions to be performed with substantially higher throughput.

Next, we investigated the conversion of **2.6** to **2.2** (**Scheme 2-2**) 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 (**2.2**), crystallization failed whenever the compound was not properly dried. As an alternative way to avoid using dilute H2SO4, we tested oxalic acid as a calcium exchange reagent in situ. We found that refluxing a heterogeneous mixture of **2.6** in acetonitrile with 1 equiv of oxalic acid and a catalytic amount of para-toluenesulfonic acid produced the desired L-threonolactone (**2.2**) and calcium oxalate as an insoluble side product precipitate that could be easily separated by filtration. Using this approach, pure **2.2** is obtained as a white solid in 93% yield after co-evaporation 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 (**2.2**) without the need for rigorous drying of aqueous solution to an anhydrous state.

Scheme 2-2. 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 **2.7**.

(Scheme was used with permission from *J.Org. Chem*, **2016**, 81, 2302-2307)

A major limitation of the original synthetic route developed by Eschenmoser and colleagues was the low regioselectivity observed when α-L-threofuranosyl nucleosides **2.4** are reacted with DMT-Cl (**Scheme 2-1**) 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 **2.5a** and **2.5b**. In principle, this synthetic strategy would have been ideal if **2.5a** and **2.5b** 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 (**2.5b**). 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 **2.5a** into **2.5b**.[14]

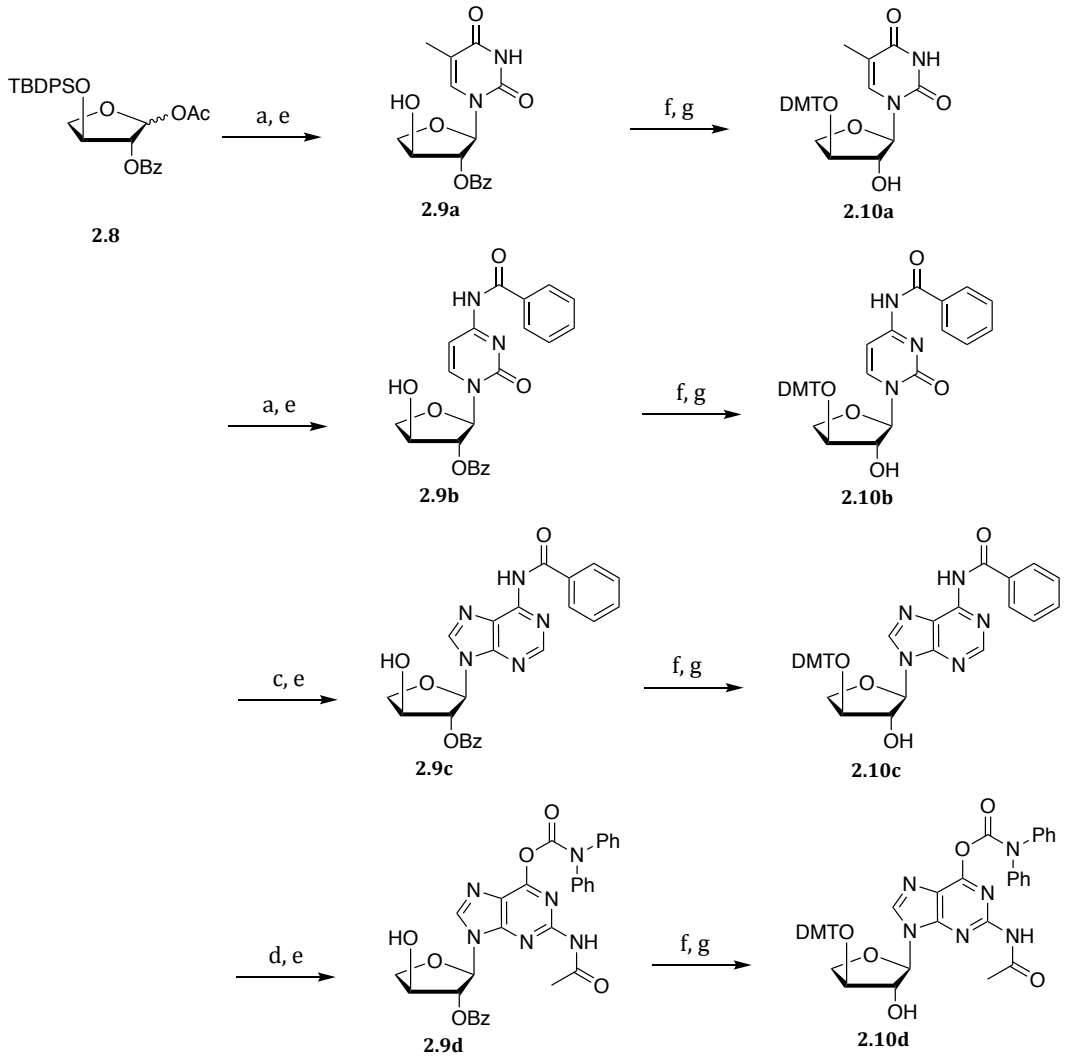
In an effort to generate antiviral compounds based on the structural framework of TNA nucleosides, Herdewijn and colleagues developed a regioselective strategy that defined the 2′ and 3′ hydroxyl positions early in the synthetic pathway. Accordingly, the authors reported that 2 could be selectively converted **2.2** to the 2-O -benzoyl L -threonolactone derivative **1.7** by the addition of benzoyl chloride to a solution of **2.2** and imidazole in acetonitrile after 8 h of stirring from − 5 ° C to room temperature (**Scheme 2-2)**. Although this method proved unsatisfactory in our hands, selective benzoylation of **2.2** was achieved with the addition of 1 equiv. of benzoyl chloride to **2.2** in 1:10 pyridine− dichloromethane solution after stirring for 30 min at 0 ° C. Following the reaction, we further discovered that **2.7** 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,**2.8** compound **2.7** was converted to the universal glycosyl donor **2.8** following three successive chemical transformations. We introduced the tert-butyldiphenylchlorosilyl protection at the 3'-hydroxy position of **2.7**, followed by DIBAL-H reduction of the lactone to the lactol, and finally acetylation at the anomeric position to give **2.8** in 96% yield from **2.7**. Filtration of **2.8** through a short-pad of silica gel proved highly facile and efficient due to the strong polarity differences between **2.8** and the nonwashable reagents. From **2.8**, we constructed a complete set of TNA nucleosides for thymidine (T, **2.9c**), cytidine (C, **2.9b**), adenosine (A, **2.9a**), and guanosine (G, **2.9d**) by applying a Vorbrüggen glycosylation (**Scheme 2-3**) 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, N4-benzoyl cytosine, and N^2-acetyl-O^6-diphenylcarbamoyl guanine to give **2.9a**, **2.9b**, and **2.9d**, respectively, in 64–70% yield after crystallization. However, glycosylation with *N6*-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 *N9*-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 *N9*-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 *N9*-regioisomer as the major synthesized product. Crystallization after 3'-desilylation gave pure *N9*-adenosine nucleoside (**2.9c**) in 43.5% yield.

Compounds **2.9a–d** are key intermediates in the divergent synthesis of TNA nucleoside-3′-triphosphates and 2′-phosphoramidites. For L-threofuranosyl nucleoside 3′-triphosphates, compounds **2.9a–d** can be phosphorylated using the standard Ludwig and Eckstein method, followed by treating the concentrated NH4OH to remove the sugar and nucleobase protecting groups.14 For L-threofuranosyl nucleoside 2′-phosphoramidites, compounds **2.9a–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 **2.10a–d**

Scheme 2-3. 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, CH3CN, 2 h, 60 °C; (c) *N6*-benzoyladenine, BSA, TMSOTf, toluene, 2.5 h, 95 °C; (d) (i) *N*^{2}-acetyl- *O6*-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 was used with permission from *J.Org. Chem*, **2016**, 81, 2302-2307)

(**Scheme 2-3**).[17] While the tritylation of **2.9a–c** proceeded efficiently under standard conditions, tritylation of **2.9d** 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 **2.9a–d** using the above-mentioned methodology. In a typical synthesis run, 125 g of **2.1** is converted into 80 g of **2.7** with an overall yield of 49%. In contrast to previously methods, modifications to the purification protocol of **2.6** and **2.2** makes it possible to generate large quantities of highly pure **2.7** in just 5 days. The conversion of **2.7** to **2.8** is generally performed with 16.5 g of **2.7** to obtain 35 g of **2.8** with an overall yield of 96% in 24 h. The *N*-glycosylation of **2.8**, followed by desilylation, affords nucleosides **2.9a–d** in 64–70% yield, with the noted exception of **2.9c**, which is obtained in slightly lower yield (43.5%). We have found that the synthesis of **2.9a–d** from **2.7** 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.

2.5 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 **2.8** was prepared from **2.7** in three chemical steps, followed by filtration through a short pad of silica gel. The complete set of TNA nucleosides (**2.9a–d**) were synthesized in two reactions from **2.8** and purified by crystallization. Final tritylation, followed by debenzoylation, provided the desired compounds (**2.10a–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.

2.6 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. 1H and 13C NMR spectra were obtained using 400 and 500 MHz NMR spectrometers. Chemical shifts are reported in parts per million (ppm, δ) referenced to the 1H resonance of TMS. 13C spectra are referenced to the residual 13C resonance of the solvent (DMSO-d^6, 39.52 ppm, CD_3OD, 49.00 ppm). Splitting patterns are designated as follows: s,; br,; d,; dd,; t,; q,; m,.

Calcium-L-threonate (2.6).[15]

To a cold (0−5 °C) solution containing 125 g (0.71 mol) of L-ascorbic acid (**2.1**) in 1 L of H_2O was slowly added 125 g (1.3 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 (**2.6**) 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 (2.2).[19]

To a suspension of calcium L-threonate (**2.6**) (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 (**2.2**) as a white solid: yield 66.5 g (93%); silica gel TLC (EtOAc): R_f = 0.36; 1H 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 (2.7).[18]

To a cold (0−5 °C) solution containing 66.5 g (0.56 mol) of L-threonolactone (**2.2**) 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 (**2.7**) was obtained as a white solid: yield 80.1 g (64%); silica gel TLC (1:1 hexanes−EtOAc): R_f = 0.30; 1H 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 (2.8).[18]

To a cold (0−5 °C) solution containing 16.5 g (74.3 mmol) of 2-*O*-benzoyl-L-threonolactone (**2.7**), 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) R_f = 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 (R_f 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 H2O (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 (**2.8**) 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 (2.9a).

To a solution containing 34.0 g (67.4 mmol) of 1-*O*-acetyl-2-*O*-benzoyl-3-*O*-tert-butyldiphenylsilyl-L-threofuranose (**2.8**) and 8.9 g (70.5 mmol) of thymine in 140 mL of anhydrous acetonitrile was added 45 mL (145 mmol) of BSA, 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): R_f = 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 TBAF (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 (**2.9a**) was obtained as a white solid: yield 14.44 g (64%); silica gel TLC (1:2 hexanes–EtOAc): R_f = 0.35; 1H 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); 13C 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 (2.9b).**

To a solution containing 33.5 g (66.3 mmol) of 1-*O*-acetyl-2-*O*-benzoyl-3-*O*-tert-butyldiphenylsilyl-L-threofuranose (**2.8**) 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 BSA, 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): R_f = 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 TBAF (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. *N4*-Benzoyl-1-(2′-*O*-benzoyl-α-L-threofuranosyl)cytosine (**2.9b**) was obtained as a white solid: yield

19.4 g (70%); silica gel TLC (1:2 hexanes–EtOAc): R_f = 0.40; 1H 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); 13C 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 (2.9c).[18]

A mixture of 40.5 g (80.0 mmol) of 1-*O*-acetyl-2-*O*-benzoyl-3-*O*-tert-butyldiphenylsilyl-L-threofuranose (**2.8**) and 19.1 g (80.0 mmol) of N^6-benzoyladenine was co-evaporated 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 BSA, 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 TBAF (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 (**2.9c**) was obtained as a white solid: yield 15.5 g (43.5%); silica gel TLC (EtOAc): Rf = 0.35; 1H 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 2.9d).

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 (**2.8**) (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 (**2.8**). 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): R_f = 0.7.

To a cold (0−5 °C) solution containing 16.12 g of crude *N*^{2}-acetyl-*O6*-diphenylcarbamoyl-9-(2′-*O*-benzoyl-3′-*O*-tert-butyldiphenylsilyl-α-L-threofuranosyl) guanine in THF (50 mL) was added dropwise TBAF (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 (**2.9d**) was obtained as a white solid: yield 8.9 g (69%); silica gel TLC (1:2 hexanes−EtOAc): Rf = 0.25; 1H NMR (400 MHz, DMSO-*d6*) 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); 13C 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 (2.10a).[15]

A mixture containing 3.2 g (9.6 mmol) of 1-(2′-*O*-benzoyl-α-L-threofuranosyl)thymine (**2.9a**), 4.88 g (14.4 mmol, 1.5 equiv) of DMT-Cl, and 100 mg of DMAP (0.82 mmol) was co-evaporated 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 (**2.10a**) 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 (2.10b).[15]

A mixture of 3.18 g (7.7 mmol) of N^4-benzoyl-1-(2′-*O*-benzoyl-α-L-threofuranosyl)cytosine (**2.9b**), 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-threofuranosyl)cytosine (**2.10b**) was obtained as a white foam: yield 3.12 g (65%); 1H 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 (2.10c).[15]

A mixture of 3.39 g (7.6 mmol) of *N*^{6}-benzoyl-9-(2′-*O*-benzoyl-α-L-threofuranosyl)adenine (**2.9c**), 3.86 g (11.4 mmol, 1.5 equiv) of DMT-Cl, and 93 mg (0.76 mmol, 0.1 equiv) of DMAP was co-evaporated 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%); 1H 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 (2.10d).[15]

5.2 g (8.74 mmol) of *N*^{2}-acetyl-*O*^{6}-diphenylcarbamoyl-9-(2′-*O*-benzoyl-α-L-threofuranosyl)guanine (**2.9d**) 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–Et3N 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–Et3N and (65:34:1) hexane–EtOAc–Et3N. *N*^{2}-Acetyl-*O*^{6}-diphenylcarbamoyl-9-(3′-*O*-dimethoxytrityl-α-L-threofuranosyl)guanine(**2.10d**) was obtained as a pale yellow powder: yield 1.9 g (71%); 1H 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.; Chattopadhaya, 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 3

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*

Synthes is of α-L-Threofuranosyl Nucleoside-3′-Monophosphates, 3′-Phosphoro(2-Methyl)imidazolides, and 3′-Triphosphates. *J. Org. Chem*, 2017, 82, 5910−5916. († equal contribution)

3.1 Contribution Statement

Bala, S. and Chaput, J. C. 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.

3.2 Abstract of the Chapter

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.

3.3 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 3-1**).

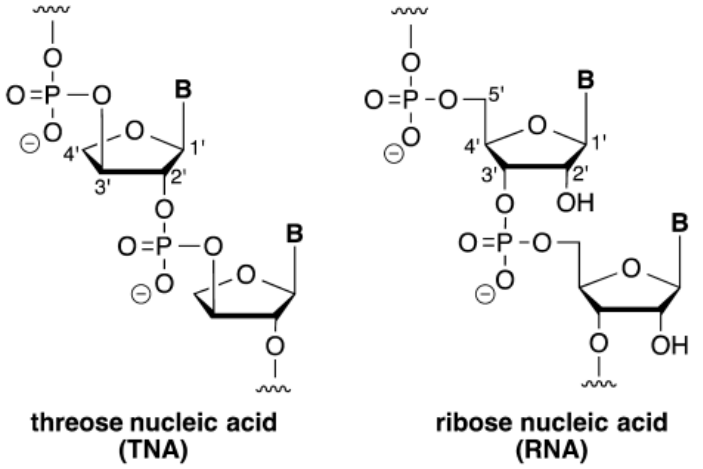


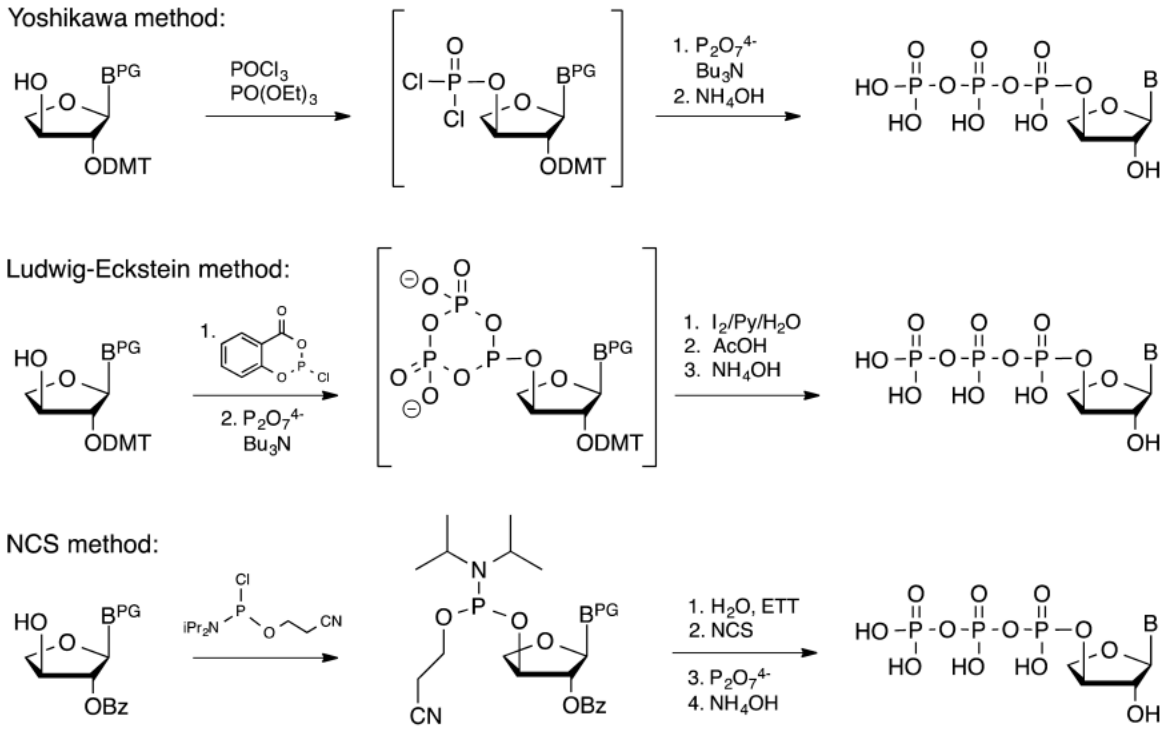
Figure 3-1. Molecular structure of TNA and RNA.

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 3-1**) 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 [27] 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 3-1. Previous Strategies for Synthesizing TNA Triphosphates



(Scheme was used with permission from . *J. Org. Chem*, 2017, 82, 5910−5916)

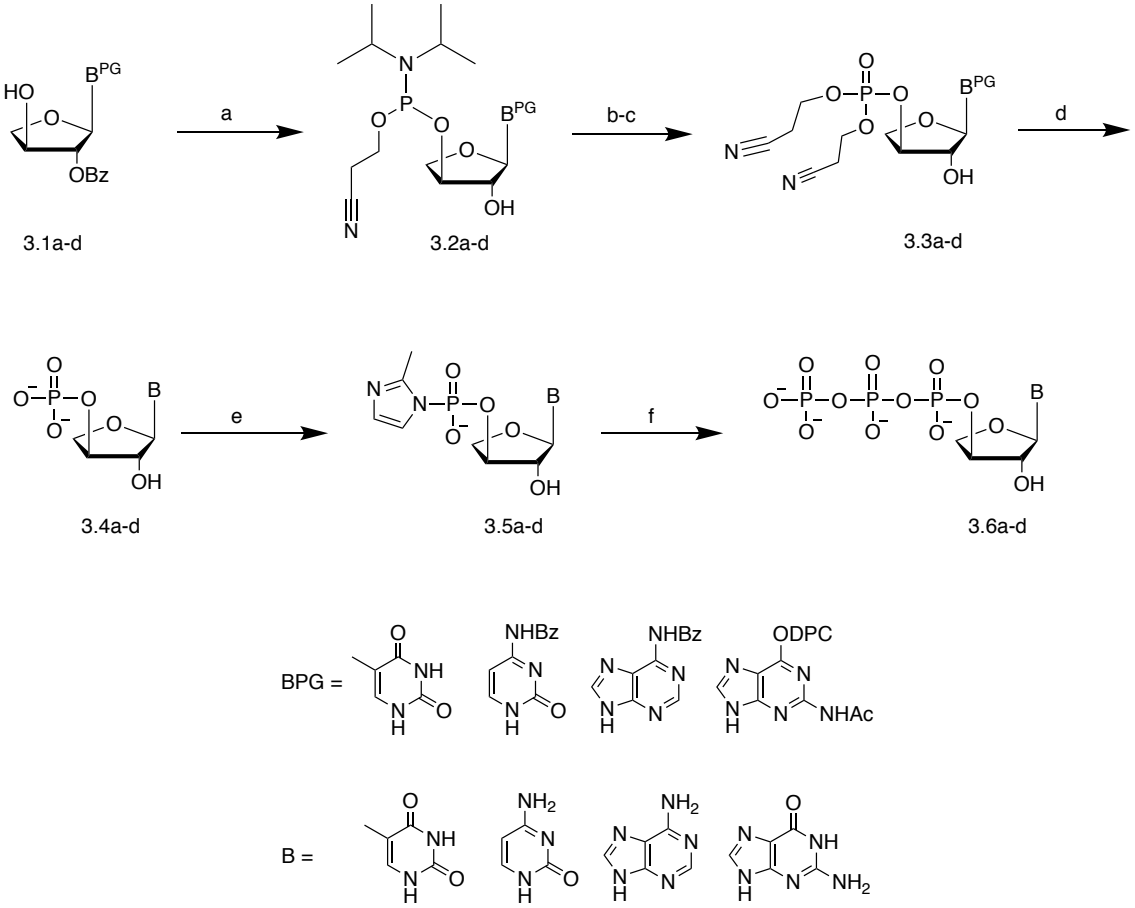
To overcome the shortcomings of previous methods, we developed a robust strategy for generating TNA nucleoside monophosphates, monophosphate derivatives, and nucleoside triphosphates that is straightforward to perform and could be transferrable to other modified nucleosides. Our approach consists of four synthetic steps to generate TNA [28] 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.

3.4 Results and Discussion

We designed a synthetic approach (**Scheme 3-2**) for constructing 3′-phosphorylated TNA compounds from 2′-*O*-benzoyl TNA nucleosides **3.1a–d**. TNA nucleosides **3.1a–d** were obtained from L-ascorbic acid using known methodology (**Scheme 3-2 and 3-3**) 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 (**3.4a–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 **3.1a–d** were converted to their corresponding 3′-*O*-phosphoramidite derivatives **3.2a–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 **3.2a–d** in high yield. Intermediates **3.2a–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 **3.3a–d**, which were purified by silica gel

Scheme 3-2. Synthesis of TNA Triphosphates Using Activated TNA Monophosphates



Reagents and conditions: (a) 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite, 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).

(Scheme was used with permission from . *J. Org. Chem*, 2017, 82, 5910−5916)

chromatography. The resulting TNA nucleoside 3′-monophosphates **3.4a–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 30 found that the higher reactivity of the chlorophosphoramidite reagent by P(III) chemistry, coupled with the ability to purify the nucleoside monophosphate precursor **3.3a-d** by silica gel chromatography, provided an efficient method for obtaining large quantities (grams) of highly pure TNA 3′-monophosphates.

Next, TNA monophosphates **3.4a–d** were converted to their 3′-phosphoro(2- methyl)-imidazolide derivatives **3.5a–d** using classic methodology previously developed by Mukaiyama and co-workers.[26] Accordingly, a Mitsunobu-like reaction is performed by incubating **3.4a–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 **3.5a–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 3-2**)

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

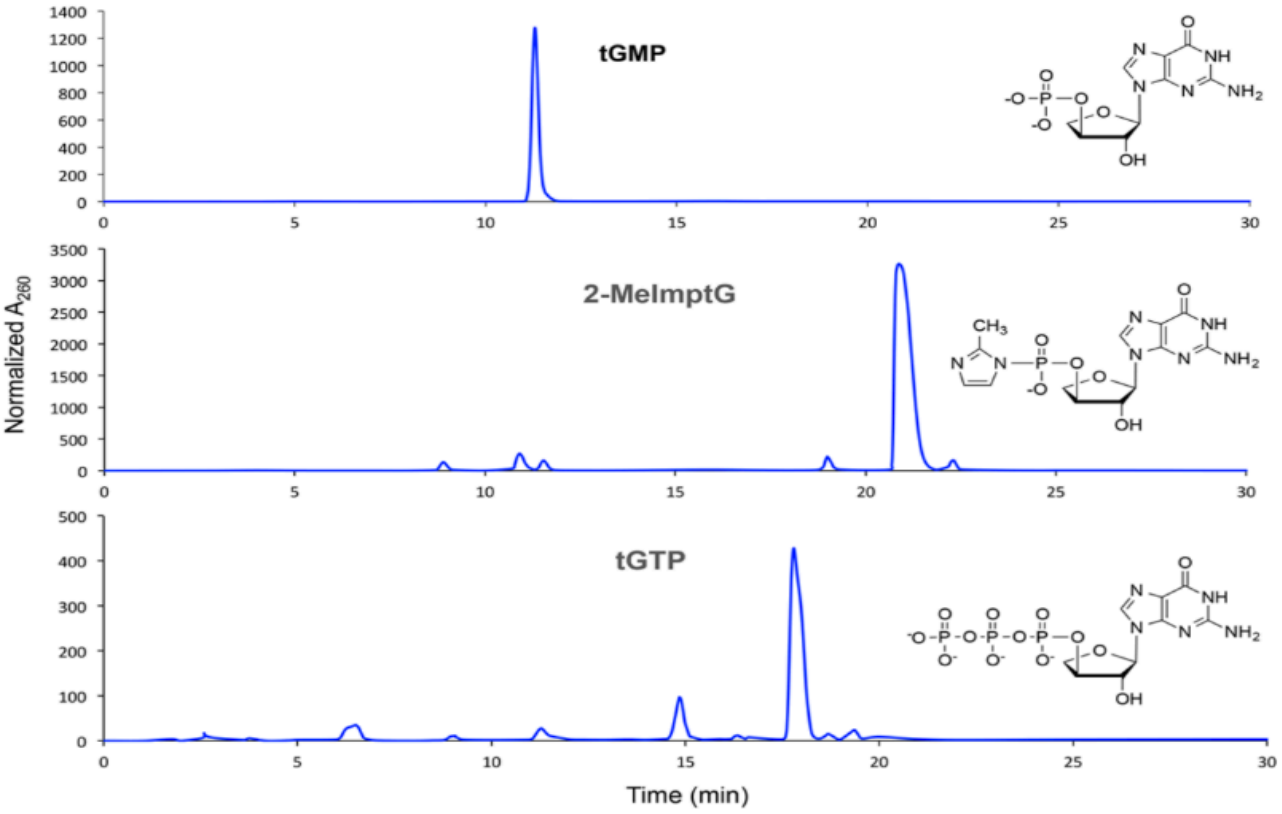


Figure 3-2. HPLC analysis of the crude reaction for tGMP, 2-MeImptG, and tGTP.

(Figure was used with permission from . *J. Org. Chem*, 2017, 82, 5910−5916)

TNA nucleoside 3′-triphosphates **3.6a−d** were synthesized from the activated nucleoside monophosphates **3.5a−d** by displacing the activated 2-methylimidazole leaving group with pyrophosphate. In these reactions, TNA nucleosides **3.5a−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 3-2**). In

all cases, >90% of the starting nucleoside was consumed within 12 h. The desired TNA nucleoside triphosphates **3.6a–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 (**3.6a–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 32 aliquots of TNA solutions that were stored at temperatures of 4, 24, and 37 °C. A time course analysis performed over 8 days (**Table 3-1**) 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 3-1. Thermal Stability of Purified TNA Triphosphates[a]

(Table was used with permission from . *J. Org. Chem*, 2017, 82, 5910−5916.)

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 3-3**) 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.

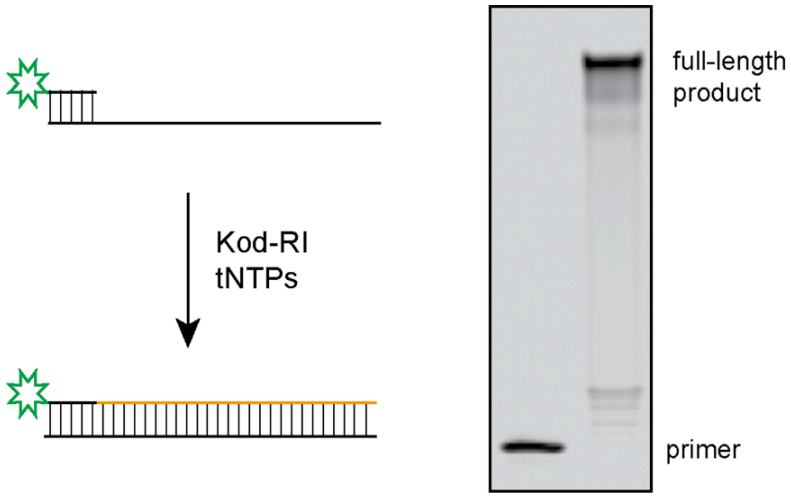


Figure 3-3. Primer-extension analysis

(Figure was used with permission from . *J. Org. Chem*, 2017, 82, 5910−5916.)

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.

3.5 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.

3.6 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 Nano Drop 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 35 (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 (3.3a).

To a stirring solution containing 1 g (3.01 mmol) of 1-(2′-O-benzoyl-α-L-threofuranosyl) thymine (**3.1a**)[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 (**3.2a**)[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 (**3.2a**)[23] in 10 mL of anhydrous acetonitrile was added 0.62 mL (9.0 mmol) of 3-hydroxypropionitrile followed by the addition of 26.8 mL of tetrazole (0.45 M solution in anhydrous acetonitrile, 12.0 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 (**3.3a**) as a white foam: yield 0.87 g (56%); silica gel TLC (DCM/MeOH, 9:1) R_f = 0.35; ^{1}H NMR (500 MHz, $CDCl_3$) 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); 13C 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 (3.3b).

To a stirring solution containing 1 g (2.3 mmol) of N^4-benzoyl-1-(2′-*O*-benzoyl-α-L-threofuranosyl)cytosine (**3.1b**)[24] and 290 mg (2.3 mmol) of DMAP in 12 mL anhydrous CH_2Cl_2 was added 1 mL (5.9 mmol) of *N,N*′-diisopropylethylamine followed by the addition of 1.05 mL (4.7 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 (**3.2b**)[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 re-dissolved 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 (3.3b) as a white foam: yield 0.68 g (47.2%); silica gel TLC (DCM/acetone, 10:1) R_f = 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-d6) 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); 31P 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 (3.3c).

To a stirring solution containing 1 g (2.2 mmol) of N^6-benzoyl-9-(2′-*O*-benzoyl-α-L-threofuranosyl) adenine (**3.1c**)[24] and 273 mg (2.2 mmol) of DMAP in 12 mL anhydrous CH_2Cl_2 was added 0.78 mL (4.4 mmol) of *N,N′*-diisopropylethylamine followed by the addition of 0.85 mL (3.8 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 (**3.2c**)[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 (6.7 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 re-dissolved 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 (**3.3c**) as a white foam: yield 0.65 g (46%); silica gel TLC (DCM/MeOH, 9:1) R_f = 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(3.3d).

To a stirring solution containing 1 g (1.7 mmol) of N^2-acetyl-*O6*-diphenylcarbamoyl-9-(2′-*O*-benzoyl-α-L-threofuranosyl) guanine (**3.1d**)[24] and 210 mg (1.68 mmol) of DMAP in 12 mL anhydrous CH_2Cl_2 was added 0.58 mL (3.4 mmol) of *N,N*′-diisopropylethylamine followed by the addition of 0.821 mL (3.7 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 (**3.2d**)[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 (**3.2d**)[23] in 10 mL of anhydrous acetonitrile was added 0.36 mL (5.0 mmol) of 3-hydroxypropionitrile followed by the addition of 14.9 mL of tetrazole (0.45 M solution in anhydrous acetonitrile, 6.7 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 (**3.3d**) as a white foam: yield 0.59 g (45%); silica gel TLC (DCM/acetone, 10:1) Rf = 0.35; 1H 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); 13C 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 (3.4a−d).

In a sealed tube, 1-(2′-*O*-benzoyl-3′-*O*-[bis(2-cyanoethyoxy)-phosphoryl]-α-L-threofuranosyl) nucleosides (0.32 mmol) (**3.3a−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 (**3.4a−d**) were obtained as the ammonium salt (white solid) in near quantitative yield.

α -L-Threofuranosyl thymine-3′-monophosphate (3.4a).

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 (3.4b).

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 (3.4c).

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; 31P 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 (3.4d).

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 (3.5a–d).

To a solution containing α-L-threofuranosyl thymidine-3′-monophosphate (**3.4a–d**) (0.27 mmol) and 2-methylimidazole (2.7 mmol) in 5 mL of anhydrous DMSO were added triethylamine (2.7 mmol), triphenylphosphine (1.1 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 (**3.5a–d**) as the sodium salt.

α -L-Threofuranosyl thymine-3′-phosphor-2-methylimidazolide (3.5a).

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 (3.5b).

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 (3.5c).

Product yield: 100.5 mg (97.7%, ε_{259} = 15200); 31P 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 (3.5d).

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 (3.6a–d).[21–23]

To a solution containing 0.1 mmol of α-L-threofuranosyl nucleosides 3′-phosphor-2-methylimidazolides (**3.5a–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 (**3.6a–d**) solution.

α -L-Threofuranosyl thymine-3′-triphosphate (3.6a).

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

α -L-Threofuranosyl cytosine-3′-triphosphate (3.6b).

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

α -L-Threofuranosyl adenine-3′-triphosphate (3.6c).

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

α -L-Threofuranosyl guanine-3′-triphosphate (3.6d).

Product yield after HPLC purification: 44.2 mg (89.7%, ε_{253} = 13700).

Thermal Stability Analysis of TNA Triphosphates.

A small volume (30 μL) of 4 mM of 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 2-1.

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 1Å~ 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 2.6 a–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. *Natur*, **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**; Vol. 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 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

Bala, S. and Liao, J.-Y. designed the synthetic strategy. Bala, S. and Liao, J.-Y. 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 of the Chapter

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]

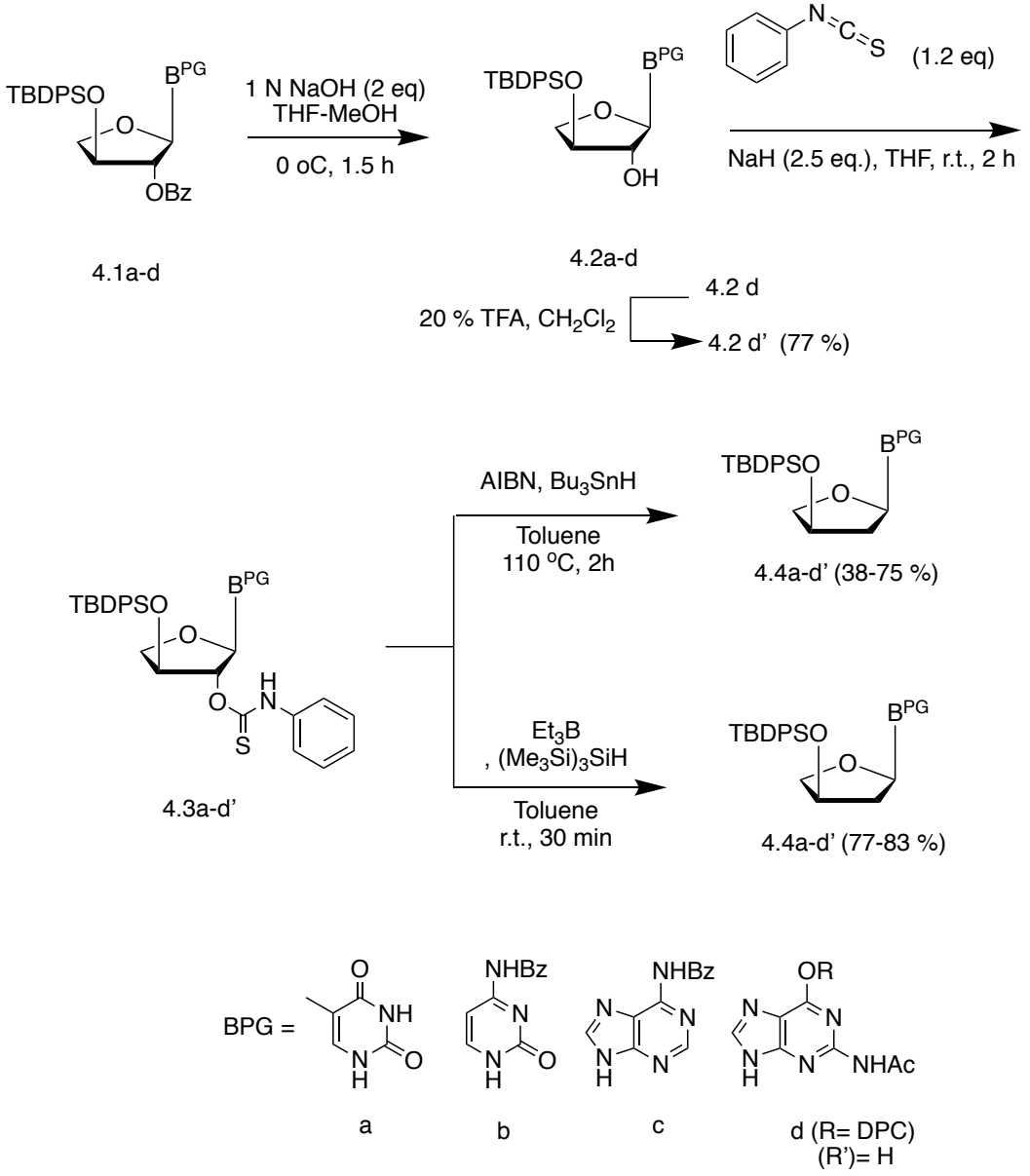
An alternative approach is to use chain-terminating nucleoside triphosphates that undergo a cycle of catalysis prior to crystallization. This strategy has been used successfully to capture the closed ternary structures of DNA polymerases by trapping the second incoming DNA nucleoside triphosphate in 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′



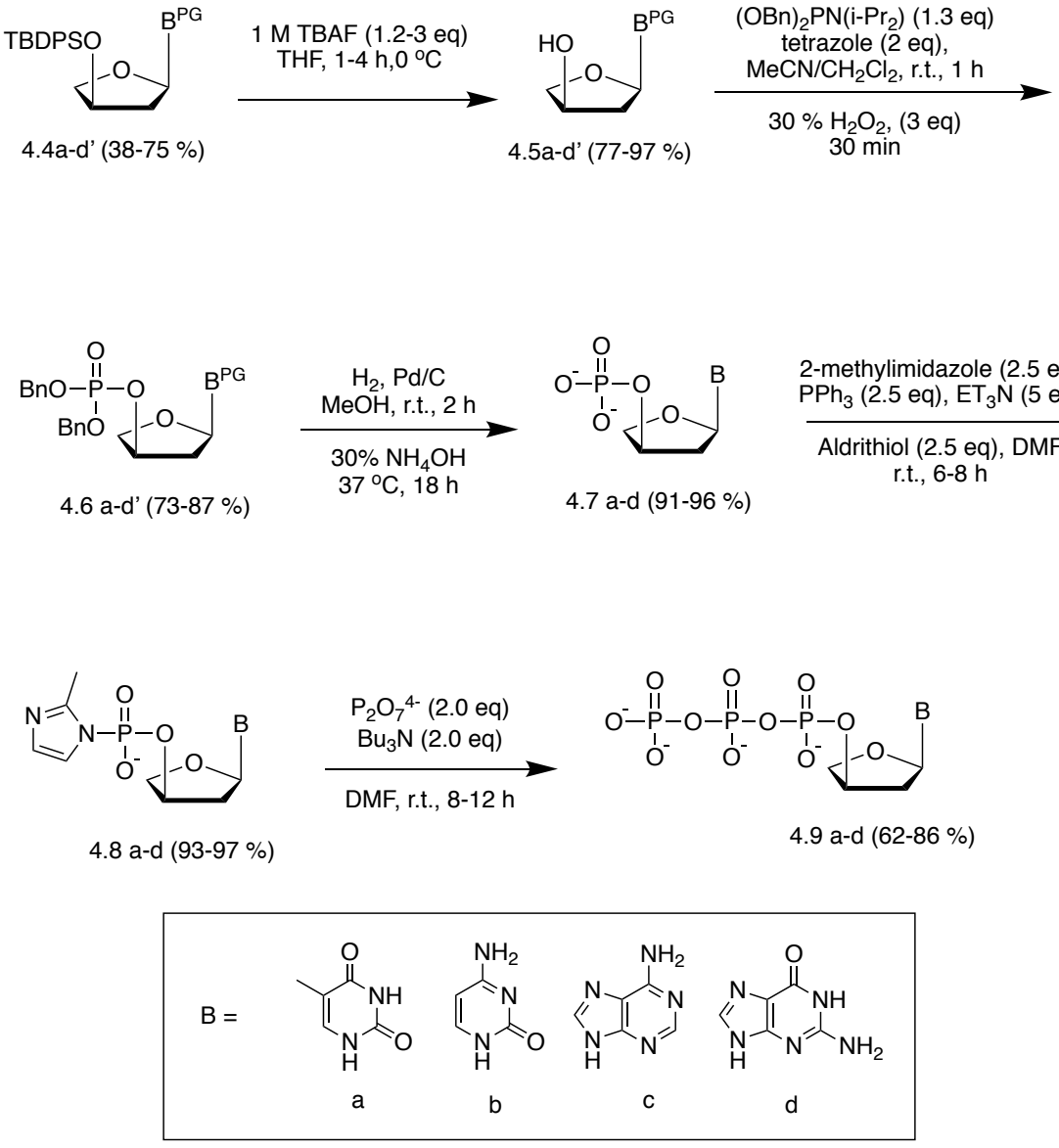
(Scheme was used with permission from . *J. Org. Chem*. **2018**, *83*, 8840−8850)

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 3-2**). 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)



(Scheme was used with permission from . *J. Org. Chem*. **2018**, *83*, 8840−8850)

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′**.

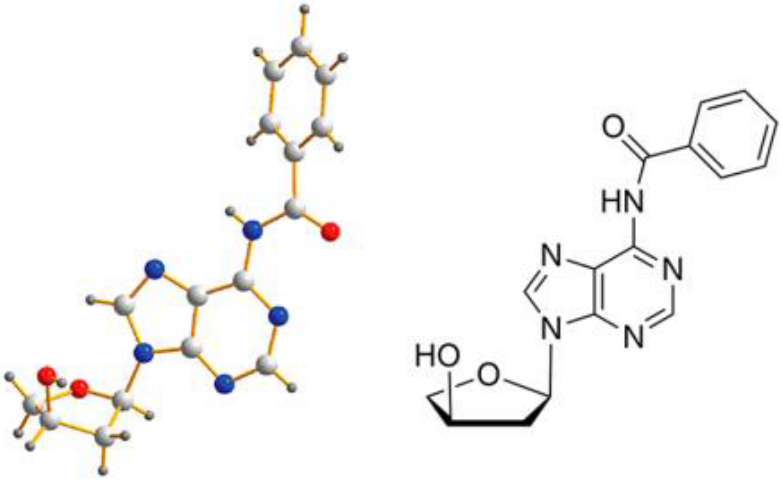


Figure 4.1 X-ray crystal structure of 6-benzoyl-2′-deoxy-α-L-threofuranosyl adenosine nucleoside 4.5c.

(Figure was used with permission from . *J. Org. Chem*. **2018**, *83*, 8840−8850)

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. 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.

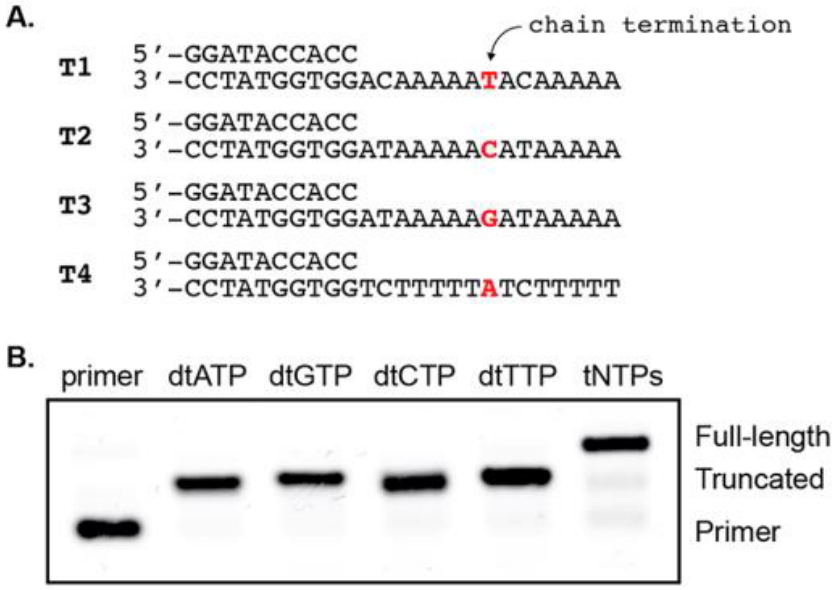


Figure 4-2. Sequence and primer extension results.

(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

(Figure was used with permission from . *J. Org. Chem*. **2018**, *83*, 8840–8850.)

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.

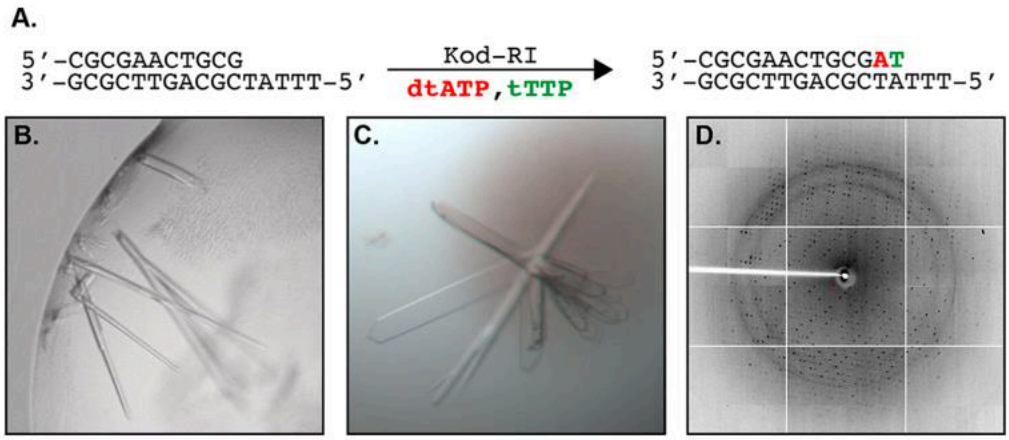


Figure 4-3. X-ray crystallography of Kod-RI bound to 4.9c.

(A) 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.

(Figure was used with permission from . *J. Org. Chem*. **2018**, *83*, 8840–8850.)

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 (40−63 μm 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 31P 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 mm2 column with 5 μm particle size, and nucleoside triphosphates were purified on a semipreparative reverse phase C18 250 × 9.4 mm2 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, ε267: 9600 $M^{-1}cm^{-1}$. C, $_{\varepsilon 271}$: 91 00 $M^{-1}cm^{-1}$. G, $_{\varepsilon 259}$: 15200 $M^{-1}cm^{-1}$. A, $_{\varepsilon 259}$: 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*-*tert*-butyldiphenylsilyl-α-L-threofuranosyl)thymine (**4.1a**) as a white solid: silica gel TLC (hexane/EtOAc, 1:1) R_f = 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) R_f = 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) R_f = 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-3′-*O*-*tert*-butyldiphenylsilyl-α-L-threofuranosyl)guanine (**4.1d**) as a yellow foam: silica gel TLC (hexane/EtOAc, 1:1) R_f = 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 4.73 g (6.5 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 (6.5 mmol) of ice-cold 1 N NaOH(aq) dropwise while stirring. After 15 min, additional 3.27 mL (3.3 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) R_f = 0.65; ^1H 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, CDCl3) *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 (10.3 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 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 1.6 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) R_f = 0.45; 1H 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 (5.8 mmol) of N^6-benzoyl-9-(3′-*O*-*tert*-butyldiphenylsilyl-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 (4.2c) as a white solid: yield 2.6 g (76.7%); silica gel TLC (EtOAc/CH_2Cl_2, 2:1) R_f = 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 (7.2 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) R_f = 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 (4.9 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 ($MeOH/CH_2Cl_2$) to afford N^2-acetyl-9- (3′-O-*tert*-butyldiphenylsilyl-α-L-threofuranosyl)guanine (4.2d′) as a pale-yellow solid: yield 2.4 g (77%); TLC ($MeOH/CH_2Cl_2$, 1:10) R_f = 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 (8.0 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) R_f = 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 (4.7 mmol) of *N*^{4}-benzoyl-1-(3′-*O*-*tert*-butyldiphenylsilyl-α-L-threofuranosyl)cytosine (**4.2b**) and 319 mg (12.6 mmol) of NaH in 47 mL of anhydrous THF was added dropwise 813 μL (6.6 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/CH_2Cl_2, 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) R_f = 0.4; ^{1}H 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/CH_2Cl_2, 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 (2.5 mmol) of **4.2d′** and 0.27 g (6.7 mmol) of NaH in 20 mL of anhydrous THF was added dropwise 0.37 mL (2.9 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) R_f = 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 (2.1 mmol) of **4.3a** in 15 mL of anhydrous toluene was added 2.08 mL (5.4 mmol) of tris(trimethylsilyl)silane followed by 2.6 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) R_f = 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, CDCl3) 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 (**3.3b**) was coevaporated with 10 mL of anhydrous toluene twice under reduced pressure. To a solution containing 2.0 g (2.9 mmol) of **4.3b** in 58 mL of

anhydrous toluene was added 2.18 mL (7.3 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) R_f = 0.28; ^1H 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, CDCl3) 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 (2.5 mmol) of **4.3c** in 25 mL of anhydrous toluene was added 1.95 mL (6.3 mmol) of trimethylsilylsilane followed by adding dropwise 3.02 mL (3.0 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 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) R_f = 0.3; ^1H 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 co-evaporated twice with 3 mL of anhydrous toluene under reduced pressure and dried overnight under vacuum. To a solution of 0.82 g (1.3 mmol) of 4.3d′ in 15 mL of anhydrous toluene was added 1.3 mL (3.1 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) R_f = 0.55; ^1H 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 (1.6 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) R_f = 0.35; ^1H 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); 13C 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 (2.2 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 6.7 mL (6.7 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 **4.5b** was obtained as a white solid: yield 650 mg (97.0%); silica gel TLC ($MeOH/CH_2Cl_2$, 1:20) R_f = 0.31; ^1H 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 (2.3 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) R_f = 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 (1.0 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 **4.5d'** 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) $C_{11}H_{13}N_5O_4Na$ $[M + Na]^+$ 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 (4.7 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) R_f = 0.55; ^1H 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 (1.0 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) R_f = 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 (1.0 mmol) of tetrazole in acetonitrile, 228 μL (0.7 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) R_f = 0.35; ^{1}H 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 (2.2 mmol) of **4.5d'**, 10 mL (4.5 mmol) of tetrazole, 0.87 mL (2.6 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, CDCl3) 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 **3.7a–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 (0.8 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 δ : −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); 13C NMR (125.8 MHz, DMSO-d_6) δ 164.7, 154.4, 142.0, 93.7, 85.6, 75.3, 72.9, 40.4; 31P 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 (2.2 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 (1.4 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) R_f = 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: ^{1}H 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 (1.8 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 (0.4 mmol) of **4.7b**, 2 mL of anhydrous DMSO, 250 μL (1.8 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 (0.8 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 (0.8 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); ^{13}C 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).

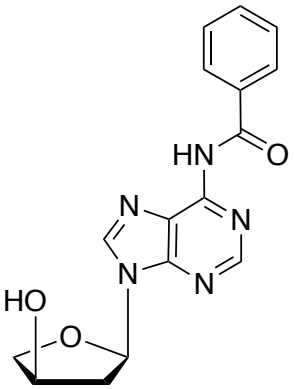
Product yield after HPLC purification was: 9.3 mg (62%, ε_{259}: 15400); ^{1}H NMR (400 MHz, D_2O) δ 8.17 (s, 1H), 6.27 (d, 1H, *J* = 7.2 Hz), 5.26 (m, 1H), 4.42 (d, 1H, *J* = 10.4 Hz), 4.20 (dd, 1H, *J* = 10.4, 2.8 Hz), 2.94 (m, 1H), 2.73 (m, 1H); ^{13}C NMR (150 MHz, D_2O) δ 161.4, 156.3, 151.4, 138.0, 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 MgSO4, 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.



3.5C

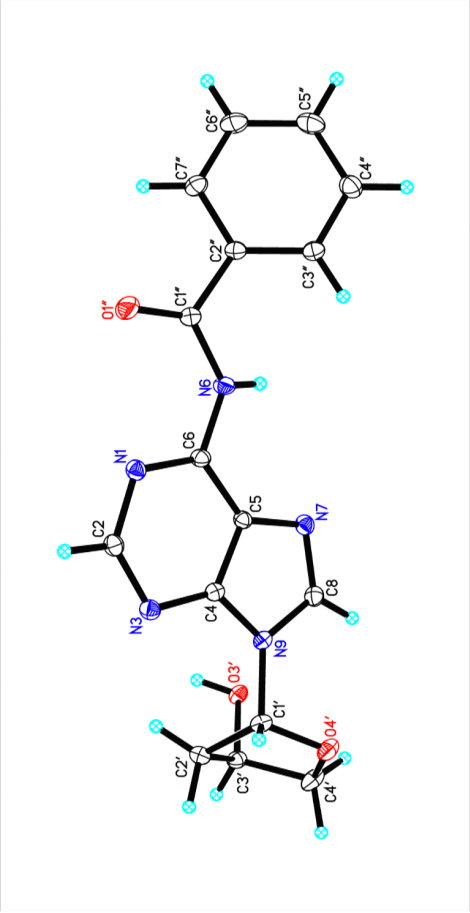


Figure 4-4. ORTEP representation of 6-benzoyl-2′-deoxy-α-L-threofuranosyl adenosine nucleoside 4.5c.

(Figure was used with permission from . *J. Org. Chem*. **2018**, *83*, 8840−8850.)

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 4.1. Crystal data and structure refinement for jcc5.

Identification code	jcc5	
Empirical formula	C16 H15 N5 O3	
Formula weight	325.33	
Temperature	88(2) K	
Wavelength	0.71073 Å	
Crystal system	Monoclinic	
Space group	*P*2**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 F2
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 4.2. Atomic coordinates (x 104) and equivalent isotropic displacement parameters (Å2x 103) for jcc5. U(eq) is defined as one third of the trace of the orthogonalized Uij 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 4.3. 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 4.4. Anisotropic displacement parameters (Å2x 103) for jcc5. The anisotropic displacement factor exponent takes the form: -22[h2 a*2U11 + ... + 2 h k a* b* U12]

	U11	U22	U33	U23	U13	U12
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 4.5. Hydrogen coordinates (x 104) and isotropic displacement parameters (Å2x 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 4.6. 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 4.7. 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 3.5c.

References

1. Schöning, K. U.; Scholz, P.; Guntha, S.; Wu, X.; Krishnamurthy, R.; Eschenmoser, A. Chemical etiology of nucleic acid structure: the α -threofuranosyl-(3′–>2') oligonucleotide system. *Science*, **2000**, *290*, 1347–1351.
2. Eschenmoser, A. Chemical etiology of nucleic acid structure. *Science,* **1999**, *284*, 2118–2124.
3. Schöning, K.-U.; Scholz, P.; X, W.; Guntha, S.; Delgado, G.; Krishnamurthy, R.; Eschenmoser, A. The α -L-Threofurnaosyl-(3′-2')- oligonucleotide system ('TNA'): synthesis and pairing properties. *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. The structural diversity of artificial genetic polymers. *Nucleic Acids Res.* **2016**, *44*, 1007–1021.
5. Larsen, A. C.; Dunn, M. R.; Hatch, A.; Sau, S. P.; Youngbull, C.; Chaput, J. C. A general strategy for expanding polymerase function by droplet microfluidics. *Nat. Commun*. **2016**, *7*, 11235.

6. Dunn, M. R.; Otto, C.; Fenton, K. E.; Chaput, J. C. Improving Polymerase Activity with Unnatural Substrates by Sampling Mutations in Homologous Protein Architectures. *ACS Chem. Biol.* **2016**, *11*, 1210−1219.
7. Yu, H.; Zhang, S.; Chaput, J. C. Darwinian evolution of an alternative genetic system provides support for TNA as an RNA progenitor. *Nat. Chem.* **2012**, *4*, 183−187.
8. Dunn, M. R.; Chaput, J. C. Reverse transcription of threose nucleic acid by a naturally occurring DNA polymerase. *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. Synthesis and Evolution of a Threose Nucleic Acid Aptamer Bearing 7-Deaza-7-Substituted Guanosine Residues. *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. Evaluating TNA stability under simulated physiological conditions. *Bioorg. Med. Chem. Lett*. **2016**, *26*, 2418− 2421.
11. Dunn, M. R.; Jimenez, R. M.; Chaput, J. C. Analysis of aptamer discovery and technology. *Nat. Rev. Chem*. **2017**, *1*, 0076.
12. Zhou, J.; Rossi, J. Aptamers as targeted therapeutics: current potential and challenges. *Nat. Rev. Drug Discovery* **2017**, *16*, 181−202.
13. Nikoomanzar, A.; Dunn, M. R.; Chaput, J. C. Evaluating the rate and substrate specificity of laboratory evolved XNA polymerases. *Anal. Chem*. **2017**, *89*, 12622−12625.
14. Chim, N.; Shi, C.; Sau, S. P.; Nikoomanzar, A.; Chaput, J. C. Structural analysis of TNA synthesis by an Engineered TNA polymerase. *Nat. Commun*. **2017**, *8*, 1810.

15. Kielkowski, P.; Fanfrlik, J.; Hocek, M. 7-Aryl-7-deazaadenine 2'- deoxyribonucleoside triphosphates (dNTPs): better substrates for DNA polymerases than dATP in competitive incorporations. *Angew. Chem., Int. Ed*. **2014**, *53*, 7552–7555.
16. Chen, T.; Romesberg, F. E. Directed polymerase evolution. *FEBS Lett*. **2014**, *588*, 219–229.
17. Loakes, D.; Holliger, P. Polymerase engineering: towards the encoded synthesis of unnatural polymers. *Chem. Commun*. **2009**, 4619–4631.
18. Kool, E. T. Active site tightness and substrate fit in DNA replication. *Annu. Rev. Biochem*. **2002**, *71*, 191–219.
19. Houlihan, G.; Arangundy-Franklin, S.; Holliger, P. Exploring the Chemistry of Genetic Information Storage and Propagation through Polymerase Engineering. *Acc. Chem. Res*. **2017**, *50*, 1079– 1087.
20. Aschenbrenner, J.; Marx, A. DNA polymerases and biotechnological applications. *Curr. Opin. Biotechnol.* **2017**, *48*, 187–195.
21. Doublie, S.; Tabor, S.; Long, A. M.; Richardson, C. C.; Ellenberger, T. Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 A resolution. *Nature* **1998**, *391*, 251–258.
22. Kropp, H. M.; Betz, K.; Wirth, J.; Diederichs, K.; Marx, A. Crystal structures of ternary complexes of archaeal B-family DNA polymerases. *PLoS One* **2017**, *12*, e0188005.
23. 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*, 2302–2307.
24. Dumbre, S. G.; Jang, M.-Y.; Herdewijn, P. Synthesis of α -L-threose nucleoside phosphonates via regioselective sugar protection. *J. Org. Chem*. **2013**, *78*, 7137–7144.

25. Zou, K.; Horhota, A.; Yu, B.; Szostak, J. W.; McLaughlin, L. W. Synthesis of a-L-threofuranosyl nucleoside triphosphates (tNTPs). *Org. Lett*. **2005**, *7*, 1485−1487.
26. Barton, D. H. R.; McCombie, S. W. A new method for the deoxygenation of secondary alcohols. *J. Chem. Soc., Perkin Trans*. 1 **1975**, 1574−1585.
27. Robins, M. J.; Wilson, J. S. Nucleic-Acid Related-Compounds 0.32. Smooth and Efficient Deoxygenation of Secondary Alcohols – a General Procedure for the Conversion of Ribonucleosides to 2'- Deoxynucleosides*. 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. Synthesis and Biological Evaluation of 2'-Deoxy-Beta-Threo-Pentofuranosyl Nucleosides - Reversion to Starting Alcohol in Barton-Type Reductions of Thionocarbonates. *Can. J. Chem*. **1988***, 66*, 1258−1262.
29. Barton, D. H. R.; Jang, D. O.; Jaszberenyi, J. C. An improved radical chain procedure for the deoxygenation of seconary and primary alcohols using diphenylsilane as hydrogen atom donor and triethylborane-air as initiator. *Tetrahedron Lett*. **1990**, *31*, 4681−4684.
30. Oba, M.; Nishiyama, K. Radical-based deoxygenation of aliphatic alcohols via thioxocarbamate derivatives. *Tetrahedron* **1994**, *50*, 10193−10200.
31. Sau, S. P.; Chaput, J. C. A Gram-Scale HPLC-Free Synthesis of TNA Triphosphates Using an Iterative Phosphorylation Strategy. *Org. Lett*. **2017**, *19*, 4379−4382.
32. Burgess, K.; Cook, D. Syntheses of Nucleoside Triphosphates. *Chem. Rev*. **2000**, *100*, 2047−2060.
33. Hollenstein, M. Nucleoside triphosphates−building blocks for the modification of nucleic acids. *Molecules* **2012**, *17,* 13569−13591.

34. Kore, A. R.; Srinivasan, B. Recent Advances in the Synthesis of Nucleoside Triphosphates. *Curr. Org. Synth*. **2013**, *10*, 903–934.

35. Bala, S.; Liao, J. Y.; Mei, H.; Chaput, J. C. Synthesis of α-LThreofuranosyl Nucleoside 3′-Monophosphates, 3′-Phosphoro(2- Methyl)imidazolides, and 3′-Triphosphates. *J. Org. Chem.* **2017**, *82,* 5910–5916.

Chapter 5

P(V) Reagents for the Scalable Synthesis of Natural and Modified Nucleoside Triphosphates

Publication Note

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

Jen-Yu Liao†, Saikat Bala †, Arlene K. Ngor, Eric J. Yik and John C. Chaput*

P(V) Reagents for the Scalable Synthesis of Natural and Modified Nucleoside Triphosphates

***Journal of the American Chemical Society*,** 2019. *141*,34,13286-13289 († Equal Contribution)

5.1 Contribution Statement

Chaput, J. C., Bala, S. and Liao. J. Y. designed the synthetic strategy. All authors contributed to the synthesis, optimization, and characterization of the synthetic compounds. Bala, S. and Liao, J.-Y. involved in the experimental section writing and supporting information compiling. Liao, J.-Y., Bala, S. and Chaput, J. C. wrote the manuscript.

5.2 Abstract of the Chapter

Natural and chemically modified nucleoside triphosphates impact nearly every major aspect of healthcare research from DNA sequencing to drug discovery. However, a scalable synthetic route to these molecules has long been hindered by the need for purification by high performance liquid chromatography (HPLC). We describe a fundamentally different

approach that uses a novel P(V) pyrene pyrophosphate reagent to generate derivatives that are purified by silica gel chromatography and converted to the desired compounds on scales vastly exceeding those achievable by HPLC. The power of this approach is demonstrated through the synthesis of a broad range of natural and unnatural nucleoside triphosphates using efficient, inexpensive, and operationally straightforward protocols.

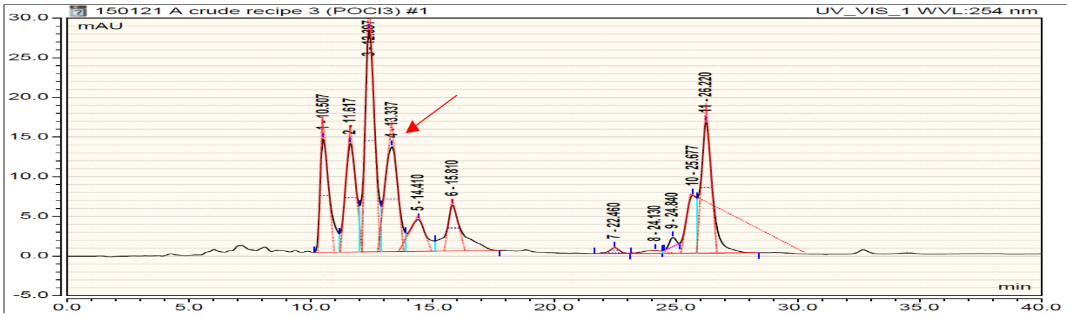
5.3 Introduction

Polymerases are among the most powerful tools in the molecular biology arsenal, permitting researchers to precisely synthesize the DNA sequences of genes, gene clusters, and entire genomes for diverse applications ranging from basic research to biotechnology and medicine. [1,2] Polymerase-mediated DNA synthesis has reduced the cost of DNA sequencing by allowing next-generation sequencing (NGS) platforms to read massive numbers of amplicons and has enabled digital archiving by compressing information into strands of DNA sequences that can be read by NGS analysis and decoded by bioinformatic assembly. [3,4] Engineered polymerases, developed by directed evolution, have also grown in demand by enabling the synthesis of artificial genetic polymers with backbone structures that are distinct from those found in nature. [5,6] Such enzymes are now used to support the evolution of affinity reagents (aptamers) and catalysts with molecular compositions that are better equipped to function in harsh biological environments. [7-10]

Today, no single strategy has been developed that can be applied universally to the synthesis of all nucleoside triphosphates. Natural DNA triphosphates (dNTPs) used to support traditional DNA synthesis applications are manufactured using enzymatic methods, while modified nucleoside triphosphates developed for biotechnology and medicinal

purposes are produced by chemical synthesis. [11-13] In general, enzymatic strategies suffer from high substrate specificity (low tolerance for analogs), while synthetic routes struggle with functional group compatibility, difficult reaction conditions, and yield. For example, attempts to synthesize α-L-threofuranosyl adenosine triphosphates using the classic Yoshikawa or Ludwig-Eckstein methods **(Fig. 5.1)** produced complex mixtures of phosphorylated compounds with the desired compound present as a minor product. [14, 15] Even more problematic is the requirement for both strategies to separate highly polar nucleoside triphosphates from polar side products using high performance liquid chromatography (HPLC). In academic environments, HPLC purification is a tedious process that limits the scale of nucleoside triphosphate synthesis to a few tens of milligrams of compound per week and requires subsequent freeze-drying steps that can lead to compound degradation.

Yoshikawa method:



Ludwig-Eckstein method

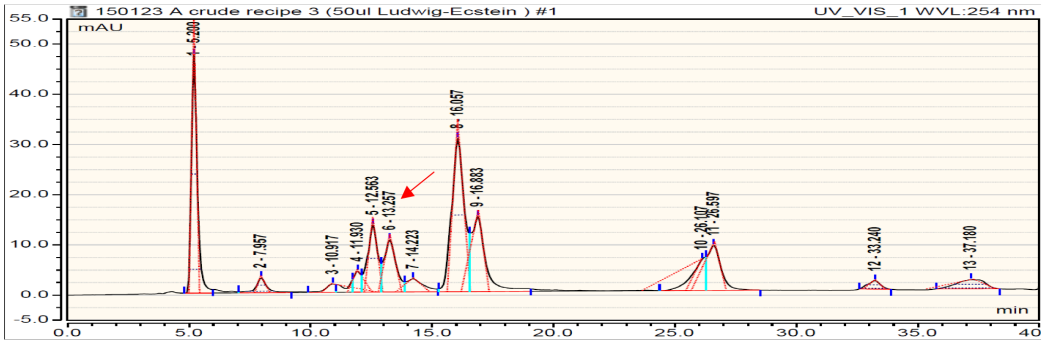


Figure 5.1. HPLC analysis of a-L-threofuranosyl adenosine triphosphate using the traditional Yoshikawa or Ludwig-Eckstein methods. Red arrow indicates the location of the desired peak in the crude reaction mixture.

(Figure was used with permission from . *J. Am. Chem. Soc.* **2019**. *141*, 34, 13286-13289)

Efforts to synthesize nucleoside triphosphates independent of HPLC purification first appeared in 2014 with a phosphorylation strategy that involves iterative cycles of coupling, oxidation, and deprotection.[16-18] This is a P(III)-based reagent method that relies on derivatized phosphoramidite reagents to facilitate P-O bond formation. Free nucleosides are converted to nucleoside triphosphates through the production of P(III)-P(V) anhydride intermediates that are oxidized and deprotected during each cycle of P(V) phosphate generation. [16] The iterative strategy benefits from high yields, rapid coupling rates, and convenient (non-dry) reaction conditions but is marked by several disadvantages that include: (1) the need for quantitative functional group transformations, which are necessary to avoid the accumulation of unwanted side products; (2) prolonged deprotection conditions due to slow removal of the second P-O protecting group on the phosphorylating reagent; (3) incompatibility with certain chemical reagents and nucleoside protecting groups; and (4) a lengthy pathway that involves 9 functional group transformations to convert each nucleoside to its corresponding triphosphate.[16-18]

In attempt to address these concerns, we conceived a fundamentally different approach for the synthesis of natural and modified nucleoside triphosphates that required invention of a novel P(V)-based organic pyrophosphate reagent to mediate P-O bond formation between the a- and b-phosphate positions. Our vision was to establish an Hoard-Ott-like

procedure (**Fig. 5.2**) using nucleoside and pyrophosphate reagents that were suitably hydrophobic so as to generate nucleoside triphosphate derivatives that were amenable to purification by silica gel chromatography but could be readily converted to the desired compound using standard deprotection and precipitation conditions.[19] We felt that this type of convergent synthesis strategy would provide a direct and scalable route to natural and chemically modified nucleoside triphosphates using operationally simple protocols that process and discovery chemists would find appealing. Critical to this effort was the need to establish a new class of organic pyrophosphate reagents that contain a large hydrophobic moiety attached to the pyrophosphate group via a cleavable linker and to demonstrate that these reagents could mediate the synthesis of nucleoside triphosphates. Reducing this concept to practice involved a systematic evaluation of three key components: (1) hydrophobic groups necessary to construct a stable organic pyrophosphate reagent that is a solid compound at room temperature and exhibits minimal hygroscopic properties; (2) a strong leaving group to activate the nucleoside monophosphate for nucleophilic attack by the pyrophosphate reagent; and (3) a suitable Lewis acid to accelerate phosphodiester bond formation (**Fig. 5.2**).

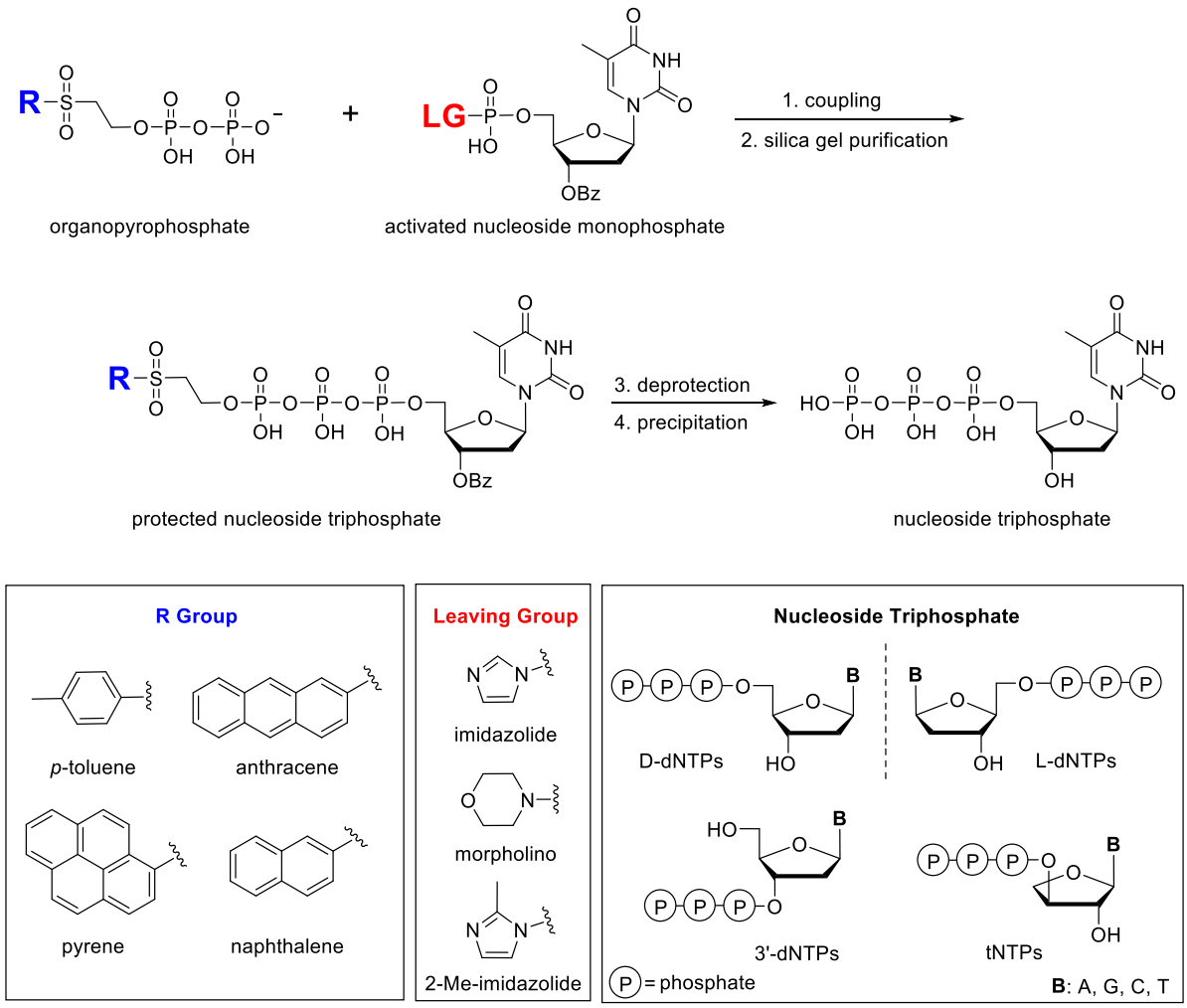


Figure 5.2. Strategy and methodology development. Vision for the scalable synthesis of natural and modified nucleoside triphosphates using a new class of organopyrophosphate reagents. R, the aromatic substitution attached to a cleavable sulfonylethyl linker. LG, leavinggroup moiety. B, nucleobase moieties: adenine (A), guanine (G), cytosine (C), thymine (T).

(Figure was used with permission from . *J. Am. Chem. Soc.* **2019**. *141*, 34, 13286-13289)

The goals of our study were ultimately met through a systematic exploration and subsequent optimization of reaction conditions which revealed that pyrene, 2-methylimidazole, and zinc chloride provided the optimal hydrophobic moiety, leaving group,

and Lewis acid catalyst, respectively, necessary to mediate the synthesis of nucleoside 5 triphosphates in their fully protected form. The crystal structure of the pyrene pyrophosphate reagent verified correct synthesis of the organic pyrophosphate reagent (**Fig 5.3**).

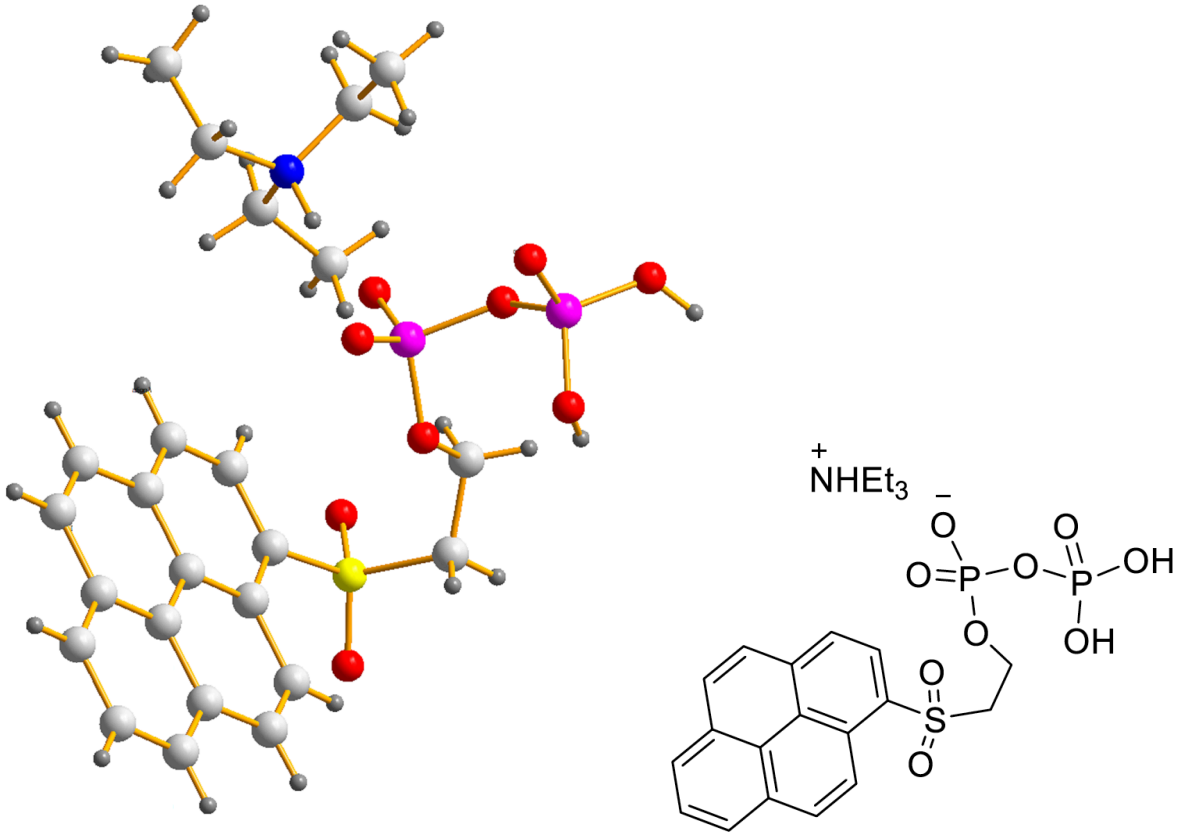


Figure 5.3. ORTEP representation of 2-(pyrenesulfonylethyl) pyrophosphate (7) with thermal ellipsoids at the 50% probability level. Crystals were prepared using the vapor diffusion method. The compound was dissolved in a minimal amount of methanol in a 2 mL open vial, which was placed in a 20 mL sealed vial containing 4 mL of anhydrous toluene. The crystal grew over a 1-month period under vapor equilibrium conditions at room temperature.

(Figure was used with permission from . *J. Am. Chem. Soc.* **2019**, *141*, 34, 13286-13289)

5.5 Results and Discussions

To test the feasibility of this approach, we applied the pyrene pyrophosphate method to the synthesis of natural thymidine-5'-triphosphate (dTTP). As shown in **Figure 5.4**, activated thymidine-5'-monophosphate is readily converted to a fully protected thymidine-5'-triphosphate derivative, which was purified by silica gel chromatography, deprotected, and precipitated to obtain the desired product as a highly pure compound in sodium form. In this reaction, 10 molar equivalents of $ZnCl_2$ enabled a near quantitative coupling (>95%) of the pyrene pyrophosphate reagent to the activated nucleoside monophosphate after 3 hours of stirring in DMF at 24 ^{o}C. The phosphorylation reaction also generated trace amounts (1-2%) of unwanted side products, including dinucleotide diphosphate, that were removed by silica gel purification using a 10% H_2O/isopropanol mobile phase containing 1% Hünig's base [diisopropylethylamine, DIPEA] as an organic counterion. Following purification, the triphosphate intermediate was deprotected with concentrated ammonium hydroxide (33% aq.) and precipitated as the sodium salt using standard conditions. Analytical HPLC analysis reveals that chemically synthesized dTTP was equivalent in purity to a commercial standard (**Fig. 5.4c**), demonstrating that the process of the nucleotide triphosphate synthesis is capable of generating material that is identical in purity to commercial compounds obtained by conventional enzymatic synthesis.

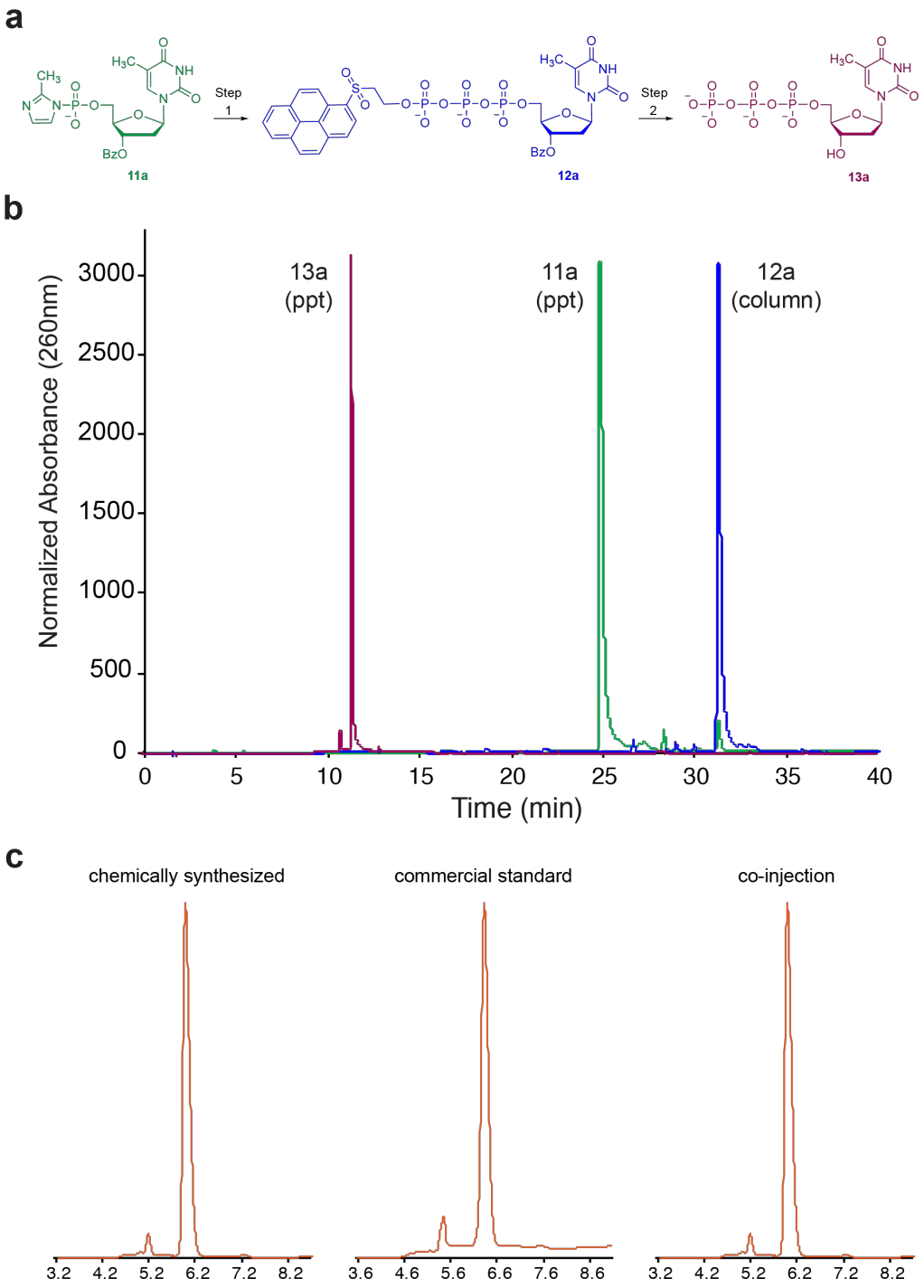


Figure 5.4. Synthesis and purity comparison of thymidine triphosphate. (a) Application of the organopyrophosphate reagent to the synthesis of natural thymidine-5'-triphosphate. (b) HPLC traces provided for the activated monophosphate, fully protected triphosphate

intermediate after silica gel purification, and thymidine triphosphate after precipitation as the sodium salt. (c) Comparison of chemically synthesized thymidine 3'-triphosphate (dTTP) with a commercial standard. thymidine-5'-triphosphate (the mirror image form of natural D-dTTP).

(Figure was used with permission from . *J. Am. Chem. Soc.* **2019**, *141*, 34, 13286-13289)

Because enzymatic approaches are not compatible with most nucleic acid analogs, we decided to broaden the scope of the reaction by applying the pyrene pyrophosphate method to a select panel of thymidine analogs (**Fig. 5.5**). The list of target molecules included thymidine-3'- triphosphate (a regioisomer of natural dTTP), **α** -L-threofuranosyl thymidine 3'-triphosphate (tTTP, a building block for threose nucleic acid, TNA), [20] and L-DNA-thymidine-5'- triphosphate

The last two examples have become valuable compounds for biomedical research due to recent advances in polymerase engineering that have enabled the synthesis of artificial genetic polymers with non-natural sugar-phosphate backbones, [21,22] some of which have been used to evolve biologically stable aptamers and catalysts. [7-10] As with natural dTTP, each unnatural thymidine triphosphate was obtained in highly pure form (>95%) as the desired sodium salt. Occasionally, we did observe a small amount of nucleoside diphosphate (dNDP) (~1-2%) after deprotection of the protected nucleoside triphosphates. dNDPs are common contaminants of commercial nucleotide triphosphates due to the slow

hydrolysis of ATP to ADP. [23] However, these molecules do not interfere with normal DNA synthesis as natural DNA polymerases are highly specific for dNTP substrates. [24]

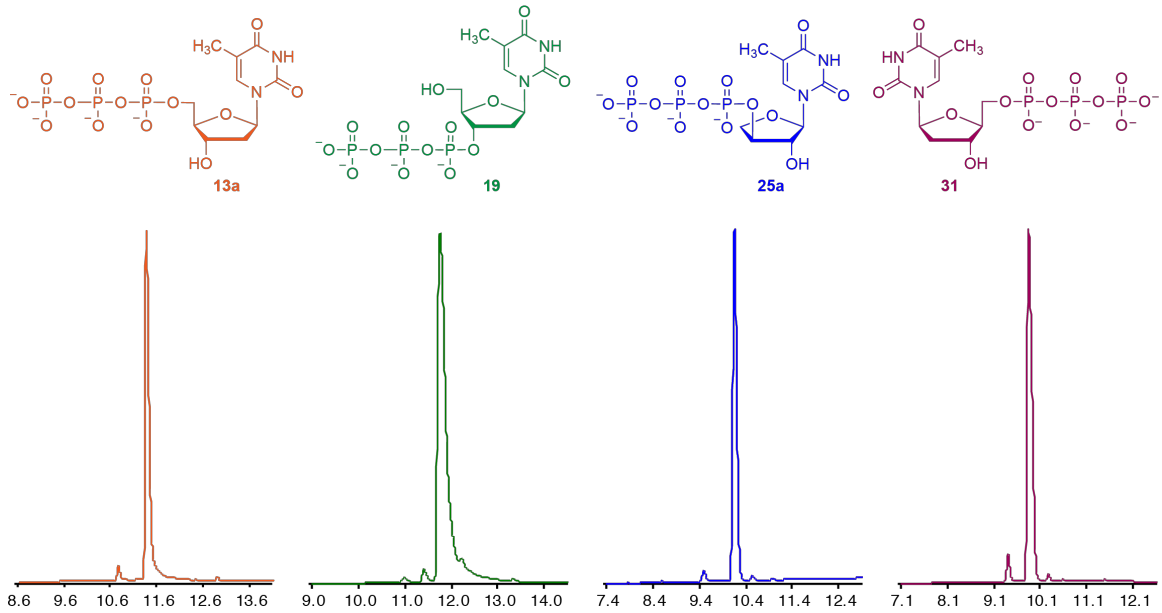


Figure 5.5. **Synthesis of thymidine triphosphate analogs**. The organopyrophosphate reagent was applied to the synthesis of thymidine 3'-triphosphate, a-L-threofuranosyl thymidine 3'- triphosphate, and L-thymidine 5'-triphosphate. HPLC traces provided after precipitation as the sodium salt.

(Figure was used with permission from . *J. Am. Chem. Soc.* **2019**, *141*, 34, 13286-13289)

Next, we prepared a complete set of DNA nucleoside triphosphates (dNTPs) using the pyrene pyrophosphate method. Each activated nucleoside monophosphate was converted to the desired nucleoside triphosphate following the standard procedure of pyrene pyrophosphate coupling, purification by silica gel chromatography, deprotection with concentrated ammonium hydroxide, and precipitation as the sodium salt (**Fig. 5.6a**). Concurrently, we also synthesized all four TNA nucleoside triphosphates (tNTPs), which are

the substrates for engineered TNA polymerases developed by directed evolution (**Fig. 5.6b**)[21]. In all cases, analytical HPLC chromatograms reveal high conversion (>95%) of activated nucleoside triphosphate into the fully protected nucleoside triphosphate. Following deprotection and precipitation, the desired DNA and TNA nucleoside triphosphates were produced in highly pure form on scales (>500 milligrams) and timeframes (2-3 days) vastly exceeding those of traditional protocols.

Since nucleoside triphosphates would ultimately be used as substrates for oligonucleotide synthesis, we chose to investigate the use of these reagents in a conventional DNA synthesis assay. One critical question that our study aimed to address is whether chemically synthesized DNA triphosphates function with same level of efficiency as enzymatically produced dNTPs obtained from commercial venders. To address this question, the polymerase chain reaction (PCR) was used to amplify an arbitrary DNA sample in reaction mixtures that contained either chemically synthesized or commercial dNTPs. We chose to amplify a 200 base pair segment that defines the finger subdomain of an archaeal DNA polymerase isolated from the thermophilic species Thermococcus kodakarensis (Kod). [25] Analysis of the resulting PCR amplicons by agarose gel electrophoresis confirmed that chemically synthesized dNTP substrates function identically to commercial dNTPs, as both reactions produce equivalent amounts of DNA at each cycle of PCR amplification (**Fig. 5.6a**). Moreover, DNA sequencing of the amplified product failed to identify any instances of insertions, deletions, or mutations, indicating that the reactions proceed with high template-sequence fidelity.

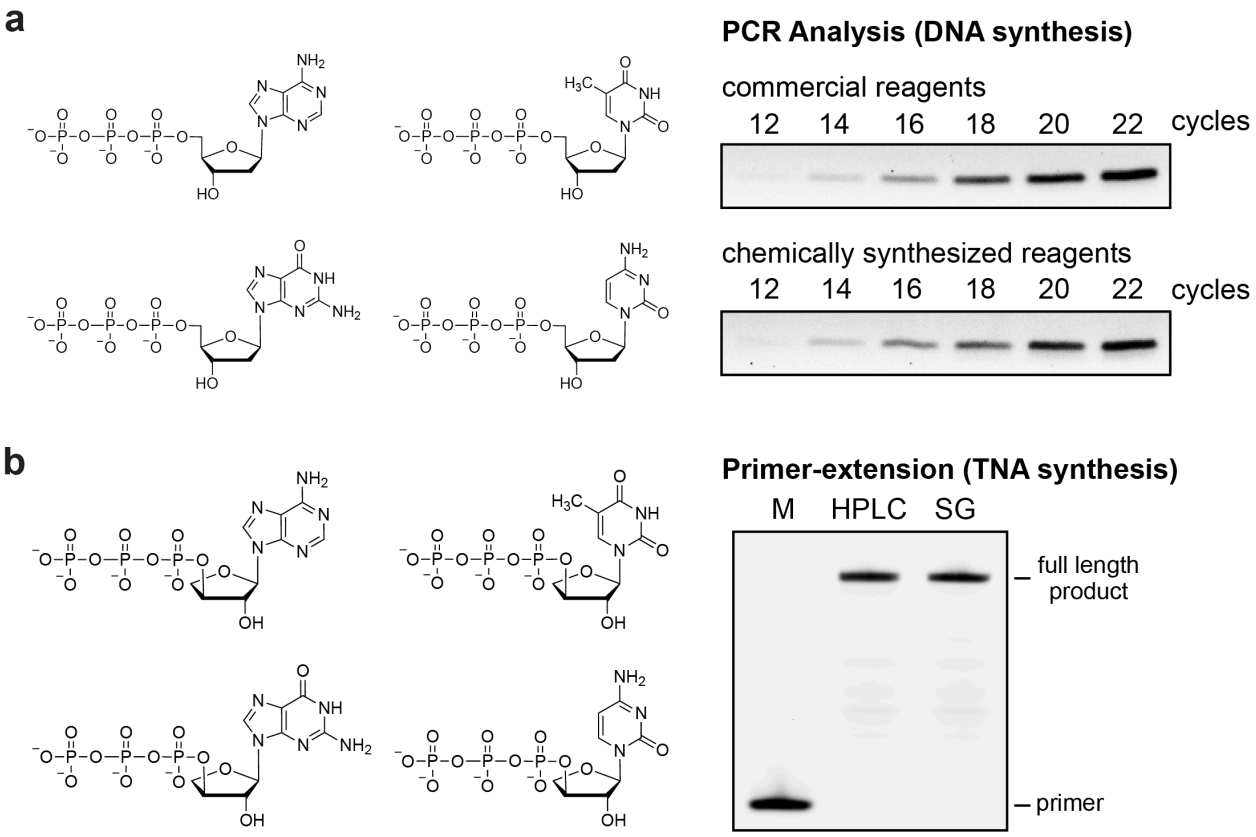


Figure 5.6. Polymerase activity assay of chemically synthesized DNA and TNA nucleoside triphosphates. All four DNA and TNA nucleoside triphosphates were constructed using the organic pyrophosphate method. (a) Chemically synthesized dNTPs were evaluated against commercial reagents in a standard PCR reaction. (b) Chemically synthesized tNTPs purified by silica gel chromatography (SG) were evaluated against HPLC purified tNTPs (HPLC) in a polymerase-mediated primer-extension reaction using Kod-RI TNA polymerase.

(Figure was used with permission from . *J. Am. Chem. Soc.* **2019**, *141*, 34, 13286-13289)

Next, we evaluated the set of chemically synthesized TNA triphosphates as substrates for TNA synthesis using an engineered TNA polymerase. [25] In this assay, a DNA primer-template complex was extended with TNA using either newly synthesized tNTPs or tNTPs obtained by a traditional Hoard-Ott-like approach that requires HPLC purification. [26] Analysis of the resulting primer-extension reactions by denaturing polyacrylamide gel electrophoresis

reveals that the two sets of TNA substrates generate equivalent amounts of full-length TNA product after 3 hours of incubation at 55 ^{o}C (**Fig. 5.6b**). This result, along with the PCR assay, confirms that the pyrene pyrophosphate method produces high quality nucleoside triphosphates that function as substrates for natural and engineered polymerases.

Historically, the synthesis of chemically modified nucleoside triphosphates has been limited by such factors as scalability, low yields, difficult reaction conditions, and tedious purifications protocols. [11-13] Efforts to overcome these problems have resulted in an astonishing number of publications, nearly all of which require HPLC purification. [11-13] Here, we describe a synthetic route to nucleic acid building blocks that (1) is not constrained to natural substrates; (2) is scalable to gram quantities of material; (3) eliminates the requirement for HPLC purification; and (4) avoids the need for freeze drying. The pyrophosphate reagent and activated nucleoside monophosphates described in this study are readily produced on scales of 5-10 grams using synthetic methodology that is inexpensive, straightforward and high yielding. The availability of these reagents makes it possible to generate 1.0 gram of nucleoside triphosphate, which is enough material to perform 5 liters of PCR or 100,000 PCR reactions with a standard 50 μL volume. Unlike conventional tributyl ammonium pyrophosphate, pyrene pyrophosphate is a solid at room temperature and stable to atmospheric conditions, allowing it to be used without specialized equipment and stored for long periods of time without loss of activity. Last, the pyrene pyrophosphate method was found to be superior for synthesizing chemically modified nucleotides that were challenging to synthesize using the classic Yoshikawa or Ludwig-Eckstein methods. [14, 15] Based on these considerations, we believe that the current

approach provides an alternative paradigm for synthesizing large quantities of nucleoside triphosphates for emerging applications in biotechnology and molecular medicine.

5.5 Experimental Details

General Information

All moisture sensitive reactions were carried out in anhydrous solvents under an argon or nitrogen atmosphere. All commercial reagents, solvents and anhydrous solvents were used without further purification. Solvents used for triphosphate purification by silica column chromatography were purchased as HPLC grade. Reaction progresses were 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 (40-63 um particle size).

^{1}H, ^{13}C, or ^{31}P Nuclear Magnetic Resonance (NMR) spectra were recorded at room temperature either on a Bruker DRX 400 or 500 MHz spectrometer at the University of California, Irvine NMR Facility. The 1H and 13C NMR chemical shifts (δ) are reported in parts per million (ppm) with tetramethylsilane or deuterium solvent as an internal reference. The 31P NMR chemical shifts are proton decoupled and reported in parts per million referenced to the deuterium solvent calibrated in ^{1}H NMR. Peaks multiplicity are designated with the following abbreviations: s, singlet; d, doublet; dd, doublet of doublets; ddd, doublet of doublet of doublets; t, triplet; m, multiplet; br, broad. X-ray crystal data was collected by Bruker SMART APEX II diffractometer at University of California, Irvine X-ray Crystallography Facility.

HPLC analysis was performed on a reverse-phase C18 150 Å~ 4.6 mm^2 column with 5 μm particle size (Agilent, USA) using a mobile phase of 100 mM triethylammonium acetate buffer (pH 7.0)/acetonitrile. 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 Core 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_{280}: 13100 $M^{-1}cm^{-1}$. G, ε_{253}: 13700 $M^{-1}cm^{-1}$. A, ε_{259}: 15200 $M^{-1}cm^{-1}$.)

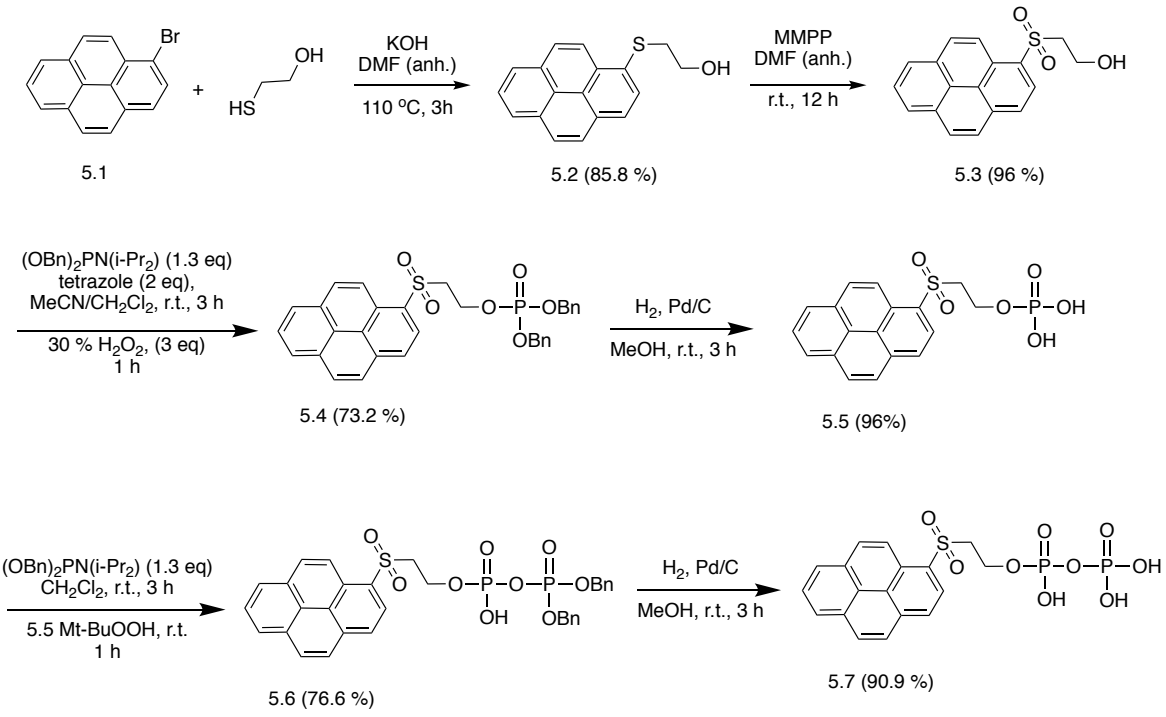
SYNTHESIS OF PYRENE PYROPHOSPHATE

2-(Pyrenethio)ethanol (5.2).

To a solution containing 9.98 g (177.8 mmol) of potassium hydroxide in 250 mL of anhydrous DMF was added 12.5 mL (177.8 mmol) of 2-mercaptoethanol. The reaction mixture was heated to 80 ^oC with stirring until potassium hydroxide dissolved. A premade solution containing 25 g (88.9 mmol) of 1-bromopyrene (5.1) in 200 mL of anhydrous DMF was dropwise added to the reaction and stirred at 110 ^oC for 3 hours. Then, the reaction was cooled to room temperature, condensed to 100 mL of solution under diminished pressure and diluted into 500 mL of CH_2Cl_2. The organic layer was washed with 200 mL of H_2O four times. And the combined aqueous layer was back extracted with 200 mL of CH_2Cl_2 two times. The organic layers were combined and evaporated to the dryness under diminished pressure. The crude product was purified by silica column chromatography with eluents (CH_2Cl_2/Hexane, 25% to 50%, then EtOAc/CH_2Cl_2, from 0% to 2%) to afford the compound

2 as a yellowish solid; yield: 21.3 g (85.8%); silica gel TLC (EtOAc/Hexane, 1:2) R_f = 0.35; ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.61 (d, 1H, J = 9.2 Hz), 8.07-8.00 (m, 4H), 7.93-7.81 (m, 4H), 3.71 (t, 2H, J = 6.0 Hz), 3.18 (t, 2H, J = 6.0 Hz); HRMS (ESI-TOF) calcd. for $C_{18}H_{14}OSNa$ $[M+Na]^+$ 301.0663; found 301.0658. This is a known compound previously reported in the literature.[1]

Scheme 5.1: Synthesis of 1-(2-(pyrenesulfonyl)ethyl)pyrophosphate



(Scheme was used with permission from . *J. Am. Chem. Soc.* **2019**, *141*,34,13286-13289)

2-(Pyrenesulfonyl)ethanol (5.3).

To a solution containing 71 g (114.9 mmol, 80% technical grade) of magnesium bis(monoperoxyphthalate) hexahydrate in 450 mL of anhydrous DMF was slowly added a pre-made solution of 21.3 g (76.6 mmol) of compound **(5.2)** in 150 mL of anhydrous DMF at 0 °C. The reaction was stirring at room temperature for 12 h. At which the TLC showed

the reaction was finished, the reaction was dropwise added to a solution containing 2 L of satd. $NaHCO_3$(aq) while stirring. The precipitate was washed with 300 mL of H_2O twice, and dried under high vacuum to afford the compound **5.3** as a yellowish solid; yield: 22.8 g (96%); silica gel TLC (EtOAc/CH_2Cl_2, 1:20) R_f = 0.22; ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.00 (d, 1H, J = 9.2 Hz), 8.71 (d, 1H, J = 8.4 Hz), 8.34-8.31 (m, 3H), 8.27-8.23 (m, 2H), 8.14-8.09 (m, 2H), 4.00 (t, 2H, J = 5.2 Hz), 3.64 (t, 2H, J = 5.2 Hz); ^{13}C NMR (125.8 MHz, $CDCl_3$-10% DMSO-d_6) δ 135.5, 130.8, 130.7, 130.6, 130.3, 129.9, 128.9, 127.5, 127.2, 127.1, 127.0, 126.9, 125.0, 124.1, 123.9, 122.4, 58.6, 56.1; HRMS (ESI-TOF) calcd. for $C_{18}H_{14}O_3SNa$ $[M+Na]^+$ 333.0561; found 333.0571.

Dibenzyl-1-(2-(pyrenesulfonyl)ethyl) monophosphate (5.4).

To a mixture containing 5 g (16.1 mmol) of 2-(pyrenesulfonyl)ethanol **(5.3),** 2.03 g (29.0 mmol) of tetrazole in 106 mL of anhydrous solution (MeCN/CH_2Cl_2, 1:1) was slowly added 6.64 mL (20.9 mmol) of $(BnO)_2PN(i\text{-}Pr)_2$ at room temperature. The reaction was sonicated until the solid disappeared (otherwise add additional 20 mL of anhydrous DMF) and the mixture was stirring at room temperature for 3 h. Then reaction was cooled to -40 ^{o}C followed by adding 15 mL of H_2O_2 (33% in H_2O) for additional 1 h stirring at room temperature. The solution was condensed to 20 mL under diminished pressure and poured to 150 mL EtOAc. The organic layer was washed with 100 mL brine, 200 mL satd. $NaHCO_3$, 200 mL H_2O, dried over $MgSO_4$ and evaporated under diminished pressure. The crude product was purified by silica column chromatography with eluent (CH_2Cl_2/Hexane, from 25% to 100%, then MeOH/CH_2Cl_2, from 1% to 3%) to afford the yellowish solid product **5.4**; yield: 6.73 g (73.2 %); TLC (EtOAc/CH_2Cl_2, 1:20) R_f = 0.34; ^{1}H

NMR (400 MHz, $CDCl_3$) δ 8.89 (d, 1H, J = 9.2 Hz), 8.60 (d, 1H, J = 8.0 Hz), 8.15-8.11 (m, 3H), 8.01-7.94 (m, 3H), 7.82 (d, 1H, J = 8.8 Hz), 7.20-7.16 (m, 6H), 7.06-7.04 (m, 4H), 4.67 (d, 4H, *J* = 8.4 Hz), 4.36 (dd, 2H, *J* = 14, 6.0 Hz), 3.70 (t, 2H, J = 6.0 Hz); ^{13}C NMR (125.8 MHz, $CDCl_3$) δ 135.2, 135.2, 135.1, 130.5, 130.4, 129.9, 129.5, 128.7, 128.3, 128.3, 127.6, 127.4, 127.1, 127.0, 126.7, 126.6, 124.5, 123.9, 123.4, 122.0, 69.2 (d, $J_{C,P}$ = 5.8 Hz), 60.7 (d, $J_{C,P}$ = 5.0 Hz), 56.3 (d, $J_{C,P}$ = 7.2 Hz); ^{31}P NMR (162 MHz, $CDCl_3$) δ -0.41; HRMS (ESI-TOF) calcd. for $C_{32}H_{27}O_6PSNa$ $[M+Na]^+$ 593.1164; found 593.1158.

1-(2-(Pyrenesulfonyl)ethyl) monophosphate (5.5).

To a solution containing 6.73 g (11.8 mmol) of dibenzyl-1-(2-(pyrenesulfonyl)ethyl) monophosphate **(5.4)** in 50 mL of MeOH was purged with nitrogen gas. The solution was added to 0.1-0.2 mass equivalent of 10% Pd/C, and repurged with nitrogen gas followed by adding hydrogen gas. The mixture was stirred at room temperature for 3 h with monitoring by TLC ($MeOH/CH_2Cl_2$, 1:10 with 1% Et_3N). The resulting suspension was filtered over a pad of celite, washed with 100 mL of MeOH four times, evaporated, and then coevaporated with dry $C_2H_4Cl_2$ under reduced pressure to afford the product **5.5** as a yellowish solid; yield 4.42 g (96%); TLC ($MeOH/CH_2Cl_2$, 1:10 with 1% triethylamine) R_f = 0; 1H NMR (400 MHz, CD_3OD) δ 8.72 (d, 1H, *J* = 9.6 Hz), 8.44 (d, 1H, *J* = 8.4 Hz), 8.00- 7.90 (m, 4H), 7.79-7.75 (m, 2H), 7.66 (d, 1H, *J* = 8.8 Hz), 4.34 (dd, 2H, *J* = 12.4, 6.0 Hz), 3.82 (t, 2H, *J* = 6.0 Hz); ^{13}C NMR (125.8 MHz, CD_3OD+10% DMSO-d_6) δ 136.6, 131.7, 131.5, 131.5, 131.4, 130.8, 129.8, 128.4, 128.2, 128.0, 127.8, 127.7, 125.5, 125.2, 124.4, 123.2, 60.9 (d, $J_{C,P}$ = 4.2 Hz), 57.6 (d, $J_{C,P}$ = 7.2 Hz); ^{31}P NMR (162 MHz, CD_3OD) δ -0.56. HRMS (ESI-TOF) calcd. for $C_{18}H_{14}O_6PS$ $[M-H]^-$ 389.0249; found 389.0234.

1-(2-(Pyrenesulfonyl)ethyl)-(β-dibenzyl)pyrophosphate (5.6).

To the 4.42 g (11.3 mmol) of 1-(2-(pyrenesulfonyl)ethyl) monophosphate **(5.5)** in 39.3 mL of anhydrous dichloromethane was added 5.33 mL (15.8 mmol) of $(BnO)_2PN(i\text{-}Pr)_2$ at room temperature under a nitrogen atomosphore. After two hours stirring, the reaction was slowly added to 6.16 mL (33.9 mmol) of tert-butyl hydrogen peroxide (5.5 M in decane) at -40 ^{o}C and stirred for an additional 30 minutes at room temperature. The reaction was evaporated under diminished pressure and the crude product was purified by silica column chromatography with the eluents (MeCN then MeOH/CH_2Cl_2 from 1% to 3% containing 1% triethylamine) to afford the solid product **5.6**; yield: 6.5 g (76.6%); TLC (2:100 MeOH-CH_2Cl_2 with 1% triethylamine) R_f = 0.26; ^{1}H NMR (400 MHz, CD_3OD) δ 8.90 (d, 1H, *J* = 9.6 Hz), 8.56 (d, 1H, *J* = 8.0 Hz), 8.15-8.11 (m, 3H), 8.06-8.00 (m, 2H), 7.94 (t, 1H, *J* = 8.0 Hz), 7.87 (d, 1H, *J* = 8.8 Hz), 7.19 (m, 10H), 4.94 (d, 4H, *J* = 7.6 Hz), 4.32 (dd, 2H, *J* = 15.2, 6.8 Hz), 3.75 (t, 2H, *J* = 6.8 Hz); ^{13}C NMR (125.8 MHz, $CDCl_3$) δ 136.0, 135.9, 135.6, 130.9, 130.7, 130.7, 130.6, 130.0, 129.1, 128.5, 128.3, 127.8, 127.7, 127.3, 127.2, 127.1, 127.0, 125.1, 124.3, 124.0, 122.6, 69.3 (d, $J_{C,P}$ = 5.5 Hz), 59.7 (d, $J_{C,P}$ = 5.5 Hz), 57.0 (d, $J_{C,P}$ = 6.9 Hz); ^{31}P NMR (162 MHz, $CDCl_3$) δ -11.2 (d, *J* = 18.5 Hz), - 11.6 (d, *J* = 18.5 Hz); HRMS (ESI-TOF) calcd. for $C_{32}H_{27}O_9P_2SNa_2$ $[M\text{-}H+2Na]^+$ 695.0646; found 695.0660.

1-2-(Pyrenesulfonyl)ethyl)pyrophosphate (5.7).

To a solution containing 6.5 g of 1-(2-(pyrenesulfonyl)ethyl)-(β-dibenzyl)pyrophosphate **(4.6)** in 30 mL of MeOH was purged with nitrogen gas. The solution was added to 0.1-0.2 mass equivalent of 10% Pd/C and purged with nitrogen gas followed by adding hydrogen gas. The mixture was stirred at room temperature for 3 h with monitoring by TLC

(MeOH/CH_2Cl_2, 1:10 with 1% Et_3N). The resulting suspension was filtered over a pad of celite, washed with 100 mL of MeOH three times, evaporated, and coevaporated with dry $C_2H_4Cl_2$ under reduced pressure to afford the product as a yellowish solid **5.7**; yield: 4.27 g (90.9%); TLC (MeOH/CH_2Cl_2, 1:10 with 1% triethylamine) R_f = 0; ^{1}H NMR (400 MHz, CD_3OD) δ 8.93 (d, 1H, *J* = 9.2 Hz), 8.61 (d, 1H, *J* = 8.0 Hz), 8.29-8.22 (m, 4H), 8.13 (d, 1H, *J* = 8.8 Hz), 8.05-8.00 (m, 2H), 4.35 (m, 2H), 3.91 (t, 2H, *J* = 10.4 Hz); ^{13}C NMR (125.8 MHz, D_2O) δ 135.0, 130.7, 130.1, 129.7, 129.2, 129.0, 128.3, 127.3, 127.1, 127.0, 126.6, 126.3, 124.4, 123.6, 122.4, 121.8, 60.6, 57.3 (d, $J_{C,P}$ = 6.0 Hz); ^{31}P NMR (162 MHz, D_2O) δ -10.0 (d, *J* = 16.8 Hz), -11.5 (d, *J* = 16.8 Hz); HRMS (ESI-TOF) calcd. For $C_{18}H_{14}O_9P_2SNa_3$ $[M-2H+3Na]^+$ 536.9527; found 536.9521.

SYNTHESIS OF NUCLEOSIDE TRIPHOSPHATES

General procedure A: synthesis of protected nucleoside monophosphates. To a mixture containing suitably protected free nucleoside (compounds: 5.8a-d, 5.14, 5.20a-d, and 5.26) and 1.8 equivalents of tetrazole in an anhydrous solvent (MeCN/CH_2Cl_2, 1:1) was added to 1.3 equivalents of $(BnO)_2PN(i\text{-}Pr)_2$ under a nitrogen atmosphere. The mixture was stirred at room temperature for 1-3 hours and the reaction was monitored by TLC. At which the starting material was consumed, the reaction was slowly added excess H_2O_2 (33% in H_2O) at -40 ^{o}C and the resulting mixture was stirred at room temperature for 1 hour. The solution was then diluted with 15-20 times volume of CH_2Cl_2 and the organic layer was sequentially washed with saturated $NaHCO_3$(aq), brine, and water. The organic extracts were combined, dried with $MgSO_4$, and evaporated under diminished pressure. The crude residue was

purified by silica gel chromatography and the fractions containing the product were collected and evaporated to afford the product 5.9a-d, 5.15, 5.21a-d, or 5.27.

General procedure B: synthesis of nucleoside monophosphates. To a solution containing the protected nucleoside monophosphate (compounds 5.9a-d, 5.15, 5.21a-d, and 5.27) in MeOH was purged with nitrogen gas. The solution was added 0.1-0.2 mass equivalent of 10% Pd/C, and re-purged with nitrogen gas followed by adding hydrogen gas. The mixture was stirred at room temperature for 3-5 hours with monitoring by TLC ($MeOH/CH_2Cl_2$, 1:10 with 1% triethylamine). Then the solution was purged with nitrogen, filtered over a pad of celite, and washed with MeOH or MeOH containing 2% triethylamine. The filtrate was collected, evaporated and coevaporated with dry 1, 2-dichloroethane under reduced pressure to afford the product 5.10a-d, 5.16, 5.22a-d, or 5.28.

General procedure C : synthesis of activated nucleoside monophosphates. To a solution containing the nucleoside monophosphate (compounds 5.a-d, 5.16, 5.22a-d and 5.28) in anhydrous DMF under a nitrogen atmosphere was slowly added 5 equivalents of anhydrous triethylamine at 0 ^{o}C. After 5 minutes with stirring, 2 equivalents of 2- methylimidazole was added followed by 2 equivalents of triphenylphosphine. After 10 minutes of stirring at room temperature, 2 equivalents of 2, 2'-dipyridyl disulfide was added and stirring was continued for an additional 3 hours at room temperature with monitoring by analytical HPLC (mobile phase: MeCN/0.1 M TEAA buffer, from 0% to 50% over 40 minutes). After consumption of the starting material, the product was precipitated by adding the reaction dropwise with stirring to 300 mL of diethyl ether. The precipitate was collected by centrifuging at 4400 rpm for 15 minutes at room temperature. The supernatant was discarded, and the pellet was resuspended with minimal amount of CH_2Cl_2 or anhydrous DMF (DMF is used when the

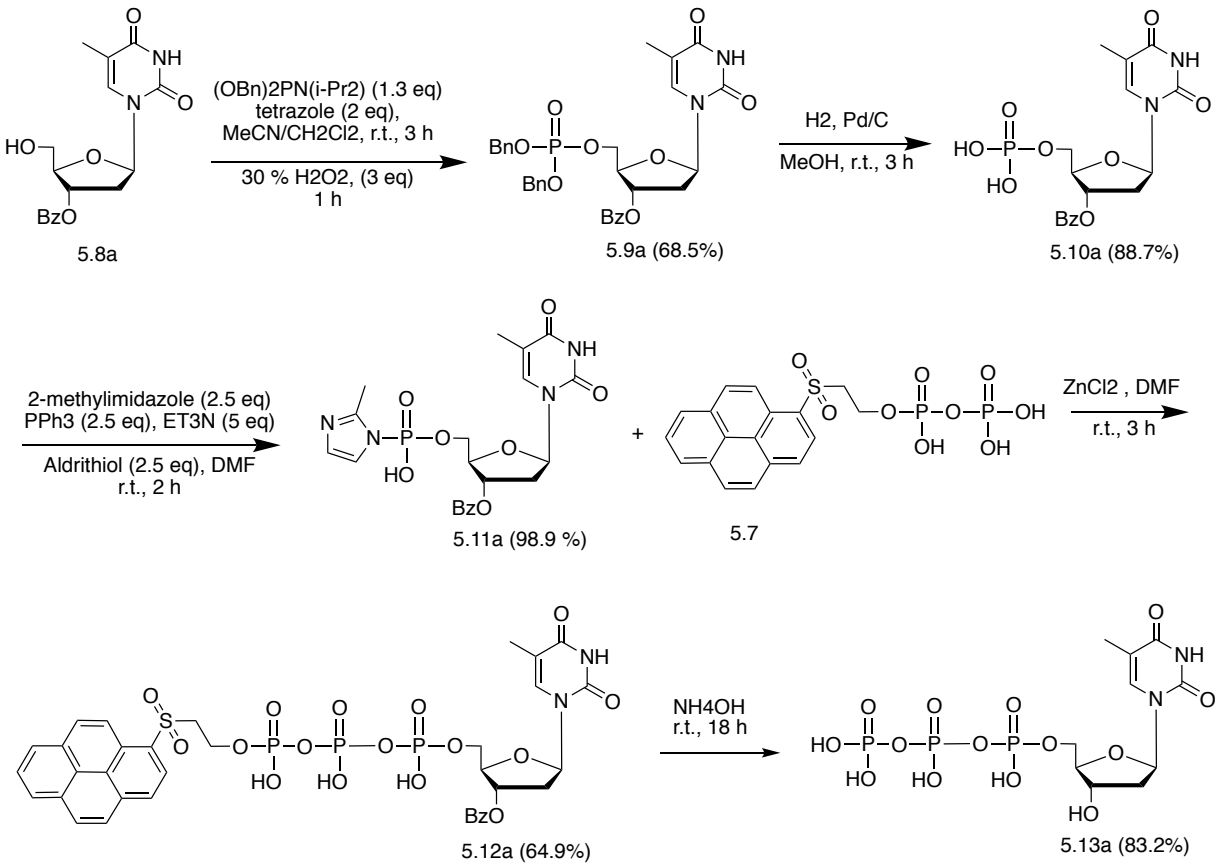
nucleoside monophosphate is not soluble in CH_2Cl_2). The solution was added dropwise to a premade solution of ether/ethyl acetate/triethylamine (5:10:1) containing 8 equivalents of sodium perchlorate for a second precipitation. The suspended solid was centrifuged at 4400 rpm for 15 minutes at room temperature, the supernatant was discarded, and the pellet was washed twice with 40 mL of mixed solvent (ether/ethyl acetate, 1:2), and dried under high vacuum to afford the product 5.11a-d, 5.17, 5.23a-d or 5.29.

General procedure D : synthesis of fully protected nucleoside triphosphates. To a mixture containing the activated nucleoside monophosphate (compounds 5.11a-d, 5.17, 5.23a-d and 5..29) and 1.2 equivalents of 1-(2-(pyrenesulfonyl)ethyl)pyrophosphate **(5.7)** was added along with 8-10 equivalents of a premade solution of $ZnCl_2$ (1.0 M in anhydrous DMF) under a nitrogen atmosphere. The mixture was stirred at room temperature for 3-5 hours and the reaction progress was monitored by HPLC (MeCN/0.1 M TEAA buffer, from 0% to 50% over 40 minutes). At which the starting material was consumed, the reaction was dropwise added to ethyl acetate or ether for precipitation. The precipitate was centrifuged and collected at 4400 rpm for 10 minutes at room temperature, and the supernatant was discarded. The pellet was resuspended by 20% H_2O in MeCN with 2%
Hunig's base and the solid was filtered by pyrex glass funnel with filter paper. The filtrate was collected and evaporated under diminished pressure and dry-packing loaded to silica gel for normal phase silica column chromatography. The fractions containing the product were collected and evaporated under diminished pressure at 30-40 ^oC. The product was resuspended with CH_2Cl_2 and insoluble silica gel was removed by filtration. The filtrate was collected and evaporated to dryness to afford the product 5.12a-d, 5.18, 5.24a-d, or 5.30.

General procedure E: synthesis of free nucleoside triphosphates. To a solution of

fully protected nucleoside triphosphate (compounds: 5.12a, 5.18, 5.24a, and 5.30) in 50 mL of 33% NH_4OH(aq) was stirred for 18 hours at room temperature or (compounds: 5..12b-d, and 5.24b-d) was stirred for 3 hours at 37 ^oC and 15 hours at room temperature in a sealed tube. After the reaction, the solvent was evaporated under diminished pressure. The solid was resuspended with Milli-Q water and the aqueous solution was washed with CH_2Cl_2 and ethyl acetate. The organic portion was discarded and the aqueous extract was collected, and evaporated under diminished pressure. The crude solid was resuspended with minimal amount of RNAse free water, filtrated by 0.22 um syringe filter and dropwise added to the forty times volume of acetone at room temperature containing 15 equivalents of sodium perchlorate. The resulting suspension was centrifuged at 4400 rpm for 15 minutes at room temperature. The supernatant was discarded and the pellet was washed with organic solution (acetone/CH_2Cl_2, 10:1) twice to afford the product 5.13a-d, 5.19, 5.25a-d,or 5.31.

Scheme 5.2: Synthesis of thymidine-5'-triphosphate (dTTP, 5.13a)



(Scheme was used with permission from . *J. Am. Chem. Soc.* **2019**, *141*, 34, 13286-13289)

3'-O-Benzoyl-2'-deoxythymidine-5'-dibenzylmonophosphate (5.9a).

General procedure A with 1g (2.9 mmol) of 3'-O-benzoyl-2'-deoxythymidine **5.8a**,[2] 364.4 mg (5.2 mmol) of tetrazole, 28.9 mL of anhydrous solution ($MeCN/CH_2Cl_2$, 1:1), 1.3 mL (3.8 mmol) of dibenzyl-N, N-diisopropylphosphoramidite for 3 hours reaction at room temperature. Then, 6 mL of 30% H_2O_2(aq) for 1 hour oxidation reaction at room temperature. Column chromatography with eluents ($MeOH/CH_2Cl_2$, from 1% to 1.4%) to afford the product **5.9a** as a white solid; yield: 1.2 g (68.5%); TLC ($MeOH/CH_2Cl_2$, 1:40) R_f = 0.23; ^1H NMR (400 MHz, CD_3OD) δ 8.03-8.00 (m, 2H), 7.62-7.58 (m, 1H), 7.48-7.45 (m, 3H),

7.36- 7.29 (m, 10H), 6.32 (dd, 1H, *J* = 8.4, 6.0 Hz), 5.42-5.40 (m, 1H), 5.12-5.07 (m, 4H), 4.34-4.28 (m, 3H), 2.47 (ddd, 1H, *J* = 14.4, 6.0, 2.0 Hz), 2.26-2.18 (m, 1H), 1.77 (d, 3H, *J* = 1.2 Hz); HRMS (ESI-TOF) calcd. for $C_{31}H_{31}N_2O_9PNa$ $[M+Na]^+$ 629.1665; found 629.1664.

This is a known compound previously reported in the literature.[2]

3'-O-Benzoyl-2'-deoxythymidine-5'-monophosphate (5.10a).

General procedure B with 1.2 g (1.9 mmol) of 5.9a, 50 mL of MeOH, and 200 mg of 10% Pd/C for 3 hours stirring at room temperature. The suspension was filtered over a pad of celite, and washed with 100 mL of MeOH containing 2% triethylamine four times to afford the product **5.10a** as a white foam of triethylammonium salt; yield: 1.11 g (88.7%); TLC ($MeOH/CH_2Cl_2$, 1:10 with 1% triethylamine) $R_f = 0$; ^1H NMR (400 MHz, DMSO-d^6) δ 11.41 (s, 1H), 8.09 (d, 2H, *J* = 8.0 Hz), 7.77-7.74 (m, 2H), 7.62 (t, 2H, J = 7.6 Hz), 6.41 (dd, 1H, J = 8.0, 6.8 Hz), 5.57 (s, 1H), 4.41 (s, 1H), 4.18 (m, 2H), 2.58-2.57 (m, 2H), 1.88 (s, 3H); ^{13}C NMR (125.8 MHz, DMSO-d^6) δ 168.6, 167.3, 152.7, 138.6, 135.6, 131.0, 130.4, 130.3, 113.0, 86.2, 85.1 (d, J C,P = 8.7 Hz), 78.0, 66.2, 37.9, 31.9, 13.2; ^{31}P NMR (162 MHz, DMSO-d^6) δ 0.24; HRMS (ESI-TOF) calcd. for $C_{17}H_{18}N_2O_9PNa_2$ $[M-H+2Na]^+$ 471.0545; found 471.0539.

3'-O-Benzoyl-2'-deoxythymidine-5'-phosphor-2-methylimidazolide (5.11a).

General procedure C with 1.11 g (1.8 mmol) of 4.10a, 7.8 mL of anhydrous DMF, 1.63 mL (11.7 mmol) of triethylamine, 385 mg (4.7 mmol) of 2-methylimidazole, 1.23 g (4.7 mmol) of triphenylphosphine, 1.04 g (4.72 mmol) of dipyridyl disulfide for 2 hours reaction at room temperature. First precipitation was achieved with 250 mL of diethyl ether. The product was resuspended with 10 mL of DMF and dropwise added to the solution containing 2.01 g of

sodium perchlorate, 15 mL of triethylamine in 300 mL of ethyl acetate for second precipitation. The product was afforded as a white solid **5.11a**; yield: 0.89 g (98.9%); ^{31}P NMR (162 MHz, D_2O) δ -6.83; HRMS (ESI-TOF) calcd. for $C_{21}H_{23}N_4O_8PNa$ $[M+Na]^+$ 513.1151; found 513.1141.

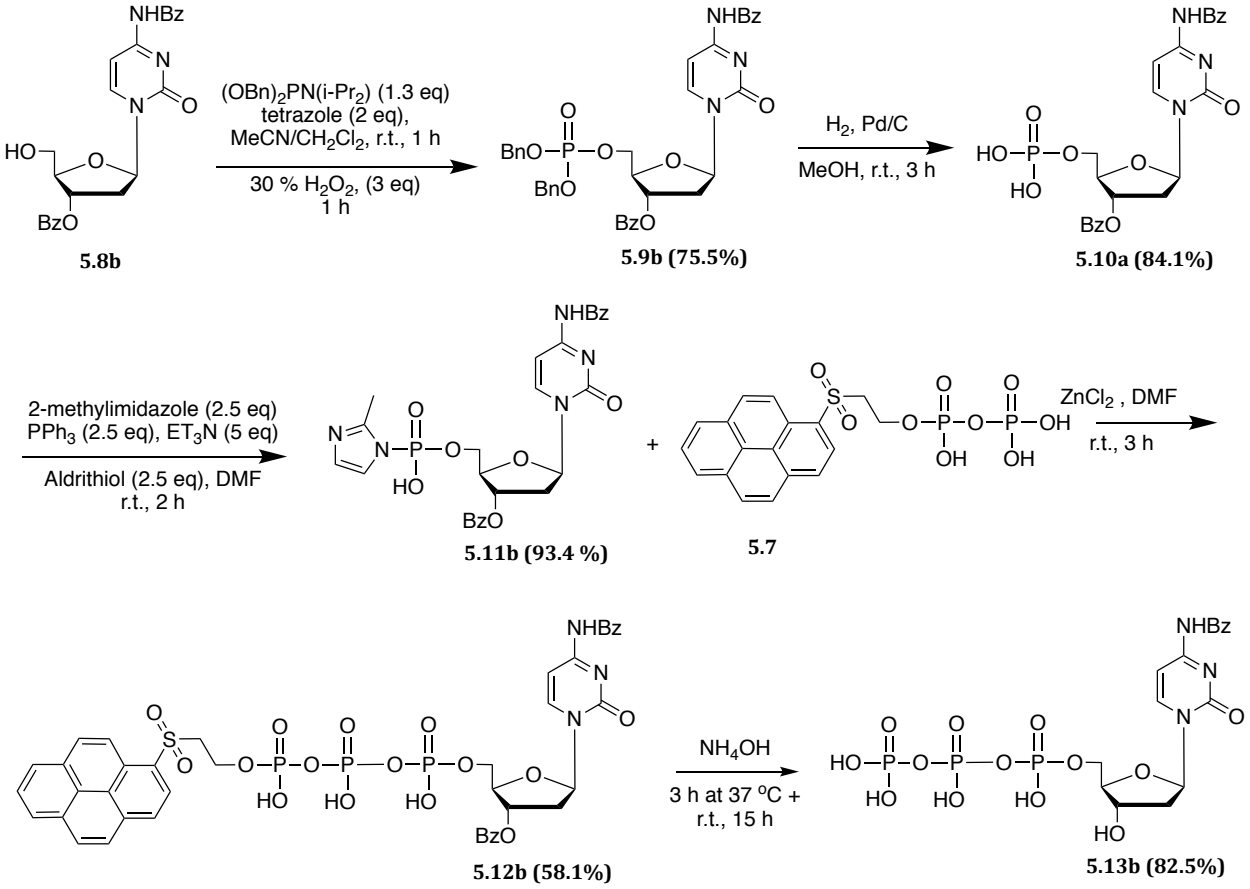
3'-O-Benzoyl-2'-deoxythymidine-5'-(γ-(2-(pyrenesulfonyl)ethyl))triphosphate (5.12a).

General procedure D with 0.89 g (1.7 mmol) of 4.11a, 1.09 g (1.91 mmol) of 1-2-(pyrenesulfonyl)ethyl) pyrophosphate (5.7) and 13.92 mL (13.92 mmol) of $ZnCl_2$ solution (1.0 M in anhydrous DMF) for 3 hours stirring at room temperature. Then crude was precipitated by dropwise adding the solution to the 300 mL of ethyl acetate with stirring. After centrifugation, the pellet was resuspended by 20% H_2O/MeCN containing 2% Hunig's base and filtered by pyrex glass funnel. The filtrate was evaporated and the crude material was purified by silica column chromatography with eluents (H_2O/(isopropanol- MeCN 1:1) from 2% to 7% containing 1% diisopropylethylamine (DIPEA) to obtain pure **5.12a**; yield: 1.43 g (64.9%); TLC (H_2O/acetone 1:10 with 2% diisopropylethylamine) R_f = 0.29; ^{1}H NMR (400 MHz, D_2O) δ 8.25 (d, 1H, *J* = 8.8 Hz), 8.10 (d, 1H, *J* = 7.6 Hz), 7.41 (d, 1H, *J* = 9.2 Hz), 7.33-7.27 (m, 3H), 7.19-7.13 (m, 4H), 6.94-6.89 (m, 3H), 6.68 (t, 2H, *J* = 6.4 Hz), 5.61 (s, 1H), 5.16 (s, 1H), 4.49 (s, 2H), 4.15-4.07 (m, 2H), 3.78 (m, 3H) 2.04- 1.94 (m, 2H), 1.70 (s, 3H); ^{31}P NMR (162 MHz, D_2O) δ -10.22 (brs, 2P), -20.50 (brs, 1P); HRMS (ESI-TOF) calcd. For $C_{35}H_{31}N_2O_{17}P_3SNa_3$ $[M-2H+3Na]^+$ 945.0250; found 945.0261.

2'-deoxythymidine-5'-triphosphate (5.13a).

General procedure E with 1.43 g (1.1 mmol) of 5.12a, 50 mL of 33% NH_4OH(aq) for 18 h stirring at room temperature. The product 4.13a was afforded as a white solid; yield: 453 mg (83.2%, T, ε_{267}: 9600 $M^{-1}cm^{-1}$); ^1H NMR (400 MHz, D_2O) δ 7.74 (s, 1H), 6.37 (t, 1H, , J = 6.0 Hz), 4.69 (s, 1H), 4.28-4.24 (m, 3H), 2.42 (s, 2H), 1.96 (s, 3H); ^{31}P NMR (162 MHz, D_2O) δ -4.06 (d, J = 16.2 Hz), -9.30 (d, J = 16.2 Hz), -17.95 (brs).

Scheme 5.3: Synthesis of 2'-deoxycytidine-5'-triphosphate (dCTP, 5.13b)



(Scheme was used with permission from . *J. Am. Chem. Soc.* **2019**. *141*, 34, 13286-13289)

3'-O, *N*^{4}-Dibenzoyl-2'-deoxycytidine-5'-dibenzylmonophosphate (5.9b).

General procedure A with 1.04 g (2.4 mmol) of 3'-O, N^4-dibenzoyl-2'-deoxythymidine **5.8b**,[2] 306 mg (4.3 mmol) of tetrazole, 16 mL of anhydrous solution (MeCN/CH_2Cl_2, 1:1), 0.99 mL (3.1 mmol) of dibenzyl-N, N-diisopropylphosphoramidite for 1 hour stirring at room temperature followed by 6 mL of H_2O_2 for 1 hour oxidation reaction. Column chromatography with eluents (MeOH/CH_2Cl_2, from 0% to 12.5%) to afford the product **5.9b** as a white solid; yield: 1.26 g (75.5 %); TLC (MeOH/CH_2Cl_2, 1:60) Rf = 0.32; ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.12 (d, 1H, *J* = 7.6 Hz), 8.02 (d, 2H, *J* = 7.6 Hz), 7.96 (d, 2H, *J* = 7.6 Hz), 7.59-7.52 (m, 2H), 7.48-7.41 (m, 5H), 7.36-7.28 (m, 10H), 6.37 (dd, 1H, *J* = 8.4, 5.6 Hz), 5.41 (d, 1H, *J* = 6.4 Hz), 5.14-5.02 (m, 4H), 4.42-4.32 (m, 3H), 2.82 (dd, 1H, *J* = 14.0, 5.2 Hz), 2.05-1.97 (m, 1H); ^{13}C NMR (125.8 MHz, $CDCl_3$) δ 165.7, 162.4, 154.6, 143.8, 135.4, 135.3, 133.4, 133.1, 132.9, 129.6, 128.9, 128.7, 128.6, 128.6, 128.4, 127.9, 127.9, 127.7 96.9, 87.0, 83.6 (d, $J_{C,P}$ = 7.8 Hz), 75.0, 69.7 (d, $J_{C,P}$ = 5.3 Hz), 66.8 (d, $J_{C,P}$ = 5.2 Hz), 38.7; ^{31}P NMR (162 MHz, $CDCl_3$) δ 0.76; HRMS (ESI-TOF) calcd. for $C_{37}H_{34}N_3O_9PNa$ $[M+Na]^+$ 718.1931; found 718.1949.

3'-O, *N*^{4}-Dibenzoyl-2'-deoxycytidine-5'-monophosphate (5.10b).

General procedure B with 1.26 g (1.8 mmol) of **5.9b**, 100 mL of MeOH, and 250 mg of 10% Pd/C for 3 h stirring. The suspension was filtered over a pad of celite, washed with 100 mL of MeOH containing 2% triethylamine four times to afford the product **5.10a** as a white foam of triethylammonium salt; yield: 1.05 g (84.1%); TLC (MeOH/CH_2Cl_2, 1:10) R_f = 0; ^{1}H NMR (400 MHz, CD_3OD) δ 8.68 (d, 1H, *J* = 6.8 Hz), 8.08 (d, 2H, *J* = 7.6 Hz), 7.98 (d, 2H, *J* = 7.6 Hz), 7.71 (d, 1H, *J* = 7.2 Hz), 7.65-7.62 (m, 2H), 7.55-7.49 (m, 4H), 6.45 (t, 1H, *J* = 6.0 Hz), 5.69 (d, 1H, *J* = 4.0 Hz), 4.51 (s, 1H), 4.22 (s, 1H), 2.78 (dd, 1H, *J* = 13.6, 4.0 Hz), 2.59 -2.52 (m, 1H);

^{13}CNMR (125.8 MHz, CD_3OD) δ 168.9, 167.2, 164.9, 157.9, 146.9, 134.8, 134.5, 134.0, 131.0, 130.6, 129.8, 129.7, 129.1, 99.1, 88.7, 86.5 (d, $J_{C,P}$ = 8.4 Hz), 77.8, 65.8, 40.1; ^{31}P NMR (162 MHz, CD_3OD) δ 1.22; HRMS (ESI-TOF) calcd. for $C_{23}H_{22}N_3O_9PNa$ $[M+Na]^+$ 538.0991; found 538.0984.

3'-O, N4-Dibenzoyl-2'-deoxycytidine-5'-phosphor-2-methylimidazolide (5.11b).

General procedure C with 1.05 g (1.5 mmol) of **5.10b**, 7.3 mL of anhydrous DMF, 1.02 mL (7.3 mmol) of triethylamine, 252 mg (3.0 mmol) of 2-methylimidazole, 918 mg (3.5 mmol) of triphenylphosphine, 804 mg (3.7 mmol) of dipyridyl disulfide for 2 hours reaction at room temperature. First precipitation was achieved with 250 mL of diethyl ether. The product was resuspended with 10 mL of DMF and dropwise added to the solution containing 1.79 g of sodium perchlorate, 15 mL of triethylamine in 300 mL of mixing solution (ethyl acetate/ether 1.5:1) for second precipitation. The product was afforded as a white solid **5.11b**; yield: 0.82 g (93.4%); ^{31}P NMR (162 MHz, D_2O) δ -6.63; HRMS (ESI-TOF) calcd. for $C_{27}H_{25}N_5O_8PNa_2$ $[M-H+2Na]^+$ 624.1236; found 624.1256.

3'-O,*N*^{4}-Dibenzoyl-2'-deoxycytidine-5'-(γ-(2-(pyrenesulfonyl)ethyl))triphosphate (5.12b).

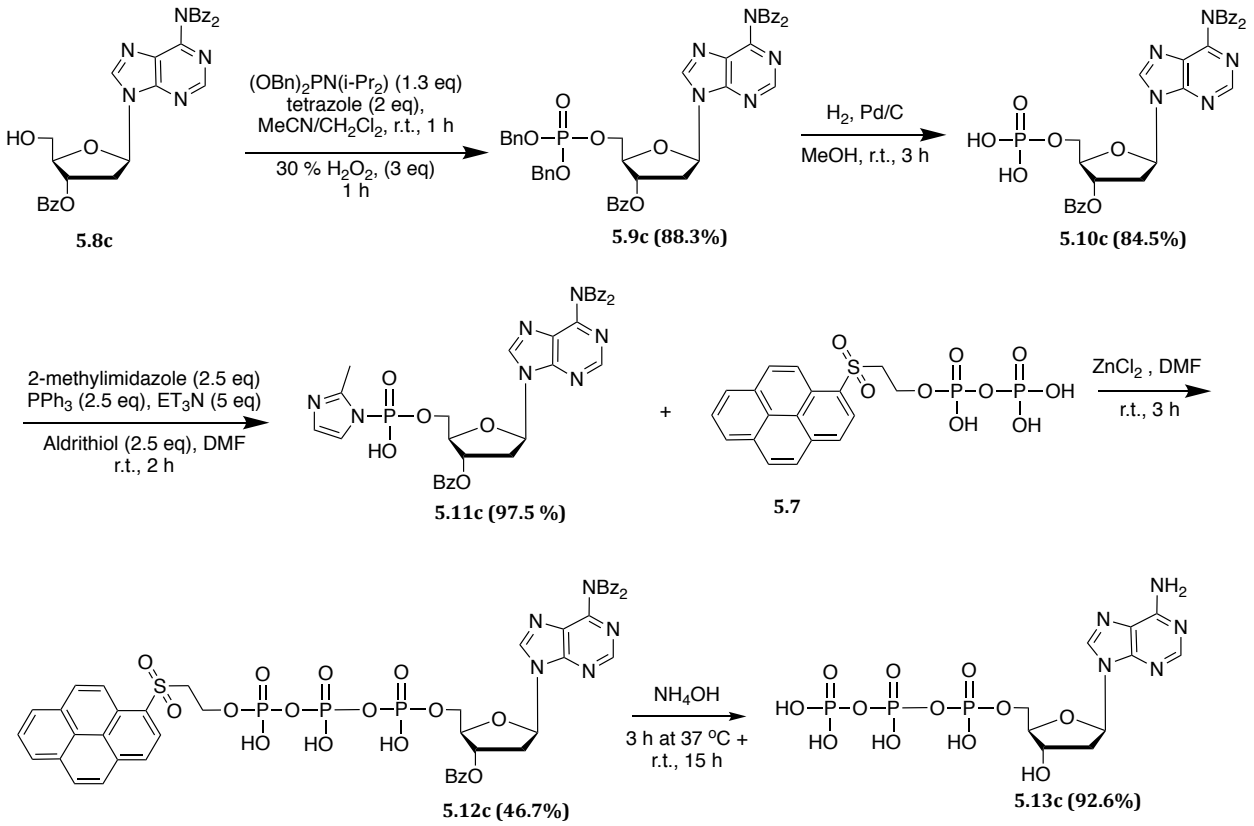
General procedure D with 0.82 g (1.4 mmol) of **5.11b**, 0.77 g (1.6 mmol) of 1-(2-(pyrenesulfonyl)ethyl) pyrophosphate (5.7) and 13.6 mL (13.6 mmol) of $ZnCl_2$ solution (1.0 M in anhydrous DMF) for 3 hours with stirring. Then crude was precipitated by dropwise adding the solution to the 300 mL of ethyl acetate with stirring for precipitation. After centrifugation, the pellet was resuspended by 20% H_2O/MeCN containing 2% Hunig's base

and filtered by pyrex glass funnel. The filtrate was evaporated and crude product was purified by silica column chromatography with eluents 7% H_2O/isopropanol containing 1% diisopropylethylamine (DIPEA) and then H_2O in (isopropanol/MeCN 1:1) from 5% to 8 % containing 1% diisopropylethylamine (DIPEA) to afford the white solid **5.12b**; yield: 1.07 g (58.1 %); TLC (H2O/acetone 1:10 containing 2% DIPEA): R_f = 0.32; ^{1}H NMR (500 MHz, D_2O/DMSO-d^6 1:1) δ 11.15 (s, 1H), 8.91 (s, 1H), 8.60-7.95 (m, 7H), 7.62- 7.33 (m, 4H), 6.25 (s, 1H), 5.47 (s, 1H), 4.33-3.90 (m, 4H), 2.34 (s, 1H); ^{31}P NMR (162 MHz, D_2O/DMSO-d^6 1:1) δ -12.13 (d, 1P, J = 17.2 Hz), -12.38 (d, 1P, J = 14.9 Hz), -22.59 (brs, 1P); HRMS (ESI-TOF) calcd. for $C_{41}H_{34}N_3O_{17}P_3SNa_3$ $[M-2H+3Na]^+$ 1034.0515; found 1034.0485.

2'-deoxythymidine-5'-triphosphate (5.13b).

General procedure E with 1.04 g (0.8 mmol) of **5.12b** in 55 mL of 33% NH_4OH(aq) for 3 hours deprotection at 37 ^{o}C and then 15 hours deprotection at room temperature. The product **5.13b** was afforded as a white solid; yield: 302 mg (82.5%, C, ε_{280}: 13100 $M^{-1}cm^{-1}$); ^{1}H NMR (400 MHz, D_2O) δ 7.96 (d, 1H, J = 7.2 Hz), 6.36 (t, 1H, J = 6.4 Hz), 6.18 (t, 1H, J = 6.8 Hz), 4.64 (s, 1H), 4.26 (s, 3H), 2.47-2.32 (m, 2H); ^{31}P NMR (162 MHz, D_2O) δ 5.04 (brs), -9.23 (brs), -18.71 (brs).

Scheme 5.4: Synthesis of 2'-deoxyadenosine-5'-triphosphate (dATP, 5.13c)



(Scheme was used with permission from . *J. Am. Chem. Soc.* **2019**, *141*, 34, 13286-13289)

3'-O, *N*^{6}, *N*^{6}-Tribenzoyl-2'-deoxyadenosine-5'-dibenzylmonophosphate (5.9c).

General procedure A with 1.03 g (1.8 mmol) of 3'-O, *N*^{6}, *N*^{6}-tribenzoyl-2'-deoxyadenosine **5.8c**,[2] 233 mg (3.3 mmol) of tetrazole, 12 mL of anhydrous solution (MeCN/CH_2Cl_2, 1:1), 0.75 mL (2.4 mmol) of dibenzyl-N, N-diisopropylphosphoramidite for 1 hour stirring at room temperature. Then, the reaction was added to 6 mL of H_2O_2 for 1 hour oxidation reaction. Column chromatography with eluents (EtOAc/hexane, from 33% to 66%) to afford the product **5.9c** as a white foam; yield: 1.33 g (88.3 %); TLC (EtOAc/hexane, 1:60) R_f = 0.32; ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.63 (s, 1H), 8.38 (s, 1H), 8.06 (d, 2H, *J* = 8.0 Hz), 7.87 (d, 3H, *J* = 8.0 Hz), 7.64-7.61 (m, 1H), 7.50-7.46 (m, 4H), 7.37-7.28 (m, 15H), 6.60-6.56 (m, 1H), 5.55 (d, 1H, *J* = 5.6 Hz), 5.07-5.01 (m, 4H), 4.40 (s, 1H), 4.34-4.24 (m, 2H), 2.72-2.67 (m, 2H); ^{13}CNMR

(125.8 MHz, $CDCl_3$) δ 172.7, 166.2, 153.2, 152.6, 152.3, 143.5, 134.5, 134.1, 133.4, 129.9, 129.5, 129.1, 129.1, 129.1, 129.1, 129.1, 129.0, 129.0, 128.6, 128.6, 128.1, 84.9, 83.9 (d, $J_{C,P}$ = 10.4 Hz), 75.0, 69.7 (d, $J_{C,P}$ = 5.3 Hz), 66.8 (d, $J_{C,P}$ = 5.2 Hz), 38.7; ^{31}P NMR (162 MHz, $CDCl_3$) δ 0.70; HRMS (ESI-TOF) calcd. for $C_{45}H_{38}N_5O_9PNa$ $[M+Na]^+$ 846.2305; found 846.2291.

3'-O, *N*^{6}, *N*^{6}-Tribenzoyl-2'-deoxyadenosine-5'-monophosphate (5.10c).

General procedure B with 1.33 g (1.6 mmol) of **5.9c**, 100 mL of MeOH, and 230 mg of 10% Pd/C for 3 hours stirring. The suspension was filtered over a pad of celite, washed with 100 mL of MeOH containing 2% triethylamine four times to afford the product **5.10c** as a white foam of triethylammonium salt; yield: 0.88 g (84.5 %); TLC ($MeOH/CH_2Cl_2$ 1:10 with 1% triethylamine) R_f = 0; 1H NMR (400 MHz, CD_3OD) δ 8.93 (s, 1H), 8.69 (s, 1H), 8.06 (d, 2H, *J* = 7.6 Hz), 7.80-7.78 (m, 3H), 7.62-7.58 (m, 1H), 7.49-7.47 (m, 4H), 7.37- 7.34 (m, 3H), 6.71 (t, 1H, *J* = 1.6 Hz), 5.77 (d, 1H, *J* = 5.6 Hz), 4.47 (s, 1H), 4.17-4.13 (m, 2H), 3.14-3.12 (m, 1H), 2.74 (dd, 1H, *J* = 13.6, 5.6 Hz); ^{13}C NMR (125.8 MHz, CD_3OD) δ 173.7, 167.2, 154.4, 153.1, 152.6, 146.3, 135.4, 134.5, 134.5, 134.2, 131.0, 130.6, 130.4, 129.8, 129.7, 129.6, 129.4, 128.9, 86.0 (d, $J_{C,P}$ = 10.4 Hz), 77.9, 66.2, 55.6, 47.5, 39.1; ^{31}P NMR (162 MHz, CD_3OD) δ 2.30; HRMS (ESI-TOF) calcd. for $C_{31}H_{25}N_5O_9PNa_2$ $[M-H+ 2Na]^+$ 688.1185; found 688.1204.

3'-O, *N*^{6}, *N*^{6}-Tribenzoyl-2'-deoxyadenosine-5'-phosphor-2-methylimidazolide (5.11c).

General procedure C with 0.88 g (1.4 mmol) of **5.10c**, 7.15 mL of anhydrous DMF, 1.0 mL (7.1 mmol) of triethylamine, 239 mg (2.9 mmol) of 2-methylimidazole, 760 mg (2.9 mmol) of triphenylphosphine, 642 mg (2.9 mmol) of dipyridyl disulfide for 2 hours reaction at room temperature. The crude was dropwise added to 300 mL of ether for precipitation and the

solid was collected by centrifugation at 4400 rpm for 10 minutes at room temperature. The crude product was resuspended with minimal volume of CH_2Cl_2 and dropwise added to 300 mL of ether containing 1.4 g of $LiClO_4$. The precipitate was collected by centrifugation at 4400 rpm for 10 minutes at room temperature to afford the white solid **5.11c**; yield: 1.08 g (97.5%); ^{31}P NMR (162 MHz, D_2O) δ -7.59; HRMS (ESI-TOF) calcd. for $C_{35}H_{30}N_7O_8PNa$ $[M+Na]^+$ 730.1791; found 730.1799.

3'-O, *N*^{6}, *N*^{6}-Tribenzoyl-2'-deoxyadenosine -5'-γ-(2-(pyrenesulfonyl)ethyl]-triphosphate (5.12c).

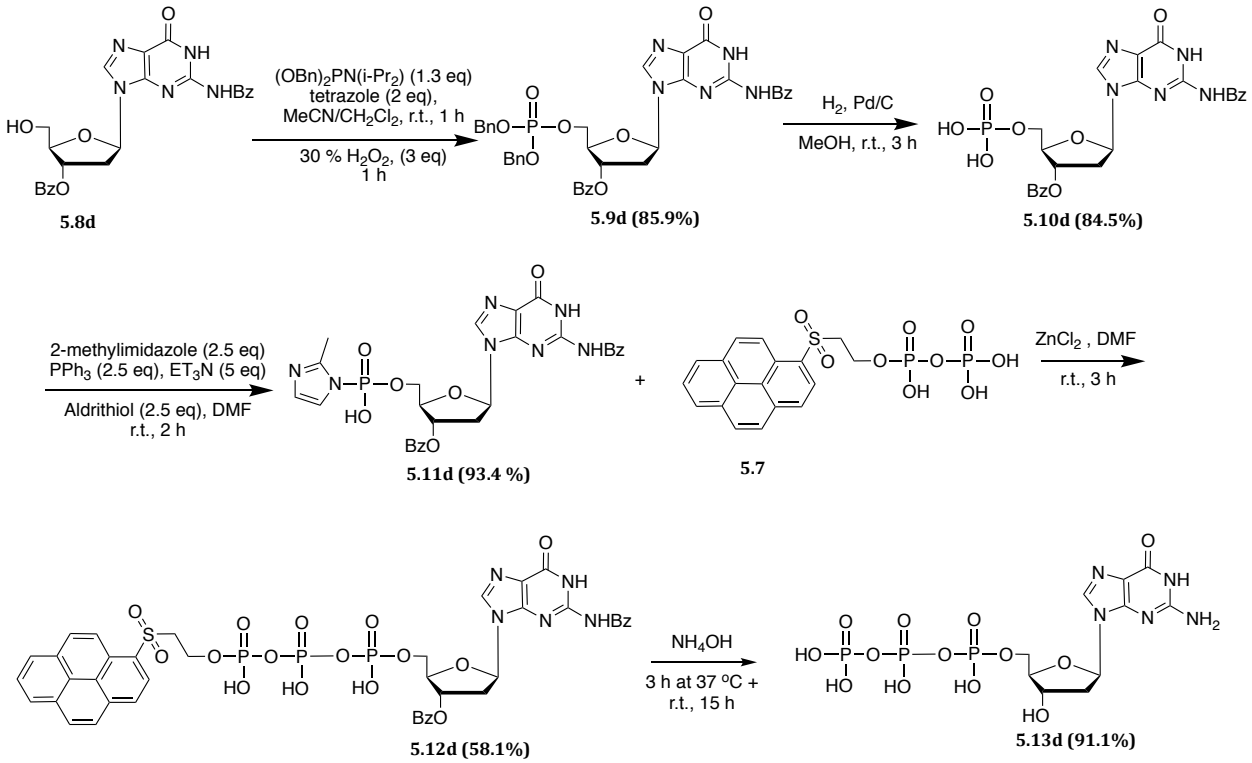
General procedure D with 1.08 g (1.3 mmol) of **5.11c**, 0.77 g (1.6 mmol) of 1-(2-(pyrenesulfonyl)ethyl)-pyrophosphate (5.**7**) and 13.6 mL (13.6 mmol) of $ZnCl_2$ solution (1.0 M in anhydrous DMF) for 3 hours with stirring. After the reaction, the solution was precipitated by 300 mL of ether. After centrifugation, the pellet was resuspended by 20% H_2O/MeCN containing 2% Hunig's base and filtered by pyrex glass funnel. The filtrate was evaporated and the crude product was purified by silica column chromatography with eluents [(8% H_2O/isopropanol+1% diisopropylethylamine), then (H_2O/(isopropanol MeCN 1:1) from 2% to 7% containing 1% diisopropylethylamine (DIPEA)] to afford the white solid **5.12c**; yield: 928 mg (46.7%); TLC (H_2O/acetone 1:10 containing 2% DIPEA): R_f = 0.28; ^1H NMR (600 MHz, D_2O/DMSO-*d*^{6} 1:1) δ 8.58-8.57 (m, 1H), 8.45 (s, 1H), 8.37-8.33 (m, 2H), 8.06-7.70 (m, 11H), 7.57-7.32 (m, 7H), 5.94 (t, 1H, *J* = 7.2 Hz), 5.31 (s, 1H), 4.20 (d, 2H, *J* = 6.6 Hz), 3.99-3.97 (m, 2H), 3.86-3.84 (m, 1H), 3.75-3.74 (m, 1H), 2.55 (s, 1H), 2.10 (dd, 1H, *J* = 13.8, 4.8 Hz); ^{31}P NMR (162 MHz, D^2O/ DMSO-*d*^{6} 1:1) δ -10.24 (d, *J* = 14.9 Hz), -10.45 (d, *J* = 14.7 Hz), -20.67 (brs, 1P); HRMS (ESI-TOF) calcd. For $C_{42}H_{34}N_5O_{16}P_3SNa$ $[M-2H+Na]^+$ 1012.0832;

found 1012.0827. During HRMS, analysis one of the two benzoyl group is removed from the exocyclic amino group.

2'-deoxyadenosine-5'-triphosphate (5.13c).

General procedure E with 928 mg (0.6 mmol) of **5.12c** in 45 mL of 33% NH_4OH(aq) in deprotection step for 3 hours at 37 ^oC and then 15 hours at room temperature. The product **5.13c** was afforded as a white solid; yield: 284 mg (92.6%, A, ε_{259}: 15200 $M^{-1}cm^{-1}$); ^1H NMR (400 MHz, D_2O) δ 8.37 (s, 1H), 8.08 (s, 1H), 6.39 (t, 1H, *J* = 6.4 Hz), 4.33 (s, 1H), 4.26 (s, 2H), 2.81-2.74 (m, 1H), 2.66-2.60 (m, 1H); ^{31}P NMR (162 MHz, D_2O) δ -3.97 (d, J = 15.7 Hz), -9.22 (d, *J* = 16.0 Hz), -17.85 (t, *J* = 13.3 Hz).

Scheme 5.5: Synthesis of 2'-deoxyguanosine-5'-triphosphate (dGTP, 5.13d)



(Scheme was used with permission from . *J. Am. Chem. Soc.* **2019**, *141*, 34, 13286-13289)

3'-O, N^2-Dibenzoyl-2'-deoxyguanosine-5'-dibenzylmonophosphate (5.9d).

General procedure A with 1.0 g (2.1 mmol) of 3'-O, N^2-dibenzoyl-2'-deoxyguanosine. 0.29 g (4.2 mmol) of tetrazole, 15 mL of anhydrous solution ($MeCN/CH_2Cl_2$, 1:1), 0.85 mL (2.5 mmol) of dibenzyl-N, N-diisopropylphosphoramidite for 1 hour stirring at room temperature. Then, the reaction was added to 6 mL of H_2O_2 for 1 hour oxidation reaction. Column chromatography with eluents (EtOAc/hexane, from 33% to 66%) to afford the product **5.9d** as a white foam; yield: 1.35 g (85.9 %); TLC (EtOAc/hexane, 1:60) R_f = 0.32; ^{1}H NMR (400 MHz, $CDCl_3$) δ 12.51 (s, 1H), 11.03 (s, 1H), 8.14 (d, 2H, *J* = 6.0 Hz), 8.06 (d, 2H, *J* = 16.0 Hz), 7.73 (s, 1H), 7.62-7.60 (m, 2H), 7.50-7.45 (m, 4 H), 7.33 (m, 5H), 7.26-7.22 (m, 1H), 7.19-7.16 (m, 1H), 7.00 (d, 2H, *J* = 16.0 Hz), 6.25 (dd, 1H, *J* = 12.0, 2.8 Hz), 5.70 (d, 1H, *J* = 4.4 Hz), 5.08 (d, 2H, *J* = 4.4 Hz), 4.80 (m, 2H), 4.56 (m, 1H), 4.43 (s, 1H), 4.35 (m, 1H), 3.60 (m, 1H), 2.51 (s, 1H); ^{13}C NMR (125.8 MHz, $CDCl_3$) δ 169.2, 166.0, 155.7, 148.2, 148.1, 139.5, 135.4, 135.1, 135.0, 133.8, 133.2, 132.1, 129.8, 129.3, 128.9, 128.8, 128.7, 128.6, 128.5, 128.3, 128.2, 128.2, 127.3, 127.3, 123.5, 86.8, 83.2, 75.8, 69.9 (d, J C,P = 5.3 Hz), 69.4 (d, $J_{C,P}$ = 5.2 Hz), 66.8, 35.5; ^{31}P NMR (162 MHz, $CDCl_3$) δ -0.55; HRMS (ESI-TOF) calcd. for $C_{38}H_{34}N_5O_9PNa$ $[M+Na]^+$ 758.1992; found 758.2000.

3'-O, N^2-Dibenzoyl-2'-deoxyguanosine-5'-monophosphate (5.10d)

General procedure B with 0.88 g (1.2 mmol) of **5.9d**, 10 mL of MeOH, and 250 mg of 10% Pd/C for 3 hours with stirring. The suspension was filtered over a pad of celite, washed with 100 mL of MeOH containing 2% triethylamine four times to afford the product **5.10d** as a

white foam of triethylammonium salt; yield: 0.72 g (84.5 %); TLC ($MeOH/CH_2Cl_2$ 1:10 with 1% triethylamine); ^1H NMR (400 MHz, DMSO-d_6) δ 8.28 (s, 1H), 8.18 (d, 2H, J = 6.4 Hz), 8.03 (m, 2H), 7.70-7.62 (m, 1H), 7.61-7.59 (m, 1H), 7.57-7.54 (m, 2H), 7.51 (m, 2H), 6.44 (m, 1H), 5.71 (d, 1H, J = 3.6 Hz), 4.38 (s, 1H), 4.12 (m, 1H), 4.02 (m, 1H), 3.35 (m, 1H), 2.56 (dd, 1H, J = 15.6, 6.8 Hz). ^{13}C NMR (125.8 MHz, DMSO-d_6) δ 169.8, 165.1, 155.2, 148.8, 139.2, 133.6, 132.8, 132.7, 129.4, 129.3, 128.8, 128.7, 128.2, 128.0, 126.8, 121.1, 84.6, 83.7, 76.8, 63.9, 30.5; ^{31}P NMR (162 MHz, CD_3OD) δ -0.34; HRMS (ESI-TOF) calcd. for $C_{24}H_{21}N_5O_9PNa_2$ $[M-H+2Na]^+$ 600.0872; found 600.0865.

3'-O, N^2-Dibenzoyl-2'-deoxyguanosine-5'-phosphor-2-methylimidazolide (5.11d).

General procedure C with 0.65 g (1.2 mmol) of **5.10d**, 10 mL of anhydrous DMF, 0.82 mL (7.2 mmol) of triethylamine, 240 mg (5.6 mmol) of 2-methylimidazole, 768.3 mg (2.9 mmol) of triphenylphosphine, 645.5 mg (2.93 mmol) of dipyridyl disulfide for 2 hours reaction at room temperature. First precipitation was achieved with 300 mL of diethyl ether. The product was resuspended with 15 mL of CH_2Cl_2 and dropwise added to the solution containing 1.17 g of sodium perchlorate, 15 mL of triethylamine in 300 mL of ethyl acetate for second precipitation. The product was afforded as a white solid **5.11d**; yield: 0.68 g (93.4%); ^{31}P NMR (162 MHz, D_2O) δ -8.70; HRMS (ESI-TOF) calcd. For $C_{28}H_{25}N_7O_8PNa_2$ $[M-H+2Na]^+$ 664.1298; found 664.1276.

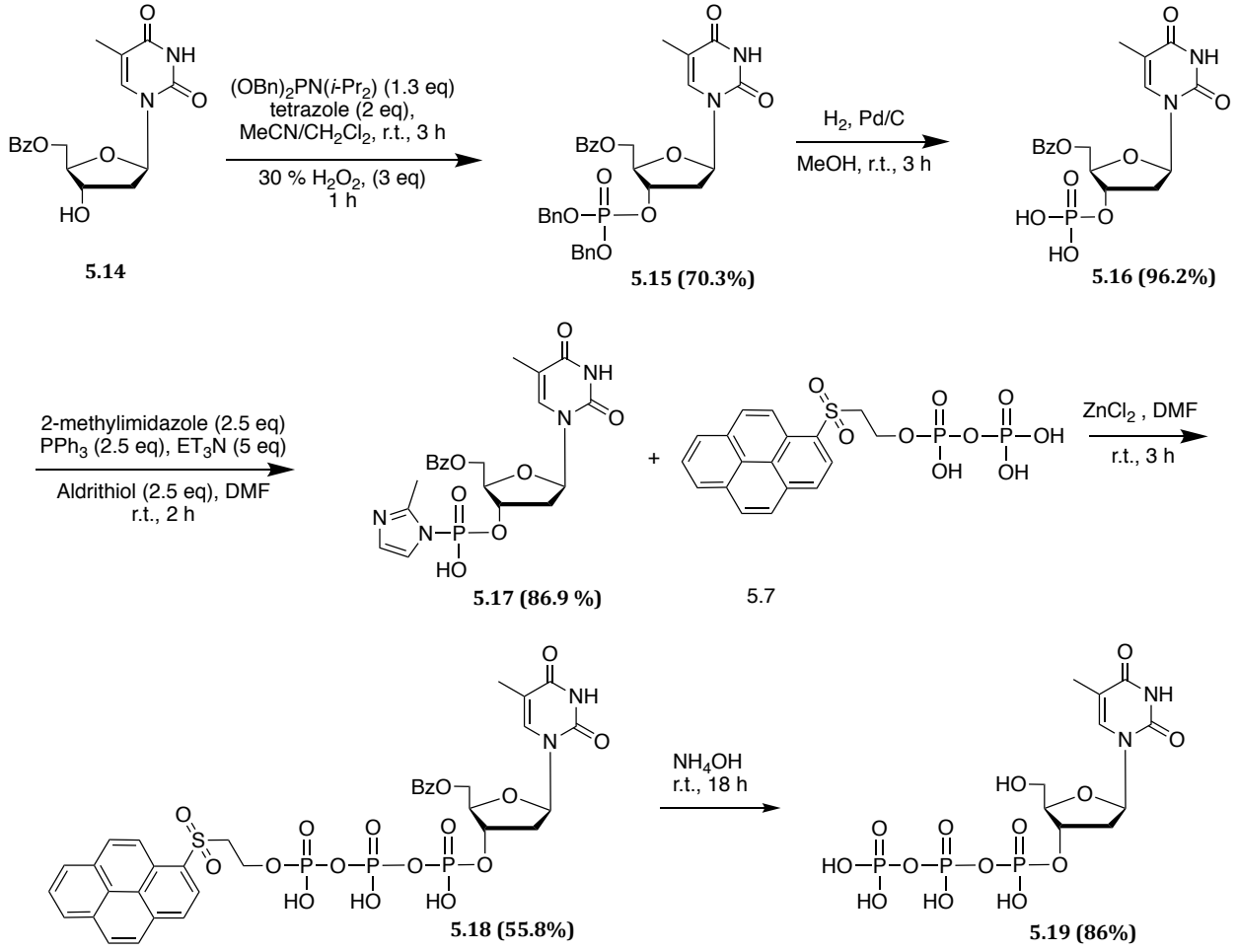
3'-O,-N^2-Dibenzoyl-2'-deoxyadenosine-5'-(γ-(2-(pyrenesulfonyl)ethyl))- triphosphate (5.12d).

General procedure D with 0.5 g (0.81 mmol) of **5.11d**, 0.495 g (1.0 mmol) of 1-(2-(pyrenesulfonyl)ethyl)-pyrophosphate (5.**7**) and 5.38 mL (8.0 mmol) of $ZnCl_2$ solution (1.5 M in anhydrous DMF) for 3 h with stirring. After the reaction, the solution was precipitated by 300 mL of ether. Silica column chromatography with eluents [(H_2O in isopropanol/MeCN) 1:1 from 3% to 8 % containing 1% diisopropylethylamine (DIPEA)] to afford the white solid **5.12d**; yield: 1.07 g (58.1 %); TLC (H_2O/acetonitrile 1:10 containing 2% DIPEA) R_f = 0.34; ^{1}H NMR (400 MHz, DMSO-d_6) δ 7.90-7.60 (m, 1H), 7.31 (s, 6H), 7.03-6.98 (m, 5H), 6.72 (m, 3H), 6.52 (s, 4H), 5.32 (s, 1H), 4.62 (s, 1H), 3.36 (brs, 2H), 3.16 (brs, 2H), 2.87 (s, 1H), 2.09 (s, 2H); ^{31}P NMR (162 MHz, DMSO-d_6) δ -11.34, -12.12, -20.44; HRMS (ESI-TOF) calcd. for $C_{42}H_{34}N_5O_{17}P_3SNa$ $[M-2H+Na]^-$ 1028.0781; found 1028.0796.

2'-deoxyguanosine-5'-triphosphate (5.13d).

General procedure E with 800 mg (0. 81 mmol) of **5.12d** in 45 mL of 33% NH_4OH(aq) for 3 hours at 37 ^{o}C and then 15 hours deprotection reaction at room temperature. The product **5.13d** was afforded as a white solid; yield: 372 mg (91.1 %, G, ε_{253}: 13700 $M^{-1}cm^{-1}$); 1H NMR (400 MHz, D_2O) δ 8.15 (m, 1H), 6.32 (s, 1H), 4.27 (s, 1H), 4.19 (m, 2H), 4.15 (s, 1H), 2.79 (brs, 1H), 2.52 (s, 1H); ^{31}P NMR (162 MHz, D_2O) δ -5.80 (d, *J* = 15.7 Hz), - 10.86 (d, *J* = 16.0 Hz), -20.59 (t, *J* = 13.3 Hz).

Scheme 5.6: Synthesis of 2'-deoxythymidine-3'-triphosphate (3'-TTP, 5.19)



(Scheme was used with permission from . *J. Am. Chem. Soc.* **2019**, *141*, 34, 13286-13289)

5'-Benzoyl-2'-deoxythymidine-3'-dibenzylmonophosphate (5.15).

General procedure A with 1.04 g (3.0 mmol) of 5'-benzoyl-2'-deoxythymidine **5.14**, 383 mg (5.4 mmol) of tetrazole, 30 mL of anhydrous solution ($MeCN/CH_2Cl_2$, 1:1), 1.23 mL (3.91 mmol) of dibenzyl-N, N-diisopropylphosphoramidite for 3 hours reaction at room temperature. Then 8 mL of 30% H_2O_2(aq) for 1 hour reaction at room temperature. Column chromatography with eluents ($MeOH/CH_2Cl_2$, from 1% to 1.4 %) to afford the product **5.15** as a white solid; yield: 1.28 g (70.3 %); TLC ($MeOH/CH_2Cl_2$, 1:40) R_f = 0.22; ^1H NMR (400 MHz, $CDCl_3$) δ 9.46 (s, 1H), 7.97 (d, 2H, *J* = 8.0 Hz), 7.61-7.57 (m, 1H), 7.45 (t, 2H, *J* = 7.2 Hz),

7.35-7.31 (m, 10H), 7.08 (s, 1H), 6.25 (dd, 1H, *J* = 7.6, 6.4 Hz), 5.11-4.97 (m, 5H), 4.56 (dd, 1H, *J* = 13.2, 3.6 Hz), 4.33-4.29 (m, 2H), 2.48 (dd, 1H, *J* = 14.4, 5.6 Hz), 2.11-2.06 (m, 1H), 1.57 (s, 3H); ^{13}C NMR (125.8 MHz, $CDCl_3$) δ 165.7, 163.8, 150.4, 135.3, 135.2, 134.3, 133.5, 129.3, 129.2, 128.7, 128.7, 128.6, 128.6, 128.1, 128.1, 111.4, 84.5, 82.6 (d, $J_{C,P}$ = 5.4 Hz), 69.8, 69.8, 69.8, 63.6, 38.6 (d, $J_{C,P}$ = 4,7 Hz), 12.0; ^{31}P NMR (162 MHz, $CDCl_3$) δ -0.31; HRMS (ESI-TOF) calcd. for $C_{31}H_{31}N_2O_9PNa$ $[M+Na]^+$ 629.1665; found 629.1666.

5'-Benzoyl-2'-deoxythymidine-3'-monophosphate (5.16).

General procedure B with 1.28 g (2.1 mmol) of **5.15**, 50 mL of MeOH, and 200 mg of 10% Pd/C for 3 hours with stirring at room temperature. The suspension was filtered over a pad of celite, washed with 100 mL of MeOH four times to afford the product **5.16** as a white foam; yield: 0.86 g (96.2%); TLC ($MeOH/CH_2Cl_2$, 1:10) R_f = 0; ^{1}H NMR (400 MHz, CD_3OD) δ 8.06 (d, 2H, *J* = 7.2 Hz), 7.64 (t, 1H, *J* = 7.6 Hz), 7.51 (t, 2H, *J* = 7.6 Hz), 7.39 (s, 1H), 6.29 (m, 1H), 5.10 (t, 1H, *J* = 6.8 Hz), 4.70 (dd, 1H, *J* = 12.0, 3.2 Hz), 4.57-4.47 (m, 2H), 2.63 (ddd, 1H, *J* = 14.0, 6.0, 2.4 Hz), 2.43 (m, 1H), 1.62 (s, 3H); ^{13}C NMR (125.8 MHz, CD_3OD) δ 167.4, 166.1, 152.1, 137.0, 134.6, 130.8, 130.5, 129.8, 111.8, 86.4, 84.6, 77.2, 65.1, 39.5, 12.2; 31P NMR (162 MHz, CD_3OD) δ 0.42; HRMS (ESI-TOF) calcd. For $C_{17}H_{18}N_2O_9PNa_2$ $[M-H+2Na]^+$ 471.0545; found 471.0551.

5'-Benzoyl-2'-deoxythymidine-3'-phosphor-2-methylimidazolide (5.17).

General procedure C with 0.86 g (2.0 mmol) of **5.16**, 10 mL of anhydrous DMF, 1.39 mL (10.0 mmol) of triethylamine, 328 mg (4.1 mmol) of 2-methylimidazole, 1.05 g (4.0 mmol) of triphenylphosphine, 903 g (4.0 mmol) of dipyridyl disulfide. First precipitation was

achieved in 250 mL of diethyl ether. The product was resuspended with 10 mL of DMF and dropwise added to the solution containing 2.0 g of sodium perchlorate, 20 mL of triethylamine in 400 mL of ethyl acetate for the second precipitation. The product was afforded as a white solid **5.17**; yield: 0.89 g (86.9%); ^{31}P NMR (162 MHz, D_2O) δ -7.57; HRMS (ESI-TOF) calcd. for $C_{21}H_{23}N_4O_8PNa$ $[M+Na]^+$ 513.1151; found 513.1158.

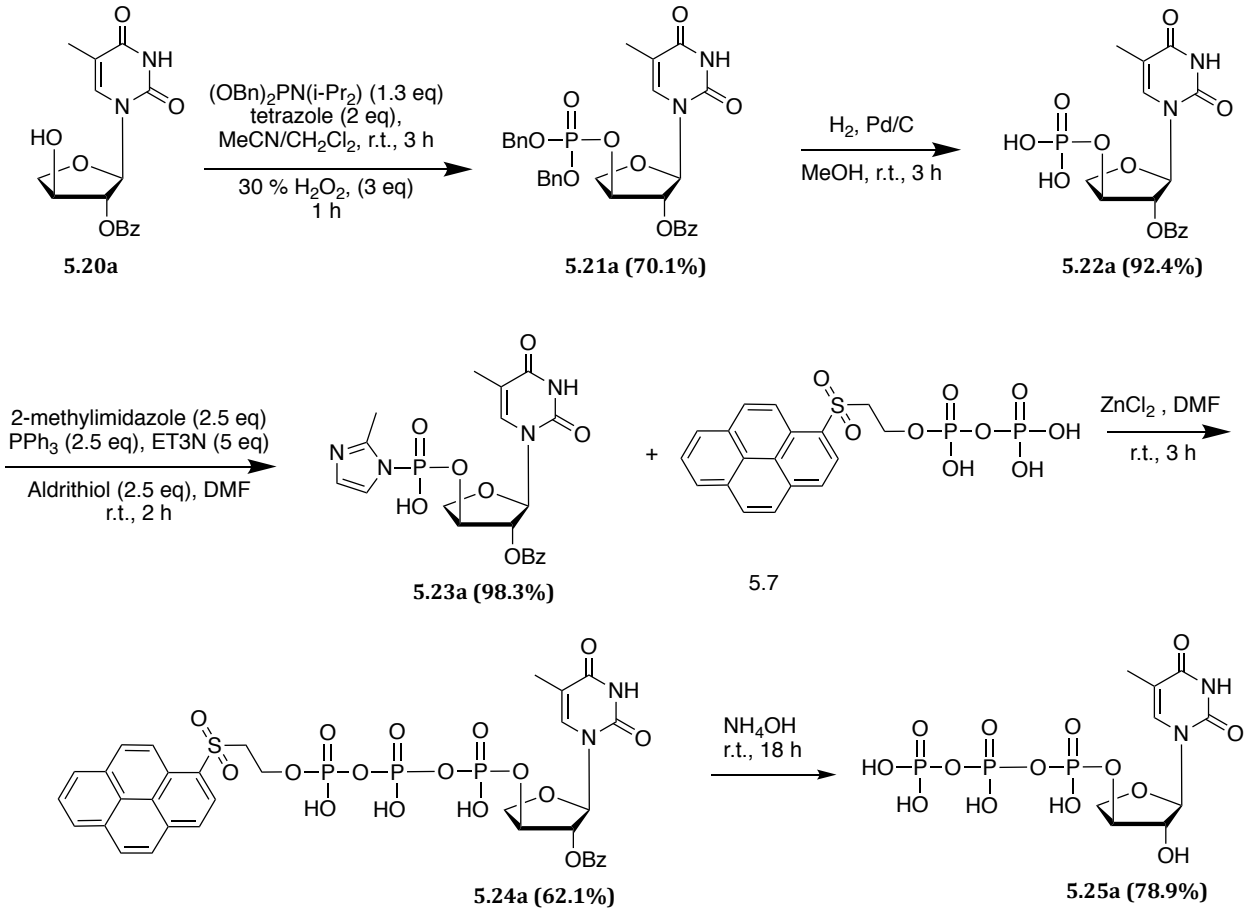
5'-Benzoyl-2'-deoxythymidine-3'-γ-(2-(pyrenesulfonyl)ethyl)triphosphate (5.18). General procedure D with 0.89 g (1.7 mmol) of **5.16**, 0.89 g (1.9 mmol) of 1-(2-(pyrenesulfonyl)ethyl)) pyrophosphate (7) and 11.5 mL (17.30 mmol) of $ZnCl_2$ solution (1.5 M in anhydrous DMF) for 3 hours with stirring at room temperature. The product was purified by 2.5 cm silica gel flash column with eluents [H_2O-isopropanol from 5% to 10 % containing 1% diisopropylethylamine (DIPEA)] to afford the yellowish solid **5.18**; yield: 1.23 g (55.8%); TLC (1:10 H_2O-acetone/2%-DIPEA): R_f =30.8; ^{1}H NMR (400 MHz, D_2O) δ 8.25 (d, 1H, *J* = 8.8 Hz), 8.11 (d, 1H, *J* = 6.4 Hz), 7.47 (d, 1H, *J* = 8.0 Hz), 7.36-7.34 (m, 4H), 7.15-7.07 (m, 2H), 6.94-6.79 (m, 5H), 6.26 (s, 1H), 5.12 (s, 1H), 4.85 (s, 1H), 4.48-4.25 (m, 5H), 3.80 (s, 2H), 2.35 (s, 1H), 1.84 (s, 1H), 0.66 (s, 3H); ^{31}P NMR (162 MHz, D_2O) δ -10.56 (d, 1P, *J* = 12.5 Hz), -11.07 (d, 1P, *J* = 14,1 Hz), -20.85 (brs, 1P); HRMS (ESI-TOF) calcd. for $C_{35}H_{30}N_2O_{17}P_3SNa_4$ $[M-3H+4Na]^+$ 967.0069; found 967.0080.

2'-deoxythymidine-3'-triphosphate (5.19).

General procedure E with 1.23 g of **5.18**, 5mL of CH_2Cl_2 and 50 mL of 33% NH_4OH(aq) for 18 hours deprotection at room temperature. The product **5.19** was afforded as a white solid; yield: 404 mg (86%, ε_{267} = 9600 $M^{-1}cm^{-1}$); ^{1}H NMR (400 MHz, D_2O) δ 7.74 (s, 1H), 6.40

(t, 1H, J = 7.6 Hz), 5.05 (s, 1H), 4.34 (d, 1H, *J* = 2.8 Hz), 3.92 (d, 2H, *J* = 3.6 Hz), 2.70- 2.65 (m, 1H), 2.54-2.47 (m, 1H), 1.96 (s, 3H); ^{31}P NMR (162 MHz, D_2O) δ -4.03 (d, 1P, *J* = 17.0 Hz), -9.99 (d, 1P, *J* = 16.7 Hz), -18.47 (brs, 1P).

Scheme 5.7: Synthesis of 1-(α-L-threofuranosyl)thymidine-3'-triphosphate (tTTP, 5.25a)



(Scheme was used with permission from . *J. Am. Chem. Soc.* **2019**, *141*, 34, 13286-13289)

1-(2'-O-benzoyl-α-L-threofuranosyl)thymidine-3'-dibenzylmonophosphate (5.21a).

General procedure A with 1g (3.0 mmol) of 1-(2'-O-benzoyl-α-L-threofuranosyl) thymine,4 379.7 mg (5.4 mmol) of tetrazole, 24 mL of anhydrous solution ($MeCN/CH_2Cl_2$, 1:1), 1.24 mL (3.9 mmol) of dibenzyl-*N, N*-diisopropylphosphoramidite for 3 hours stirring at room temperature and 6 mL of H_2O_2 for 1 hour oxidation reaction. Silica gel column chromatography with eluents ($MeOH/CH_2Cl_2$, from 0 to 2%). The product **5.21a** was acquired as a white solid; yield: 1.25 g (70.1 %). This is a known compound previously reported in the literature.[5]

1-(2'-O-benzoyl-α-L-threofuranosyl)thymidine-3'-monophosphate (5.22a)

General procedure B with 1.25 g (2.1 mmol) of **5.21a**, 50 mL of MeOH, and 200 mg of 10% Pd/C for 3 hours stirring. The suspension was filtered over a pad of celite, washed with 100 mL of MeOH four times. The product **5.22a** was afforded as a white foam; yield: 0.83 g (92.4%); TLC ($MeOH/CH_2Cl_2$, 1:10) R_f = 0; ^{1}H NMR (400 MHz, CD_3OD) δ 8.05 (d, 1H, *J* = 7.6 Hz), 7.64 (t, 1H, *J* = 7.2 Hz), 7.57 (s, 1H), 7.50 (t, 2H, *J* = 7.2 Hz), 6.05 (s, 1H), 5.60 (s, 1H), 5.01 (s, 1H), 4.52 (d, 1H, *J* = 10.4 Hz), 4.29 (d, 1H, *J* = 9.2 Hz), 1.92 (s, 3H); ^{13}C NMR (125.8 MHz, CD_3OD) δ 166.3, 152.2, 138.0, 134.9, 130.9, 130.1, 129.7, 111.7, 91.2, 82.1 (d, $J_{C,P}$ = 5.9 Hz), 79.0, 74.7, 12.5; ^{31}P NMR (162 MHz, $CDCl_3$) δ -0.23; HRMS (ESI-TOF) calcd. for $C_{16}H_{17}N_2O_9PNa$ $[M+Na]^+$ 435.0569; found 435.0557.

1-(2'-O-benzoyl-α-L-threofuranosyl)thymidine-3'-monophosphor-2-methylimidazolide (5.23a)

General procedure C with 0.83 g (2 mmol) of **5.22a**, 7.0 mL of anhydrous DMF, 1.40 mL (10.1 mmol) of triethylamine, 328 mg (4 mmol) of 2-methylimidazole, 1.05 g (4 mmol) of

triphenylphosphine, 0.88 g (4 mmol) of dipyridyl disulfide. First precipitation was achieved with 250 mL of diethyl ether. The product was resuspended with 15 mL of CH_2Cl_2 and dropwise added to the solution containing 2.01 g of sodium perchlorate, 10 mL of triethylamine in 300 mL of ethyl acetate for the second precipitation. The product **5.23a** was afforded as a white solid; yield: 0.98 g (98.3%); ^{31}P NMR (162 MHz, D_2O) δ -8.17; HRMS (ESI-TOF) calcd. for $C_{20}H_{21}N_4O_8PNa$ $[M+Na]^+$ 499.0995; found 499.0992.

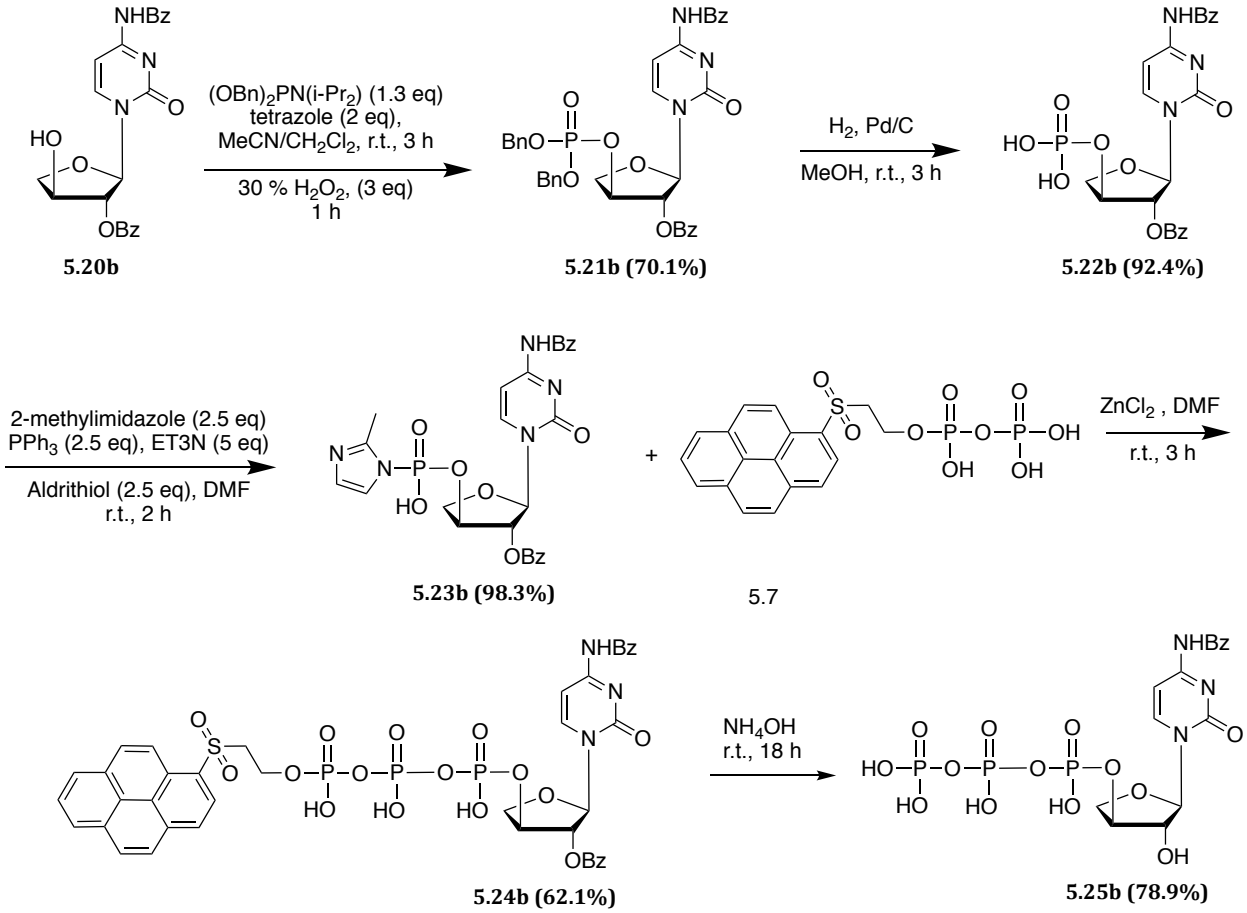
1-(2'-O-benzoyl-α-L-threofuranosyl)thymidine-3'-(γ-(2-(pyrenesulfonyl)ethyl)) triphosphate (5.24a)

General procedure D with 0.98 g (1.9 mmol) of **5.23a**, 1.11 g (2.4 mmol) of 7 and 13.13 mL (19.7 mmol, 1.5 M in anhydrous DMF) of $ZnCl_2$ solution for 3 hours stirring at room temperature. The product was purified by 2 cm silica gel column chromatography with eluents [H_2O-isopropanol from 5% to 12.5 % containing 1% diisopropylethylamine (DIPEA)] to afford the white solid **5.24a**; yield: 1.53 g (62.1%); TLC (1:10 H_2O-acetone with 2%-DIPEA): R_f = 0.25; ^{1}H NMR (400 MHz, D_2O) δ 8.22 (d, 1H, *J* = 9.2 Hz), 8.09 (d, 1H, *J* = 8.0 Hz), 7.48-7.30 (m, 7H), 7.09-6.98 (m, 4H), 6.84 (t, 2H, *J* = 6.8 Hz), 5.32 (s, 1H), 5.20 (s, 1H), 4.89 (s, 1H), 4.48-4.40 (m, 3H), 3.97 (d, 1H, *J* = 8.0 Hz), 3.76 (m, 2H), 1.50 (s, 3H); ^{31}P NMR (162 MHz, D_2O) δ -10.23 (d, *J* = 13.0 Hz), -11.65 (d, *J* = 13.9 Hz), - 20.59 (brs); HRMS (ESI-TOF) calcd. for $C_{34}H_{29}N_2O_{17}P_3SNa_3$ $[M-2H+3Na]^+$ 931.0093; found 931.0049.

1-(α-L-threofuranosyl)thymidine-3'-triphosphate (5.25a)

General procedure E with 1.53 g (1.2 mmol) of **5.24a** in 50 mL of 33% NH_4OH(aq) for 18 hours deprotection at room temperature. The product **5.25a** was afforded as a white solid; yield: 445 mg (78.9%, ε_{267} = 9600 $M^{-1}cm^{-1}$); ^1H NMR (400 MHz, D_2O) δ 7.62 (s, 1H), 5.85 (s, 1H), 4.90 (s, 1H), 4.61 (s, 1H), 4.51 (d, 1H, *J* = 10.4 Hz), 4.40 (d, 1H, *J* = 9.2 Hz), 1.95 (s, 1H); ^{31}P NMR (162 MHz, D_2O) δ -4.89 (d, *J* = 14.7 Hz), -11.02 (d, *J* = 18.63 Hz), - 19.80 (brs).

Scheme 5.8: Synthesis of 1-(α-L-threofuranosyl)cytidine-3'-triphosphate (tCTP, 5.25b)



(Scheme was used with permission from . *J. Am. Chem. Soc.* **2019**, *141*, 34, 13286-13289)

N^4-Benzoyl-1-(2'-O-benzoyl-α-L-threofuranosyl)cytidine-3'-dibenzylmonophosphate (5.21b)

Modified general procedure A[5] with 1g (2.4 mmol) of N^4-benzoyl-1-(2'-O-benzoyl-α-L-threofuranosyl) cytosine **(5.20b),**[4] 300 mg (4.3 mmol) of tetrazole, 15.8 mL of anhydrous solution ($MeCN/CH_2Cl_2$, 1:1), 0.98 mL (3.1 mmol) of dibenzyl-*N, N*-diisopropylphosphoramidite for 1 hour reaction at room temperature and 5 mL of H_2O_2 for 1 hour oxidation reaction. After the reaction, the product **5.21b** was afforded as a white solid; yield: 1.29 g (79.3 %). Additional compound characterization.[5]

N^4-Benzoyl-1-(2'-O-benzoyl-α-L-threofuranosyl)cytidine-3'-monophosphate (5.22b).

General procedure B with 1.29 g (1.9 mmol) of **5.21b**, 50 mL of MeOH, and 250 mg of 10% Pd/C for 5 hours stirring. The suspension was filtered over a pad of celite, washed with 100 mL of MeOH containing 2% triethylamine four times. The product **5.22b** was afforded as a white foam of a triethylammonium salt; yield 1.23 g (92.8%);^{1}H NMR (400 MHz, $CDCl_3$) δ 8.26 (d, 1H, *J* = 7.6 Hz), 7.99 (d, 2H, *J* = 7.2 Hz), 7.90 (d, 2H, *J* = 7.6 Hz), 7.57-7.53 (m, 2H), 7.47-7.39 (m, 5H), 6.18 (s, 1H), 5.66 (s, 1H), 4.91-4.85 (m, 2H), 4.28 (dd, 1H, *J* = 10.4, 3.2 Hz); ^{13}C NMR (125.8 MHz,$CDCl_3$) δ 164.6, 162.7, 154.9, 146.2, 133.5, 132.9, 129.9, 129.3, 128.9, 128.9, 128.5, 128.0, 128.0, 96.6, 91.4, 80.9 (d, $J_{C,P}$ = 6.5 Hz), 76.4 (d, $J_{C,P}$ = 4.7 Hz), 57.9; ^{31}P NMR (162 MHz, $CDCl_3$) δ -1.15; HRMS (ESITOF) calcd. for $C_{20}H_{20}N_3O_9PNa_3$ $[M-2H+3Na]^+$ 546.0630; found 546.0640.

N^4-Benzoyl-1-(2'-O-benzoyl-α-L-threofuranosyl)cytidine-3'-monophosphor-2-methylimidazolide (5.23b)

General procedure C with 1.02 g (1.8 mmol) of **5.22b**, 4.3 mL of anhydrous DMF, 0.87 mL (6.4 mmol) of triethylamine, 287 mg (3.5 mmol) of 2-methylimidazole, 0.94 g (3.6 mmol) of triphenylphosphine, 0.79 g (3.6 mmol) of dipyridyl disulfide for 2 hours stirring at room temperature. First precipitation was achieved with 200 mL of diethyl ether. The product was resuspended with 15 mL of CH_2Cl_2 and dropwise added to the solution containing 2.04 g of sodium perchlorate, 10 mL of triethylamine in 300 mL of ethyl acetate for the second precipitation. The product **5.23b** was afforded as a white solid; yield: 0.94 g (91.5%); ^{31}P NMR (162 MHz, D_2O) δ -7.56; HRMS (ESI-TOF) calcd. for $C_{26}H_{24}N_5O_8PNa$ $[M+Na]^+$ 588.1260; found 588.1246.

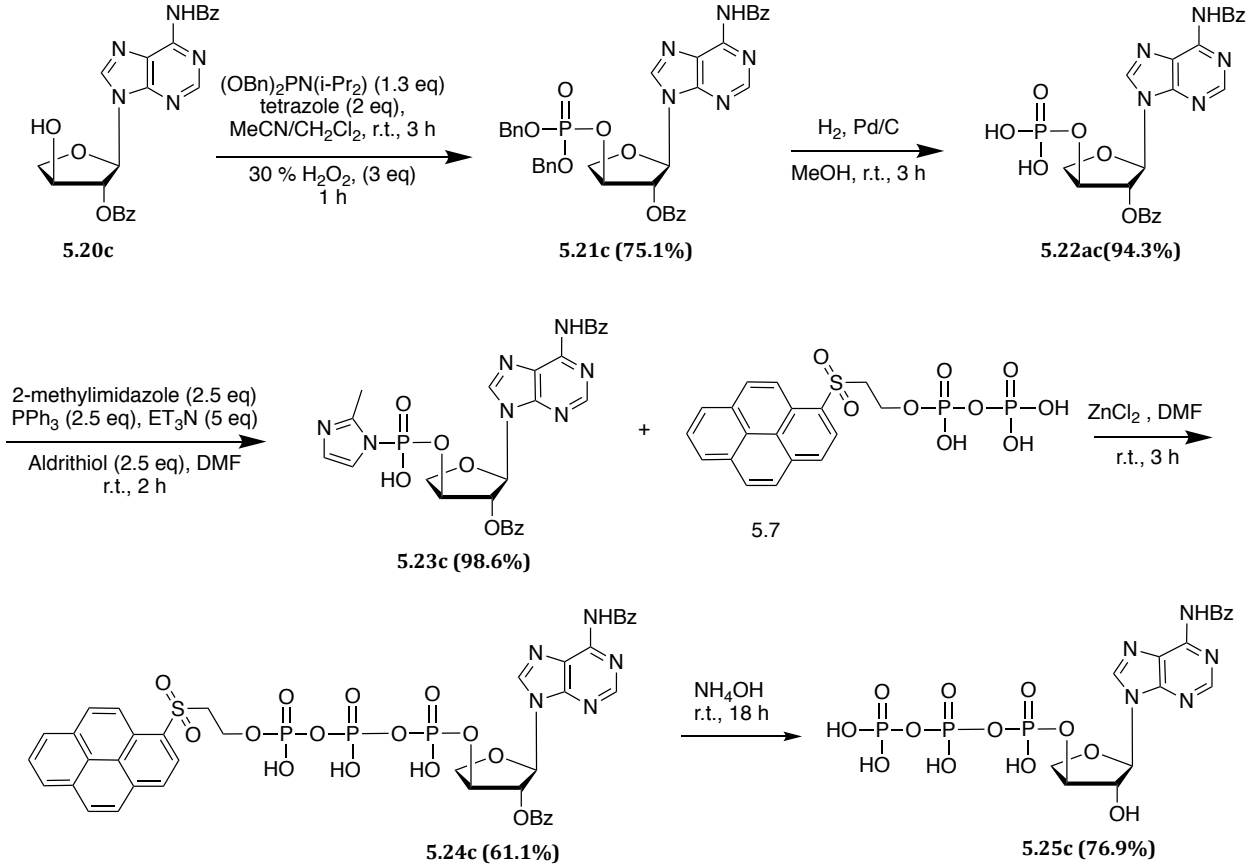
***N*⁴-Benzoyl-1-(2'-O-benzoyl-α-L-threofuranosyl)cytidine-3'-(γ-(2-(pyrenesulfonyl)ethyl) triphosphate (5.24b)**

General procedure D with 0.94 g (1.6 mmol) of **5.23b**, 0.9 g (1.9 mmol) of 7 and 10.67 mL (16 mmol, 1.5 M in anhydrous DMF) of $ZnCl_2$ solution for 5 hours stirring at room temperature. The product was purified by silica gel chromatography with eluents [(6% H_2O/isopropanol containing 1% DIPEA) and then (5% to 10% of H2O/(isopropanol-MeCN 1:1)) containing 1% diisopropylethylamine (DIPEA) to afford the white solid **5.24b**; yield: 1.43 g (66.7%); TLC (1:10 H_2O-acetone with 2%-DIPEA) R_f = 0.25; 1H NMR (400 MHz, D_2O) δ 8.18 (s, 1H), 8.04 (s, 1H), 7.63-7.25 (m, 13H), 7.14 (s, 3H), 6.95-6.91 (m, 4H), 6.64 (s, 1H), 5.57 (s, 1H), 5.41 (s, 1H), 4.94 (s, 1H), 4.63 (s, 1H), 4.44 (s, 2H), 4.23 (s, 1H), 3.74-3.63 (m, 2H); ^{31}P NMR (162 MHz, D_2O) δ -10.23 (d, *J* = 13.0 Hz), -11.65 (d, *J* = 13.9 Hz), -20.59 (brs); HRMS (ESI-TOF) calcd. for $C_{40}H_{32}N_3O_{17}P_3SNa$ $[M+Na]^+$ 974.0563; found 974.0620.

1-(α-L-threofuranosyl)cytidine-3'-triphosphate (5.25b)

General procedure E with 1.43 g (1.0 mmol) of **5.24b** in 50 mL of 33% NH_4OH(aq) for 3 hours at 37 ^oC and then 15 hours deprotection reaction at room temperature. The product **5.25b** was afforded as a white solid; yield: 412 mg (85.8%, ε_{280} = 13100 $M^{-1}cm^{-1}$); ^1H NMR (400 MHz, D_2O) δ 7.82 (d, 1H, *J* = 7.6 Hz), 6.09 (d, 1H, *J* = 7.6 Hz), 5.90 (s, 1H), 4.91 (d, 1H, *J* = 4.8 Hz), 4.58-4.53 (m, 2H), 4.42 (d, 1H, *J* = 6.8 Hz); ^{31}P NMR (162 MHz, D_2O) δ -4.26 (brs), -10.93 (d, J = 17.3 Hz), -19.15 (brs).

Scheme 5.9: Synthesis of 9-(α-L-threofuranosyl)adenosine-3'-triphosphate (tATP, 5.25c)



(Scheme was used with permission from . *J. Am. Chem. Soc.* **2019**, *141*, 34, 13286-13289)

N^6-Benzoyl-9-(2'-O-benzoyl-α-L-threofuranosyl)adenosine-3'-dibenzyl monophosphate (5.21c)

General procedure A with 1.0 g (2.3 mmol) of N^6-benzoyl-9-(2'-O-benzoyl-α-L-threofuranosyl) adenine **(5.20c),** 4 3 mg (4.5 mmol) of tetrazole, 15 mL of anhydrous solution ($MeCN/CH_2Cl_2$, 1:1), 0.91 mL (2.7 mmol) of dibenzyl-*N, N*-diisopropylphosphoramidite for 3 hours reaction at room temperature and 5 mL of H_2O_2 for 1 hour oxidation reaction. The product **5.21c** afford as a white solid; yield: 1.41 g (88.4 %). Additional compound characterization.[5]

N^6-Benzoyl-9-(2'-O-benzoyl-α-L-threofuranosyl)adenosine-3'-monophosphate (5.22c)

General procedure B with 1.3 g (1.8 mmol) of **5.21c**, 10 mL of MeOH, and 400 mg of 10% Pd/C for 12 hours with stirring. The suspension was filtered over a pad of celite, washed with 100 mL of MeOH four times. The product **5.22c** was afford as a white solid; yield: 943 mg (97.4%);^{1}H NMR (400 MHz, CD_3OD): δ 8.98 (s, 1H), 8.91 (s, 1H), 8.26-8.17 (m, 4H), 7.83-7.80 (m, 2H), 7.74-7.65 (m, 4 H), 6.68 (s, 1H), 6.18 (s, 1H), 5.33 (s, 1H), 4.88-4.85 (m, 1H), 4.66 (dd, 1H, *J* = 14, 6.4 Hz). ^{13}C NMR (125.8 MHz, DMSO-d^6): δ 164.7, 151.9, 150.4, 142.8, 133.9, 133.3, 132.4, 129.6, 129.0, 128.7, 128.4, 128.4, 125.3, 109.4, 87.0, 80.3, 76.2, 72.6; 31P NMR (162 MHz, DMSO-d^6) δ: 0.30; HRMS (ESI-TOF) calcd. For $C_{23}H_{21}N_5O_8P$ $[M + H]^+$ 526.1128; found 526.1118.

N^6-Benzoyl-9-(2'-O-benzoyl-α-L-threofuranosyl)adenosine-3'-monophosphor-2-methylimidazolide (5.23c)

General procedure C with 0.93 g (1.6 mmol) of **5.22c**, 12 mL of anhydrous DMF, 1.1 mL (7.9 mmol) of triethylamine, 284 mg (3.9 mmol) of 2-methylimidazole, 1.03 g (3.9 mmol) of triphenylphosphine, 0.87 g (3.9 mmol) of dipyridyl disulfide. First precipitation was achieved with 150 mL of diethyl ether. The product was resuspended with 5 mL of CH_2Cl_2 and dropwise added to the solution containing 1.5 g of sodium perchlorate, 6 mL of triethylamine in 150 mL of ethyl acetate for the second precipitation. The product afford as a white solid **5.23c**; yield: 0.89 g (93.0 %); ^{31}P NMR (162 MHz, DMSO-d_6) δ: -9.80; HRMS (ESI-TOF) calcd. for $C_{27}H_{24}N_7O_7PNa$ $[M + Na]^+$ 612.1373; found 612.1345.

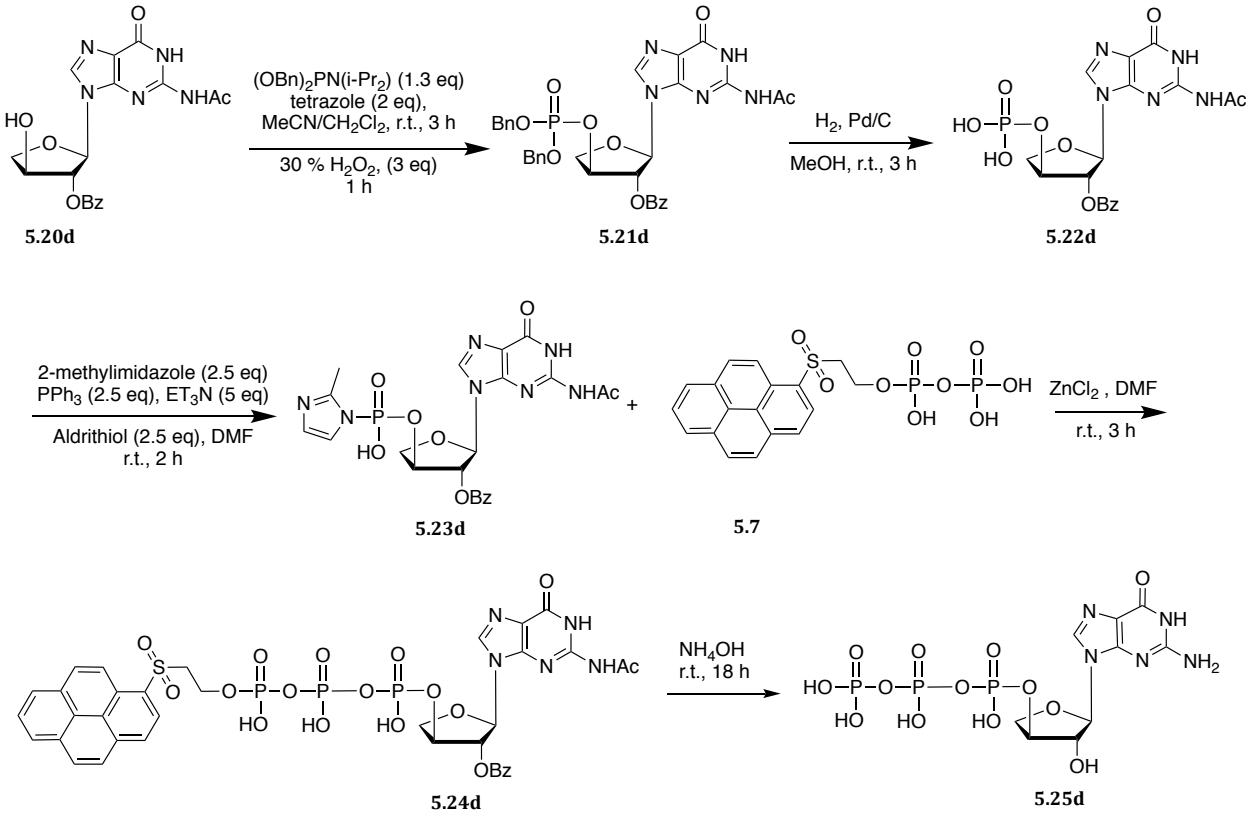
N^6-Benzoyl-9-(2'-O-benzoyl-α-L-threofuranosyl)adenosine-3'-γ-[2-(pyrenesulphonyl)ethyl] triphosphate (5.24c)

General procedure D with 0.6 g (1.02 mmol) of **5.23c**, 0.52 g (1.1mmol) of 7 and 7 mL (10.1 mmol, 1.5 M in anhydrous DMF) of $ZnCl_2$ solution for 3 hours stirring at room temperature. The product was purified by silica gel chromatography with eluents [MeOH-$CHCl_3$ from 5% to 12 % containing 1% diisopropylethylamine (DIPEA)] to afford the white solid **5.24c**; yield 0.64 g (63.1 %); TLC (1:10 H_2O-acetone with 2%-DIPEA) R_f = 0.31; ^{1}H NMR (400 MHz, D_2O) δ 9.50 (s, 1H), 9.18-7.56 (m, 7H), 7.51-7.46 (m, 6H), 7.35-7.09 (m, 9H), 6.19 (m, 1H), 5.72 (m, 1H), 4.72-4.52 (m, 1H), 3.48 (s, 2H), 2.99 (d, 2H, J=12 Hz); ^{31}P NMR (162 MHz, D_2O) δ -9.21 (d, *J* = 12.9 Hz), -10.51 (d, *J* = 13.0 Hz), -19.24 (brs); HRMS (ESI-TOF) calcd. for $C_{41}H_{32}N_5O_{16}P_3SNa$ $[M-2H+Na]^-$ 998.0676; found 998.0667.

9-(α-L-threofuranosyl)adenosine-3'-triphosphate (5.25c)

General procedure E with 0.5 g (0.5 mmol) of **5.24c**, 35 mL of 33% NH_4OH(aq) for 3 hours deprotection at 37 ^oC, and then 15 hours deprotection reaction at room temperature. The product was afforded as a white solid; yield: 0.21 g (86.2 %, ε_{280} = 15200 $M^{-1}cm^{-1}$); ^1H NMR (400 MHz, D_2O) δ 8.36 (s, 1H), 8.27 (s, 1H), 6.16 (s, 1 H), 5.09 (d, 1 H, J = 8.0 Hz), 4.96 (s, 1H), 4.50 (s, 2H); ^{31}P NMR (162 MHz, D_2O) δ -4.02 (d, *J* = 16.6 Hz), -10.51 (d, *J* = 16.2 Hz) , -17.77 (brs).

Scheme 5.10: Synthesis of 9-(α-L-threofuranosyl)guanosine-3'-triphosphate (tGTP, 5.25d)



(Scheme was used with permission from . *J. Am. Chem. Soc.* **2019**, *141*, 34, 13286-13289)

N^2-Acetyl-9-(2'-O-benzoyl-α-L-threofuranosyl)guanosine-3'-dibenzylmonophosphate (5.21d)

General procedure A with 1 g (1.7 mmol) of N^2-acetyl-9-(2'-O-benzoyl-α-L-threofuranosyl) guanine **(5.20d)**,[4] 212 mg (3.0 mmol) of tetrazole, 12 mL of anhydrous solution ($MeCN/CH_2Cl_2$, 1:1), 735 mL (2.2 mmol) of dibenzyl-N, N-diisopropylphosphoramidite for 3 hours reaction at room temperature and 6 mL of H_2O_2 for 1 hour oxidation. Column chromatography with eluents (EtOAc/Hexane, 50%; then $EtOAc/CH_2Cl_2$, from 16% to 25%) to afford the product **5.21d** as a white solid; yield: 1.14 g (79.4 %). TLC ($EtOAc/CH_2Cl_2$, 1:4) R_f = 0.44; This is a known compound previously reported in the literature.[5]

N^2-Acetyl-9-(2'-O-benzoyl-α-L-threofuranosyl)guanosine-3'-monophosphate (5.22d)

General procedure B with 1.63 g (2.5 mmol) of **5.21d**, 15 mL of MeOH, and 600 mg of 10% Pd/C for 3 hours stirring. The suspension was filtered over a pad of celite, washed with 100 mL of MeOH four times. The product **5.22d** was afford as a white solid; yield; 1.04 g (87.7 %).^{1}H NMR (400 MHz, DMSO-d^6): δ 12.1 (s, 1H), 11.68 (s, 1H), 8.13 (s, 1H), 8.04- 8.02 (m, 2H), 7.74-7.70 (m, 1H), 7.59-7.55 (m, 2H), 6.11 (d, 1H, J = 4 Hz), 5.86 (s, 1H), 5.09 (d, 1H, J = 2.4 Hz) 4.47-4.45 (m, 1H), 4.38-4.34 (m, 1H), 2.15 (s, 3H); ^{13}C NMR (125.8 MHz, DMSO-d^6): δ 173.5, 164.5, 154.8, 148.4, 148.1, 137.7, 134.1, 129.6, 128.9, 128.3, 119.9, 86.9, 80.6, 76.5, 72.8, 48.6, 23.8; ^{31}P NMR (162 MHz, DMSO-d^6): δ -0.89; HRMS (ESI-TOF) calcd. for $C_{18}H_{18}N_5O_9PNa$ $[M + Na]^+$ 502.0740; found 502.0727.

N^2-Acetyl-9-(2'-O-benzoyl-α-L-threofuranosyl)guanosine-3'-monophosphor-2-

methylimidazolide (5.23d)

General procedure C with 0.94 g (1.9 mmol) of **5.22d**, 12 mL of anhydrous DMF, 1.3 mL (9.9 mmol) of triethylamine, 355 mg (4.9 mmol) of 2-methylimidazole, 1.29 g (4.9 mmol) of triphenylphosphine, 1.086 g (4.9 mmol) of dipyridyl disulfide. First precipitation was achieved with 150 mL of diethyl ether. The product was resuspended with 5 mL of CH_2Cl_2 and dropwise added to the solution containing 1.5 g of sodium perchlorate, 6 mL of trimethylamine in 150 mL of ethyl acetate for the second precipitation. The product afford as a white solid **5.23d**; yield: 0.88 g (82.1 %); ^{31}P NMR (162 MHz, DMSO-*d*^{6}): δ -9.77; HRMS (ESI-TOF) calcd. for $C_{22}H_{22}N_7O_8PNa$ $[M + Na]^+$ 566.1165; found 566.1163.

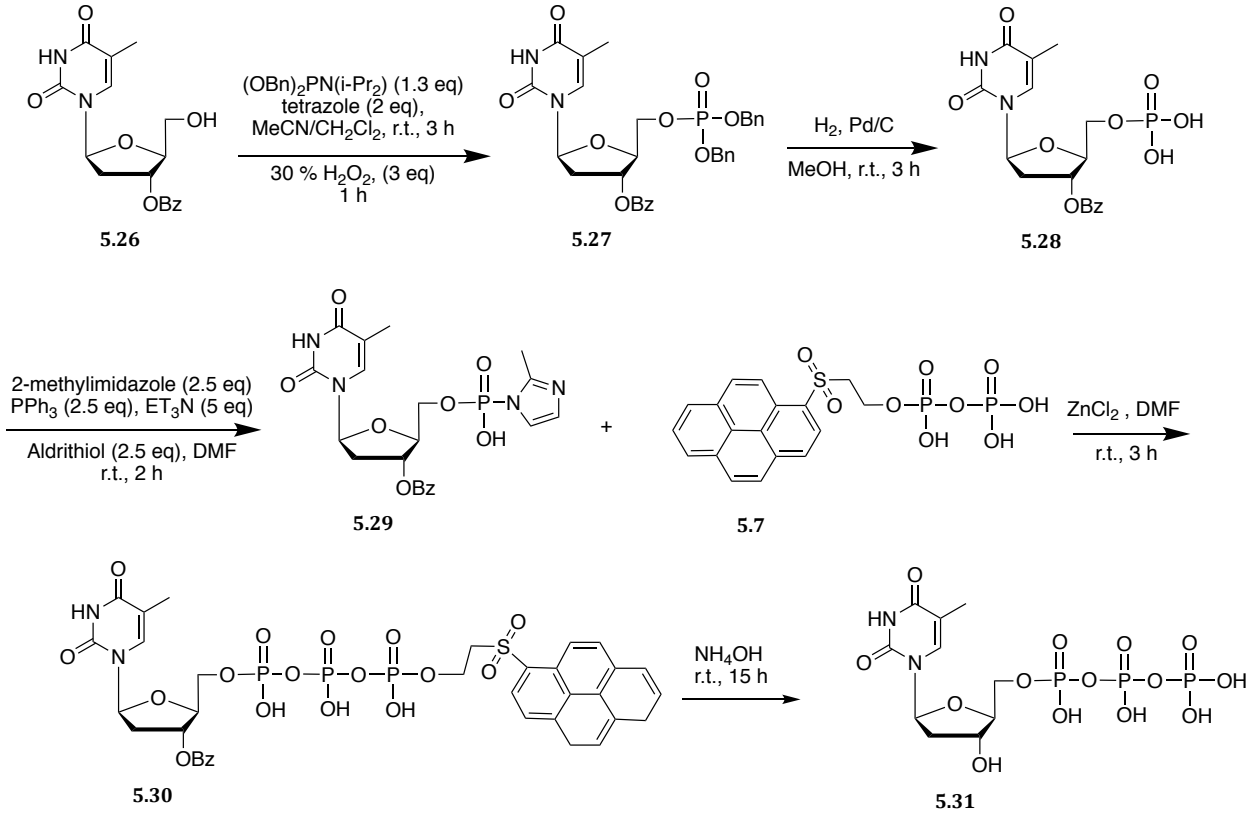
***N*^{2}-Acetyl-9-(2'-O-benzoyl-α-L-threofuranosyl)guanosine-3'-(γ-(2-(pyrenesulfonyl)ethyl)) triphosphate (5.24d)**

General procedure D with 0.5 g (0.92 mmol) of **5.23d**, 0.56 g (1.2 mmol) of **7** and 10 mL (9.2 mmol, 1.5 M in anhydrous DMF) of $ZnCl_2$ solution for 3 hours stirring at room temperature. The product was purified by silica gel chromatography with eluents [(H_2O/isopropanol from 5% to 10% then H_2O/(isopropanol-MeCN 1:1) containing 1% diisopropylethylamine (DIPEA)] to afford the white solid **5.24d**; yield: 0.74 g (80.9 %); TLC (1:10 H_2O-acetonitrile with 2% DIPEA) R_f = 0.31; ^{1}H NMR (400 MHz, D_2O) d 12.00 (m, 1H), 8.96 (d, 1H, *J* = 8.4 Hz), 8.65 (d, 1H, *J* = 8.2 Hz), 8.48-8.42 (m, 3H), 8.39-8.30 (m, 1H), 8.28 (s, 1H), 8.19-8.11 (m, 2H), 7.99-7.97 (m, 1H), 7.51 (brs, 2H), 7.25 (s, 2H), 7.13 (s, 2H), 7.01 (s, 1H), 6.01 (s, 1H), 5.85 (s, 1H), 5.09 (s, 1H), 4.63 (s, 1H), 4.19 (s, 2H), 3.90 (d, 2H, *J* = 5.6 Hz), 3.07 (s, 2H), 1.93 (s, 3H); ^{31}P NMR (162 MHz, D_2O) δ -12.87(brs, 2P), -20.71 (brs,1P); HRMS (ESI-TOF) calcd. for $C_{36}H_{30}N_5O_{17}P_3SNa$ $[M-2H+Na]^-$952.0473; found 952.0466

9-(α-L-threofuranosyl)guanosine-3'-triphosphate (5.25d)

General procedure E with 0.55 g of **5.24c**, 50mL of 33% NH_4OH(aq) for 3 hours deprotection at 37 ^oC and then 15 hours at room temperature. The product was afforded as a white solid; yield: 0.24 g (87.2 %, ε_{253} = 13700 $M^{-1}cm^{-1}$); 1H NMR (400 MHz, D_2O) δ 7.84 (m, 1H), 6.06 (d, 1H, J = 7.2 Hz), 5.84 (1 H), 4.53 (d, 1 H, J = 8 Hz), 4.41 (m, 1H), 4.31 (s, 2H),3.35 (m, 2H), 2.96 (m, 1H); ^{31}P NMR (162 MHz, D_2O) δ -9.75 (d, *J* = 20.4 Hz), - 11.27 (d, *J* = 20.4 Hz), -22.16 (d, *J* = 16.2 Hz).

Scheme 5.11: Synthesis of L-2'-deoxythymidine-5'-triphosphate (L-dTTP, 5.31)



(Scheme was used with permission from . *J. Am. Chem. Soc.* **2019**, *141*, 34, 13286-13289)

3'-Benzoyl-2'-deoxy-L-thymidine-5'-dibenzylmonophosphate (5.27)

General procedure A with 200 mg (0.6 mmol) of 26, 72.9 mg (1.18 mmol) of tetrazole, 12 mL of anhydrous solution (MeCN/CH_2Cl_2, 1:1), 260 μL (0.75 mmol) of dibenzyl-*N, N* diisopropylphosphoramidite for 3 hours reaction at room temperature. Then 2 mL of 30% H_2O_2(aq) for 1 hour oxidation reaction at room temperature. Column chromatography with eluents (MeOH/CH_2Cl_2, from 1% to 1.4 %) to afford the product **5.27** as a white solid; yield: 0.32 g (68.2%); TLC (MeOH/CH_2Cl_2, 1:40) R_f = 0.23; ^{1}H NMR (400 MHz, CD_3OD) δ 7.96 (d, 2H, *J* = 7.6 Hz), 7.55-7.52 (m, 1H), 7.41-7.38 (m, 3H), 7.31-7.25 (m, 10H), 6.30 (dd, 1H, *J* = 14.4, 2.4 Hz), 5.35 (d, 1H, *J* = 6.8 Hz), 5.06-5.03 (m, 4H), 4.31-4.28 (m, 2H), 4.24 (brs, 1H), 2.42 (dd, *J* = 19.2, 8.4 Hz), 2.21-2.14 (m, 1H), 1.72 (s, 3H); ^{13}CNMR (125.8 MHz, CD_3OD) δ 167.9, 166.7, 152.9, 137.8, 137.7, 137.7, 137.7, 137.6, 135.4, 131.5, 131.3, 130.6, 130.5, 130.5, 130.0, 130.0, 112.9, 87.0, 84.7 (d, $J_{C,P}$ = 6.5 Hz), 76.9, 71.8 (t, $J_{C,P}$ = 4.5 Hz), 69.4 (d, $J_{C,P}$ = 4.6 Hz), 38.6, 13.4 (d, $J_{C,P}$ = 2.7 Hz); ^{31}P NMR (162 MHz, CD_3OD) δ 0.19; HRMS (ESI-TOF) calcd. for $C_{31}H_{31}N_2O_9PNa$ $[M+Na]^+$ 629.1665; found 629.1685.

3'-Benzoyl-2'-deoxy-L-thymidine-5'-monophosphate (5.28)

General procedure B with 300 mg (0.49 mmol) of **5.27**, 15 mL of MeOH, and 80 mg of 10% Pd/C for 3 hours stirring at room temperature. The suspension was filtered over a pad of celite, and washed with 60 mL of MeOH containing 2% triethylamine four times to afford the product **5.28** as a white foam of triethylammonium salt; yield: 190 mg (91.3 %); TLC (MeOH/CH2Cl2, 1:10 with 1% triethylamine) R_f = 0; ^{1}H NMR (400 MHz, CD_3OD) δ 8.05- 7.89 (m, 3H), 7.64-7.50 (m, 3H), 6.47 (s, 1H), 5.66 (s, 1H), 4.40 (brs, 2H), 3.33-3.26 (m, 2H), 2.53

(s, 2H), 1.96 (s, 3H); ^{13}C NMR (125.8 MHz, CD_3OD) δ 168.1,167.3, 153.4, 138.9, 135.3, 131.8, 131.4, 130.5, 113.2, 86.9, 86.3, 78.8, 67.0, 48.2, 39.3, 13.4; ^{31}P NMR (162 MHz, CD_3OD) δ 1.02 ; HRMS (ESI-TOF) calcd. for $C_{10}H_{14}N_2O_8P$ $[M-H]^-$ 321.0488; found 321.0493.

3'-Benzoyl-2'-deoxy-L-thymidine-5'-phosphor-2-methylimidazolide (5.29)

General procedure C with 180 mg (0.42 mmol) of **5.28**, 8 mL of anhydrous DMF, 300 μL (2.1 mmol) of triethylamine, 86 mg (1.0 mmol) of 2-methylimidazole, 275 mg (1.0 mmol) of triphenylphosphine, 242 mg (1.0 mmol) of dipyridyl disulfide for 2 hours reaction at room temperature. First precipitation was achieved with 50 mL of diethyl ether. The product was resuspended with 10 mL of DMF and dropwise added to the solution containing 700 mg of sodium perchlorate, 15 mL of triethylamine in 100 mL of ethyl acetate for the second precipitation. The product was afforded as a white solid **5.29**; yield: 145 mg (96.9%); ^{31}P NMR (162 MHz, CD_3OD) δ -7.85; HRMS (ESI-TOF) $C_{21}H_{23}N_4O_8PNa$ $[M+Na]^+$ 513.1151; found 513.1146

3'-Benzoyl-2'-deoxy-L-thymidine-5'-(γ-(2-(pyrenesulfonyl)ethyl))triphosphate (5.30)

General procedure D with 170 mg (0.44 mmol) of **5.29**, 248 mg (0.53 mmol) of 2-(pyrenesulfonyl)ethyl]-pyrophosphate (**7**) and 3 mL (4.40 mmol) of $ZnCl_2$ solution (1.5 M in anhydrous DMF) for 3 hours stirring. After the reaction, the solution was precipitated by 150 mL of ether. Silica column chromatography with eluents (H_2O/(isopropanol-MeCN 1:1) from 2% to 7% containing 1% diisopropylethylamine (DIPEA) to afford the white solid **5.30**; yield: 150 mg (46.7%); TLC (H_2O/isopropanol 1:10 containing 2% DIPEA): R_f = 0.35; ^{1}H NMR

(400 MHz, DMSO-d^6+ D_2O (1:2)) δ : 8.99 (brs, 1H), 8.65 (d,1H, *J* = 8 Hz), 8.32- 8.21 (m, 5 H), 8.15-8.09 (m, 2H), 7.98-7.87 (m, 2H), 7.70 (s, 1H), 7.56-7.55 (m, 1H), 7.41 (s, 1H), 6.22 (t, 1H, J = 4 Hz) , 5.10 (s, 1H), 4.35 (s, 2H), 4.22 (s, 1H), 4.11 (s, 1H), 3.92 (t, 2H, *J* = 4 Hz), 2.33 (s, 1H), 1.88 (s, 1H); ^{31}P NMR (162 MHz, DMSO-d^6 + D_2O (1:2)) δ -10.99 (brs, 1P), -11.26 (brs, 1P) -20.94 (brs, 1P); HRMS (ESI-TOF) calcd. For $C_{35}H_{31}N_2O_{17}P_3SNa$ $[M-2H+Na]^-$ 899.0459; found 899.0454.

L-2'-deoxythymidine-5'-triphosphate (5.31).

General procedure E with 20 mg (0.04 mmol) of **5.30**, 10 mL of 33% NH_4OH(aq) for 15 hours stirring at room temperature. The product **5.31** was afforded as a white solid; yield: 8 mg (83.2%); ^{1}H NMR (400 MHz, D_2O) δ 7.84 (s, 1H), 6.28 (t, 1H, *J* = 6.0 Hz), 5.17 (s, 1H), 4.54 (s, 1H), 4.15 (d, 1H, *J* = 10.4 Hz), 2.83-2.77 (m, 1H), 2.49 (s, 1H), 2.45 (s, 2H), 2.00 (s, 3H); ^{31}P NMR (162 MHz, D_2O) δ -4.38 (d, *J* = 16.2 Hz), -10.65 (d, *J* = 16.2 Hz), -19.74 (brs).

Analytical Reverse-Phase HPLC Analysis.

2 μL of reaction crude (5.10-5.11a-d, and 5.16-5.17, and 5.22-5.23a-d, and 5.28-5.29) in DMF or 1 μL of purified pyrene substituted nucleoside triphosphates (5.12a-d, and 5.18, and 5.24a-d, and 5.30) in methanol or 3 μL of nucleoside triphosphates (5.13a-d, and 5.19, and 5.25a-d, and 5.31) in H_2O was added to 50 μL of 0.1 M triethylammonium acetate (TEAA) buffer pH 7.0. The solution was centrifuged by 4000 rpm for 2 minutes at room temperature and 30 μL of supernatant was injected for HPLC analysis. Reaction progress of the synthesis of nucleoside monophosphates (5.10a-d, and 5.16, and 5.22a-d, and 5.28), nucleoside phosphor- 2-methylimidazolides (5.11a-d, and 5.17, and 5.23a-d, and 5.29), pyrene

substituted nucleoside triphosphates (5.12a-d, and 5.18, and 5.24a-d, and 5.30) and nucleoside triphosphates (5.13a-d, and 5.19, and 5.25a-d, and 5.31) in figure 5.1 to 5.10 was monitored by analytical HPLC chromatography with the gradient from 0% to 50 % of acetonitrile in 0.1 M triethylammonium acetate buffer pH 7.0 over forty minutes. Purity determination of synthesized 5'-dNTPs (5.13a-d) was monitored by the co-injection with the commercial 5'- dNTPs purchased from Sigma-Aldrich by analytical HPLC chromatography in figure S11 to S14 with the gradient from 0% to 10 % of acetonitrile in 0.1 M triethylammonium acetate buffer pH 7.0 over forty minutes of analytical run.

PCR Fidelity Assay.

A 100 µl PCR 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), 0.5 µM of Fwd [CAACCGGTCCCCACGTTGCC] and Rev [AACGGCTGGGAGAACCTGGTTCTCAATGTA] PCR primers, 400 µM dNTPs (Chemically Synthesized vs Life Technologies), 4.4 ng of pGDR11 KOD-RS plasmid [Target sequence: CAACCGGTCCCCACGTTGCCGTTGCCAAGAGGTTGGCCGCGAGAGGAGTCAAAATACGCCCTGGA ACGGTGATAAGCTACATCGTGCTCAAGGGCTCTGGGAGGATAGGCGACAGGGCGATACCGTTCG ACGAGTTCGACCCGACGAAGCACAAGTACGACGCCGAGTACTACATTGAGAACCAGGTTCTCCCA GCCGTT] with final concentration of 5 units/100 µL Taq polymerase (NEB). The PCR conditions were: 95 ^oC, 2.5 min (melt), 95 ^oC, 30 sec, 62 ^oC, 45 sec, 72 ^oC, 30 sec for 20 cycles and an additional 72 ^oC, 2 min. The amplified amplicon (200bp) was agarose purified, ligated into a TOPO-TA vector, and subsequently cloned into NEB DH5a E. coli competent cells following the manufacturer's instructions. Individual colonies were grown in liquid media

and sequenced using the M13F primer by Retrogen, San Diego, CA. DNA sequences were aligned and analyzed using MEGA7 software. Five sequences clones were analyzed for each condition to give a total of 750 nucleotide positions.

TNA Transcription Assay.

Primer-extension reactions were performed in a final volume of 20 µl. Each reaction contained 10 pmol of primer [IR680 - 5'-GTCCCCTTGGGGATACCACC-3'] annealed to 10 pmol of template [5'ATCGAGTACAGTCAGATCGATATGATCTATATATTAATTAGGTGGTATCCCCAAGGGGAC-3'], 1x ThermoPol buffer, 0.5 µM KOD-RS, 100 µM of each tNTP. Reactions were incubated for 120 min at 55ÅãC, quenched with stop buffer (40% Formamide and 1x TBE buffer, 10 mM EDTA), and analyzed by 10% denaturing urea PAGE.

Table 5.1. Crystal data and structure refinement for pyrene pyrophosphate (5.7).

Identification code	jcc7 (Liao Jen-Yu)
Empirical formula	$C_{24}H_{31}NO_9P_2S$
Formula weight	571.50
Temperature	133(2) K
Wavelength	0.71073 .
Crystal system	Orthorhombic
Space group	*Pbca*
Unit cell dimensions	a = 9.0890(17) . a= 90Å.
	b = 17.092(3) . b= 90Å.
	c = 34.022(6) . g = 90Å.

Volume	5285.2(17) .3
Z	8
Density (calculated)	1.436 Mg/m^3
Absorption coefficient	0.297 mm^{-1}
F(000)	2400
Crystal color	colorless
Crystal size	0.585 x 0.353 x 0.044 mm^3
Theta range for data collection	1.197 to 26.372Å
Index ranges	$-11 \leq h \leq 11, -21 \leq k \leq 21, -42 \leq l \leq 42$
Reflections collected	52512
Independent reflections	5402 [R(int) = 0.0393]
Completeness to theta = 25.500Å	99.9 %
Absorption correction	Semi-empirical from equivalents
Max. and min. transmission	0.8621 and 0.7124
Refinement method	Full-matrix least-squares on F2
Data / restraints / parameters	5402 / 0 / 349
Goodness-of-fit on F2	1.213
Final R indices [I>2sigma(I) = 4887 data]	R1 = 0.0569, wR2 = 0.1343
R indices (all data, 0.80 .)	R1 = 0.0628, wR2 = 0.1370
Largest diff. peak and hole	0.611 and -0.388 e^{-3}

Table 5.2. PCR Fidelity.

Commercial dNTPs			Chemically synthesized dNTPs		
Expected	**Mutation**	**Count**	**Expected**	**Mutation**	**Count**
A	C	0	A	C	0
	G	2		G	1
	T	0		T	0
C	A	0	C	A	0
	G	0		G	0
	T	0		T	0
G	A	0	G	A	0
	C	0		C	0
	T	0		T	0
T	A	0	T	A	0
	C	0		C	0
	G	0		G	0
Total nucleotide positions = 750			Total nucleotide positions = 750		

References

1. L. A. Loeb.; R. J. Monnat, Jr.; DNA polymerases and human disease. *Nat. Rev. Genet*. **2008** *9*, 594-604.
2. W. J. Wu.; W. Yang,.; M. D. Tsai,.; How DNA polymerases catalyse replication and repair with contrasting fidelity. *Nat. Rev. Chem*. **2008**, *1,* 0068.
3. S. Goodwin.; J. D. McPherson.; W. R. McCombie.; Coming of age: ten years of next generation sequencing technologies. *Nat. Rev. Genet*. **2016**, *17*, 333-351.

4. Y. Erlich.; D. Zielinski.; DNA Fountain enables a robust and efficient storage architecture. *Science* **2017**, *355*, 950-954.
5. T. Chen,.; F. E. Romesberg,.; Directed polymerase evolution. *FEBS Lett.* **2014**, *588,* 219-229.
6. D. Loakes.; P. Holliger.; Polymerase engineering: towards the encoded synthesis of unnatural polymers. *Chem. Commun*. **2009**, 4619-4631.
7. V. B. Pinheiro et al.; Synthetic genetic polymers capable of heredity and evolution. *Science,* **2012**, *336*, 341-344.
8. H. Yu.; S. Zhang.; J. C. Chaput.; Darwinian evolution of an alternative genetic system provides support for TNA as an RNA progenitor. *Nat. Chem*. **2012**, *4,* 183-187.
9. A. I. Taylor et al.; Catalysts from synthetic genetic polymers. *Nature* **2015**, *518*, 427-430
10. Y. Wang .; A. K. Ngor.; A. Nikoomanzar.; J. C. Chaput.; Evolution of a general RNAcleaving FANA enzyme. *Nat. Commun*., **2018**, *9*, 5067
11. K. Burgess.; D. Cook.; Syntheses of Nucleoside Triphosphates. *Chem. Rev*. **2000**, *100*, 2047-2060.
12. A. R. Kore.; B. Srinivasan.; Recent Advances in the Synthesis of Nucleoside Triphosphates. *Curr. Org. Syn*. **2013**, *10*, 903-934.
13. M. Hollenstein.; Nucleoside triphosphates--building blocks for the modification of nucleic acids. *Molecules .* **2012**, *17*, 13569-13591.
14. M. Yoshikawa.; T. Kato,.; T. Takenishi.; A novel method for phosphorylation of nucleosides to 5'-nucleotides. *Tetrahedron Lett*. **1967**, *8*, 5065-5068.
15. J. Ludwig.; F. Eckstein.; Rapid and efficient synthesis of nucleoside 5'-0 -(1-

thiotriphosphates), 5'-triphosphates and 2',3'-cyclophosphorothioates using 2-chloro-4H-1,3,2-benzodioxaphosphorin-4-one. *J. Org. Chem*. **1989**, *54*, 631-635.

16. G. S. Cremosnik,.; A. Hofer.; H. J. Jessen.; Iterative synthesis of nucleoside oligophosphates with phosphoramidites. *Angew. Chem. Int. Ed.* **2014**, *53*, 286-289
17. A. Hofer et al.; A Modular Synthesis of Modified Phosphoanhydrides. *Chemistry.* **2015**, *21*, 10116-10122.
18. S. P. Sau.; J. C. Chaput.; A Gram-Scale HPLC-Free Synthesis of TNA Triphosphates Using an Iterative Phosphorylation Strategy. *Org. Lett*. **2017**, *19*, 4379-4382.
19. D. E. Hoard, D. G. Ott, Conversion of Mono- and Oligodeoxyribonucleotides to 5-Triphosphates. J. Am. Chem. Soc. 87, 1785-1788 (1965).
20. K. U. Schöning et al.; Chemical etiology of nucleic acid structure: the threofuranosyl-(3'-->2') oligonucleotide system. *Science*. **2000***, 290*, 1347-1351.
21. A. C. Larsen et al.; A general strateg*y for expanding polymerase function by droplet microfluidics. Nat. Commun. **2016**, 7*, 11235.
22. Z. Wang.; W. Xu.; L. Liu,.; T. F. Zhu.; A synthetic molecular system capable of mirror-image genetic replication and transcription. *Nat. Chem*. **2016**, *8*, 698-704.
23. H. R. Hulett.; Non-enzymatic hydrolysis of adenosine phosphates. *Nature* **1970**, *225*, 1248-1249.
24. T. A. Steitz.; S. J. Smerdon.; J. Jager.; C. M. Joyce.; A unified polymerase mechanism for nonhomologous DNA and RNA polymerases. *Science,* **1994**, *266*, 2022-2025.
25. N. Chim.; C. Shi.; S. P. Sau.; A. Nikoomanzar.; J. C. Chaput.; Structural basis for TNA synthesis by an engineered TNA polymerase. *Nat Commun*. **2017,** *8*, 1810.

26. S. Bala.; J. Y. Liao,.; H. Mei.; J. 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.

Chapter 6

Conclusions

The research accomplishments presented in this thesis have described optimal synthetic strategies for α-L-threofuranosyl (TNA) monomers, nucleoside phosphoramdites, and nucleoside triphosphates synthesis with high scalability. 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.

Here 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.

Then a 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 (Chapter 4), and 7-deaza-7-phenyl TNA-Guanosine nucleoside triphosphate, Fluorescence activated tricyclic TNA-cytosine nucleoside triphosphate and several modified TNA-Uracil nucleoside triphosphates synthesis.

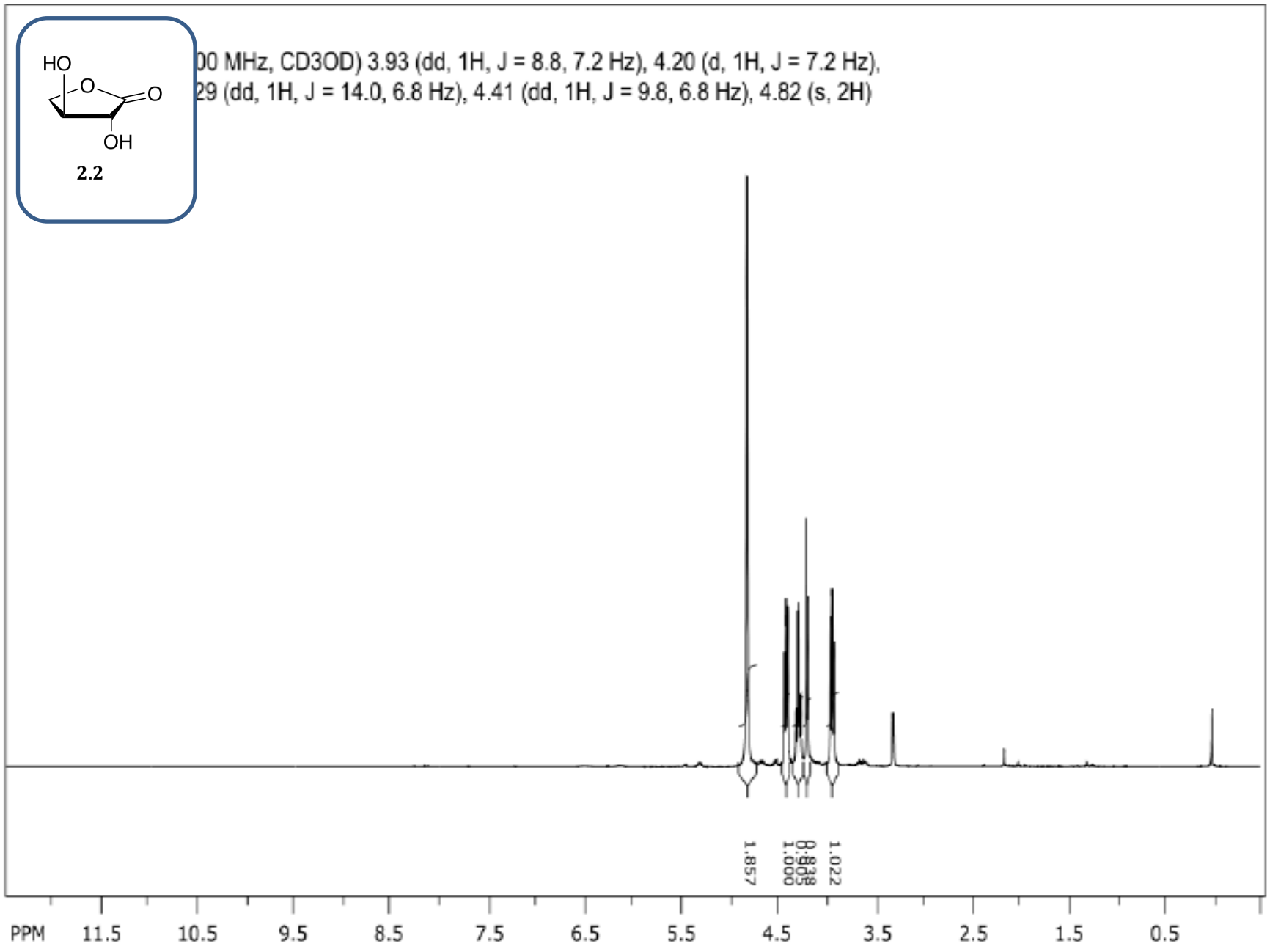
We used this idea for the chemical synthesis of 2′-deoxy-α-L-threofuranosyl nucleoside 3′-triphosphates (dtNTPs). The key synthetic challenge was resolved by introducing the triethylborane as the radical initiator for room temperature 2'-deoxygenation reaction of TNA nucleosides. Stepwise phosphorylation strategy developed in chapter 4 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.

Finally, we developed a novel pyrophosphate reagent and described a fundamentally different approach to generate triphosphate derivatives that are purified by silica gel chromatography and converted to the desired compounds on scales vastly exceeding those achievable by HPLC. The power of this approach is demonstrated through the synthesis of a broad range of natural and unnatural nucleoside triphosphates (dNTPs and xNTPs) using protocols that are efficient, inexpensive, and operationally straightforward.

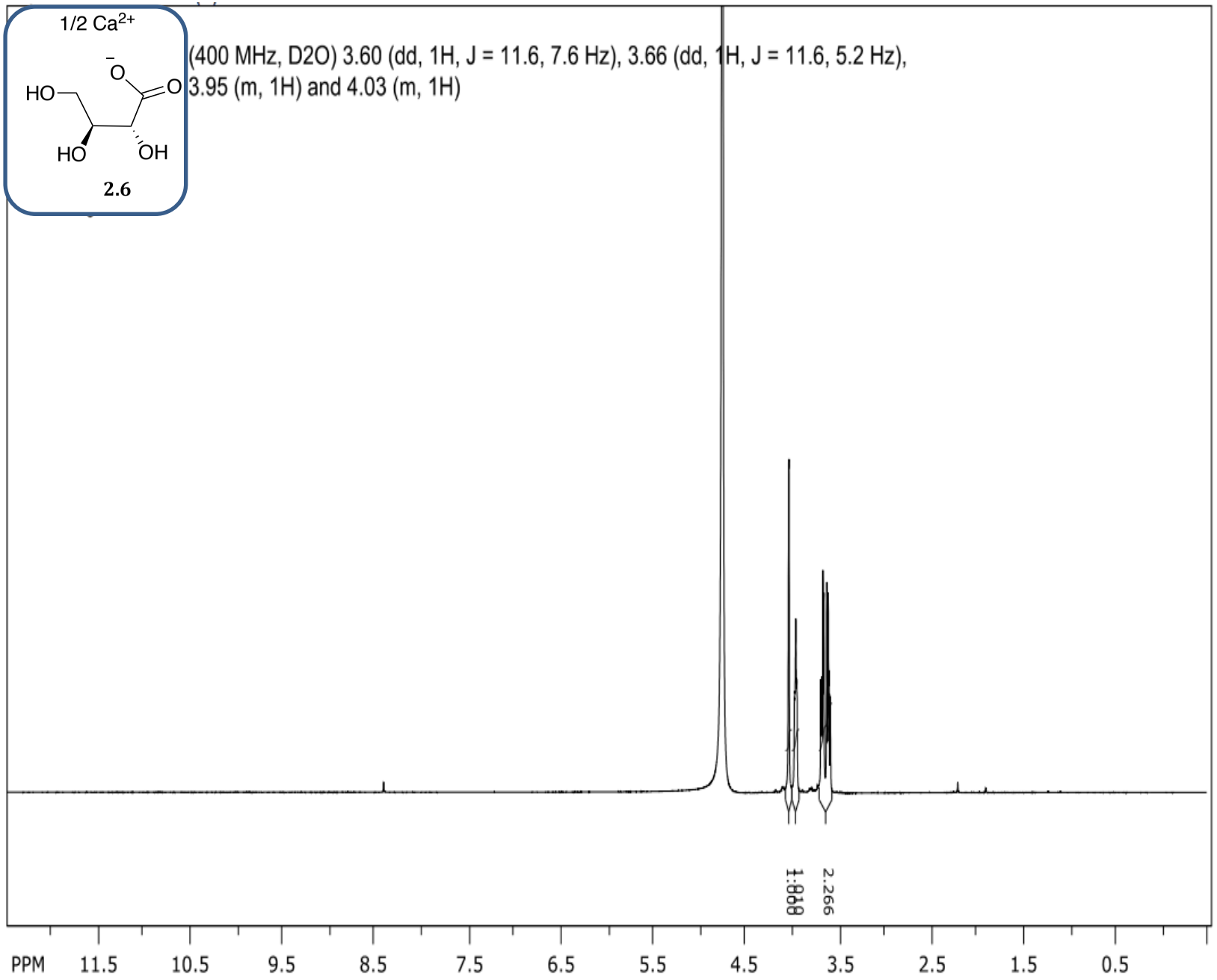
Current development of a P(V) pyrophosphate reagents in the lab has made it possible to synthesize the gram scale of natural/modified nucleoside triphosphate using a silica gel chromatography instead of using HPLC purification. Extensive applications on the other phosphorylated reagent analogues are under consideration. Look into future prospect, installation of solid support on monophosphate or pyrophosphate reagents are feasible idea to improve the purification process of phosphorylated molecules. And this kind of reagents could be used as a substrate for our engineered polymerases for checking their mechanism.

APPENDIX A

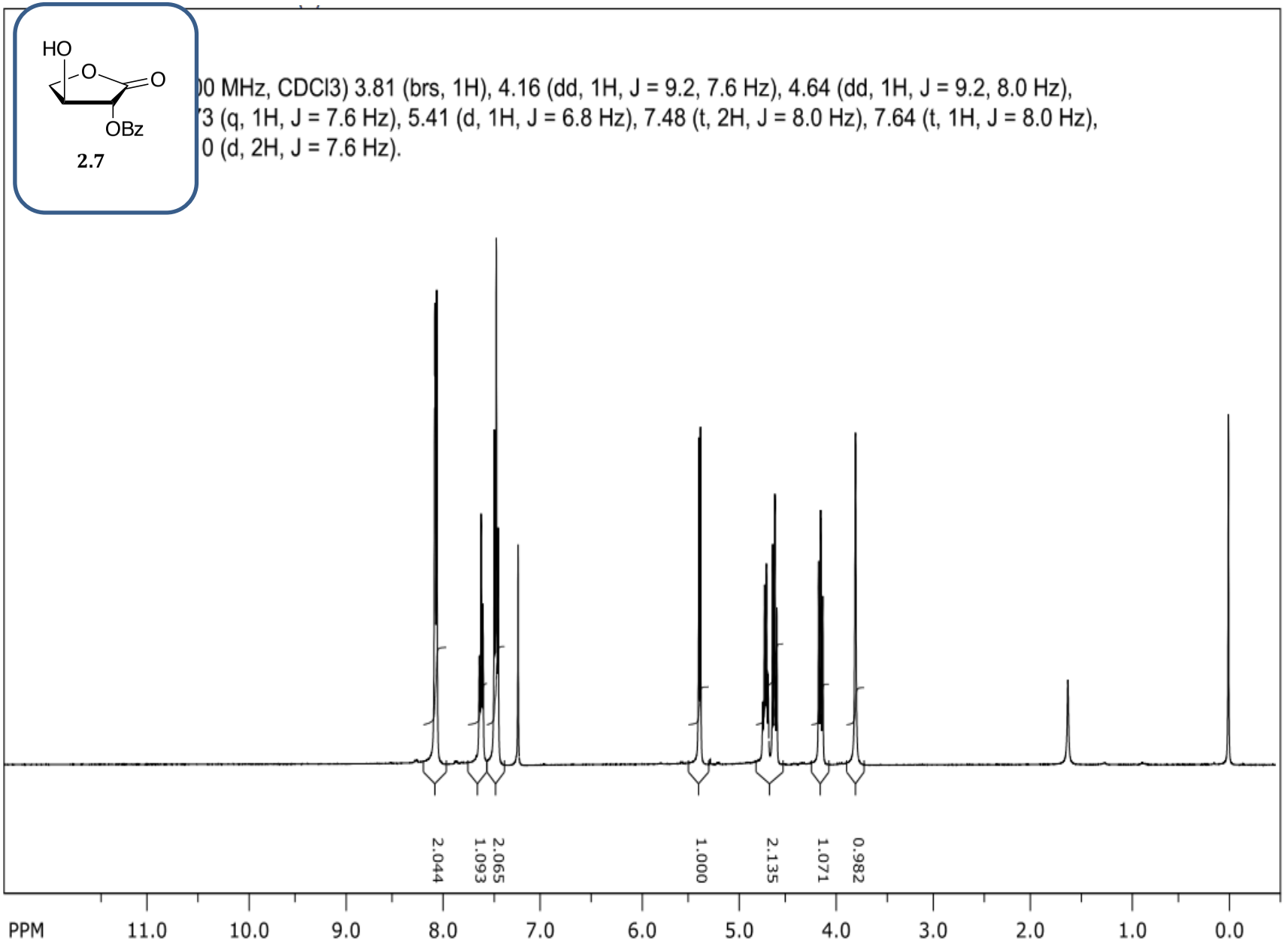
SUPPLEMENTAL NMR SPECTRA



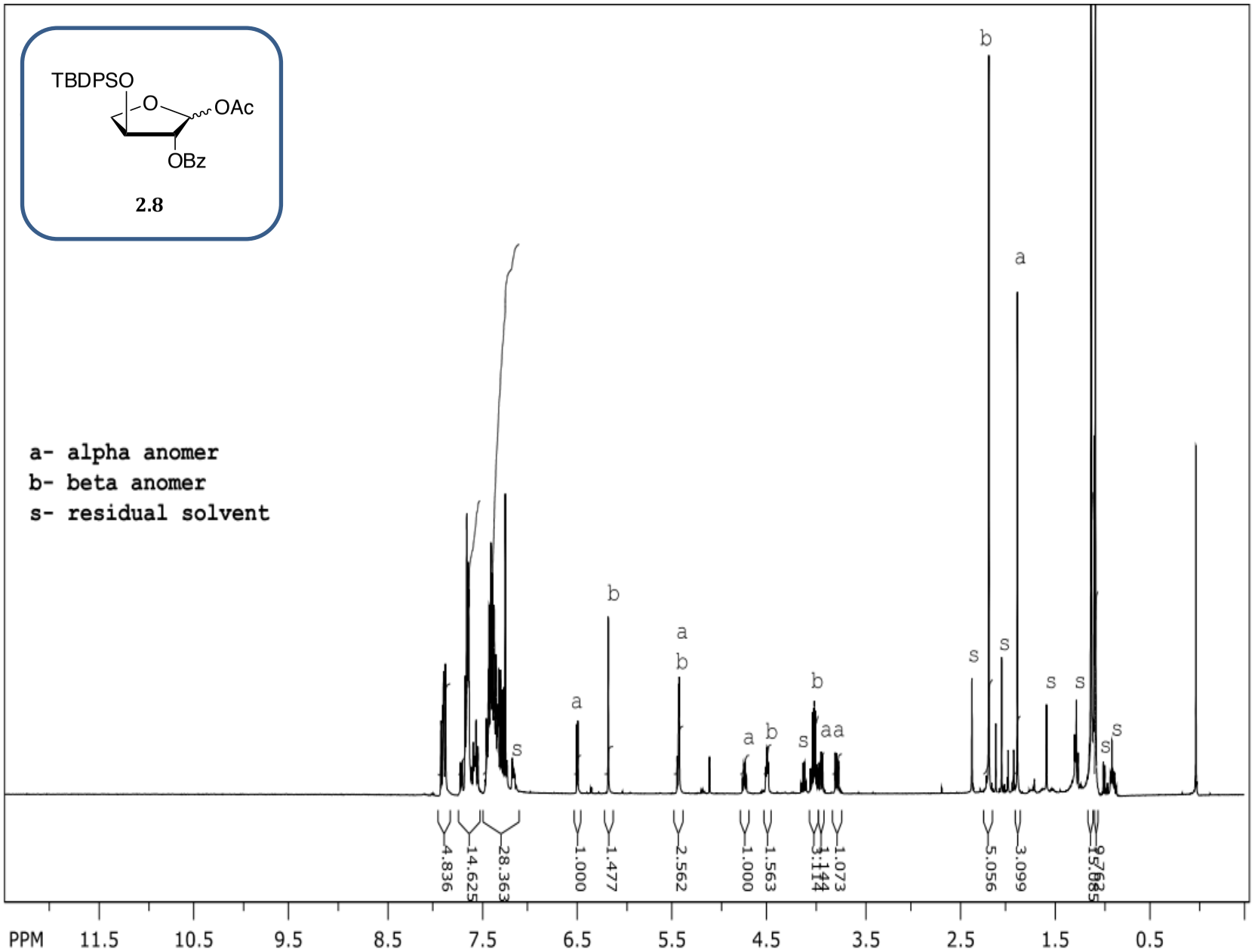
^{1}H NMR spectrum of compound **2.2** (400 MHz, $CDCl_3$)



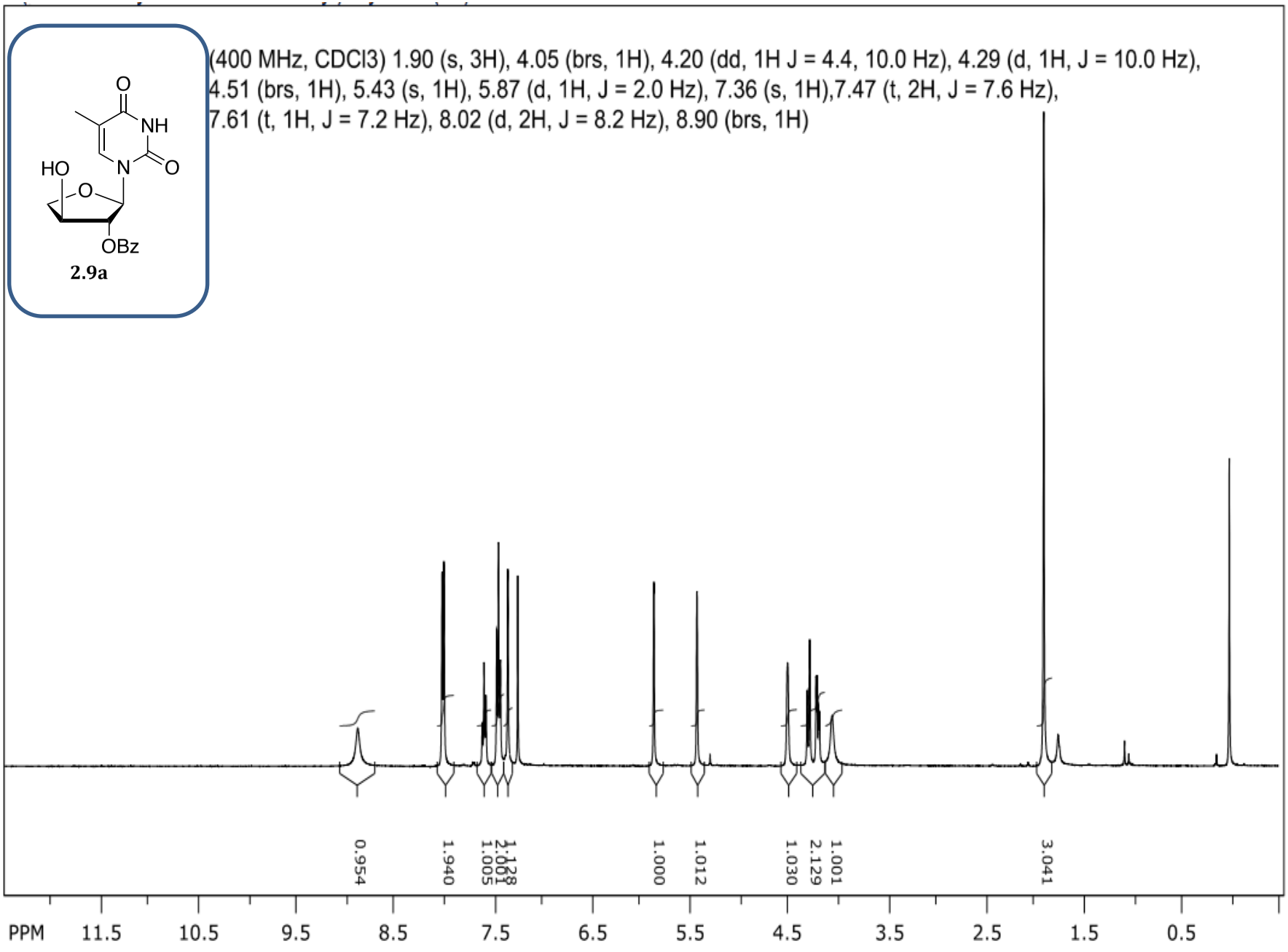
^{1}H NMR spectrum of compound **2.6** (400 MHz, D_2O)



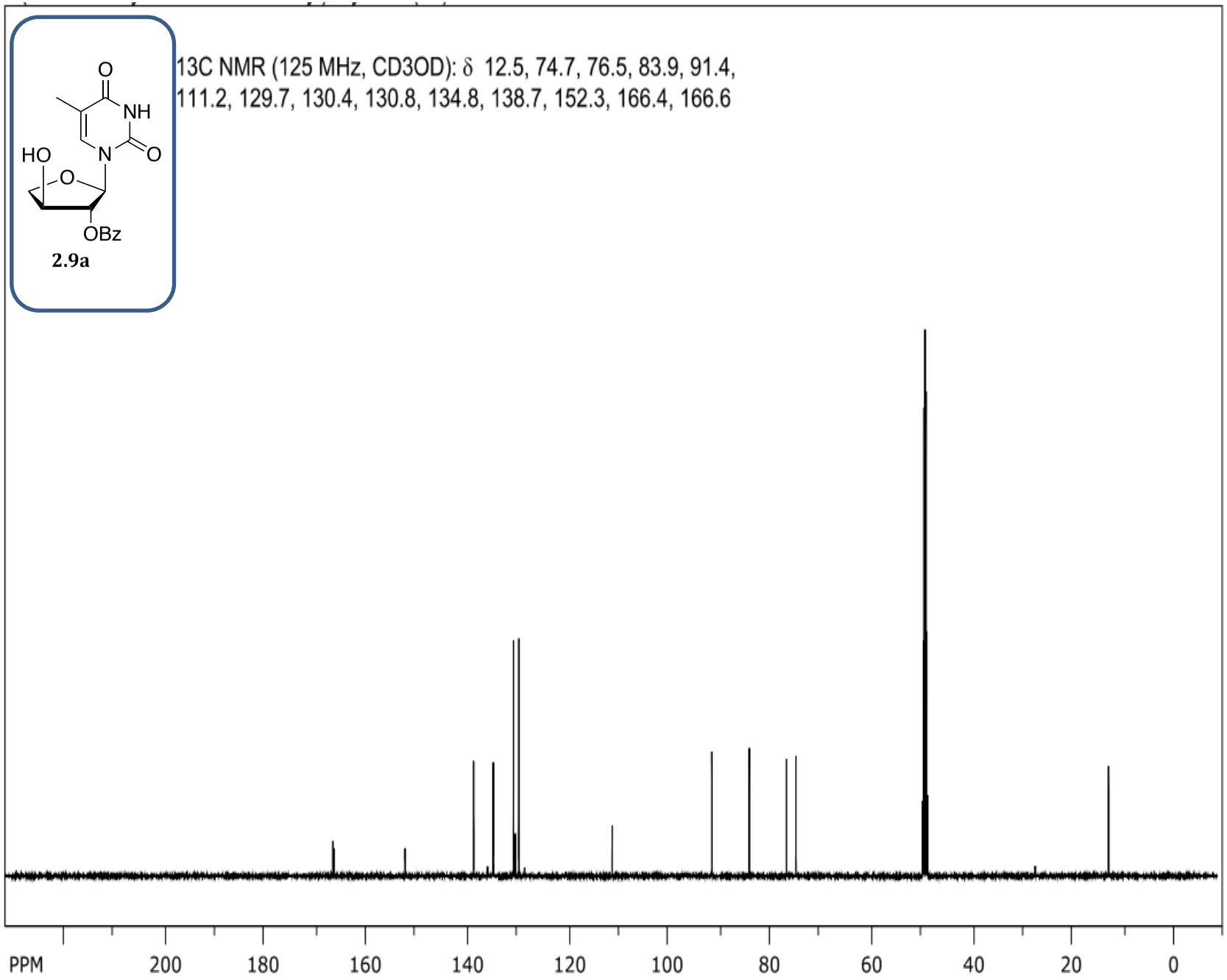
^{1}H NMR spectrum of compound **2.7** (400 MHz, $CDCl_3$)



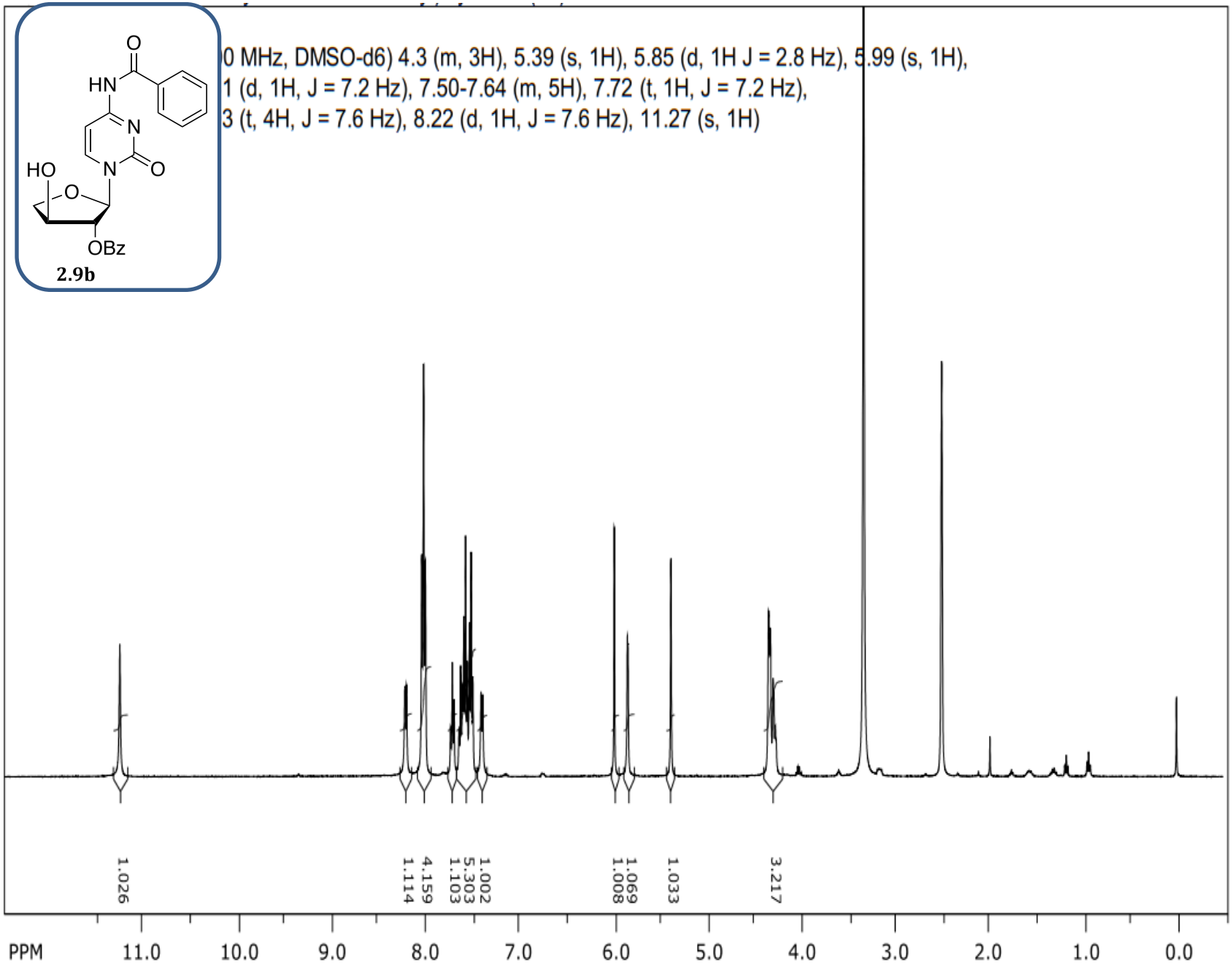
^{1}H NMR spectrum of compound **2.8** (400 MHz, $CDCl_3$)



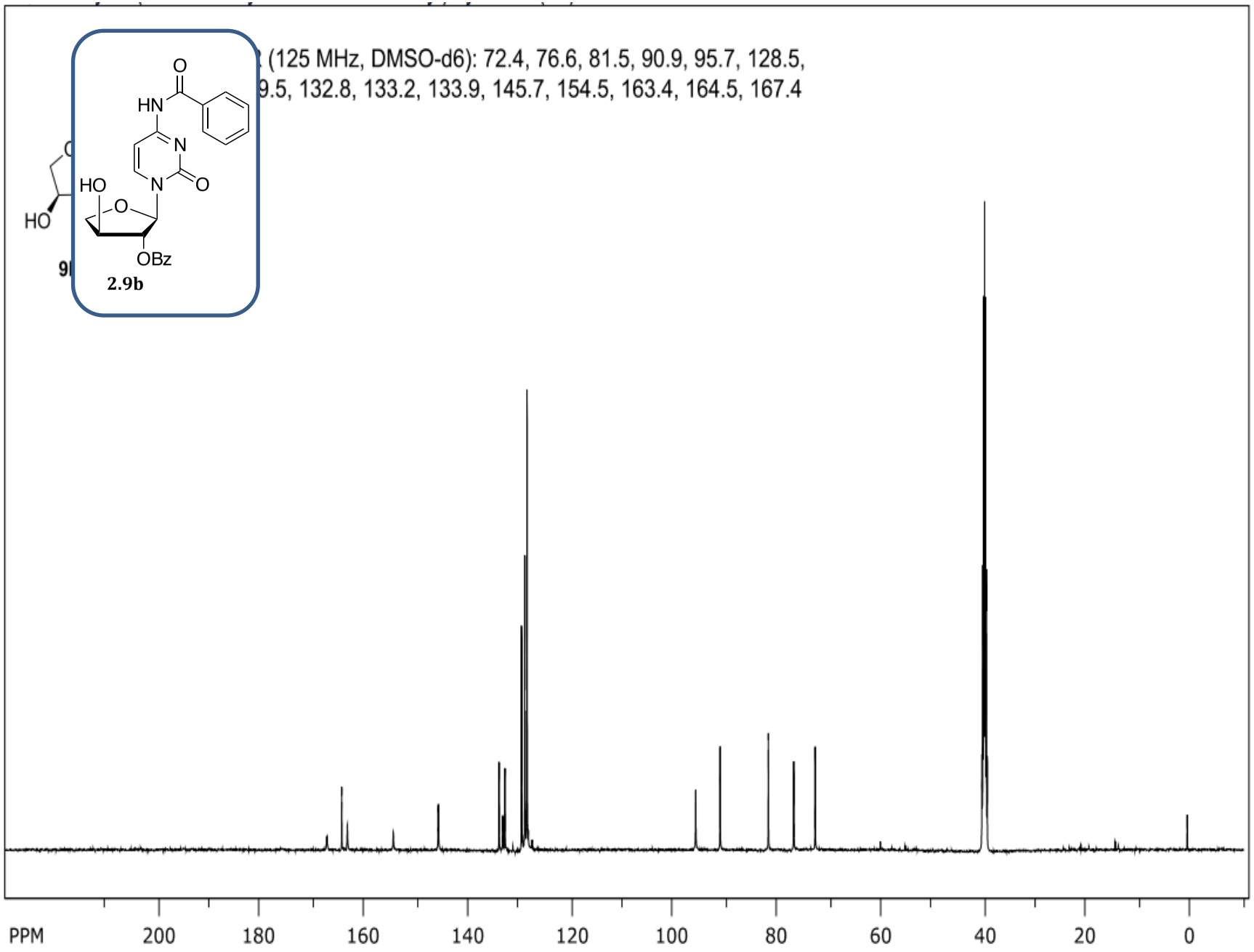
^{1}H NMR spectrum of compound **2.9a** (400 MHz, $CDCl_3$)



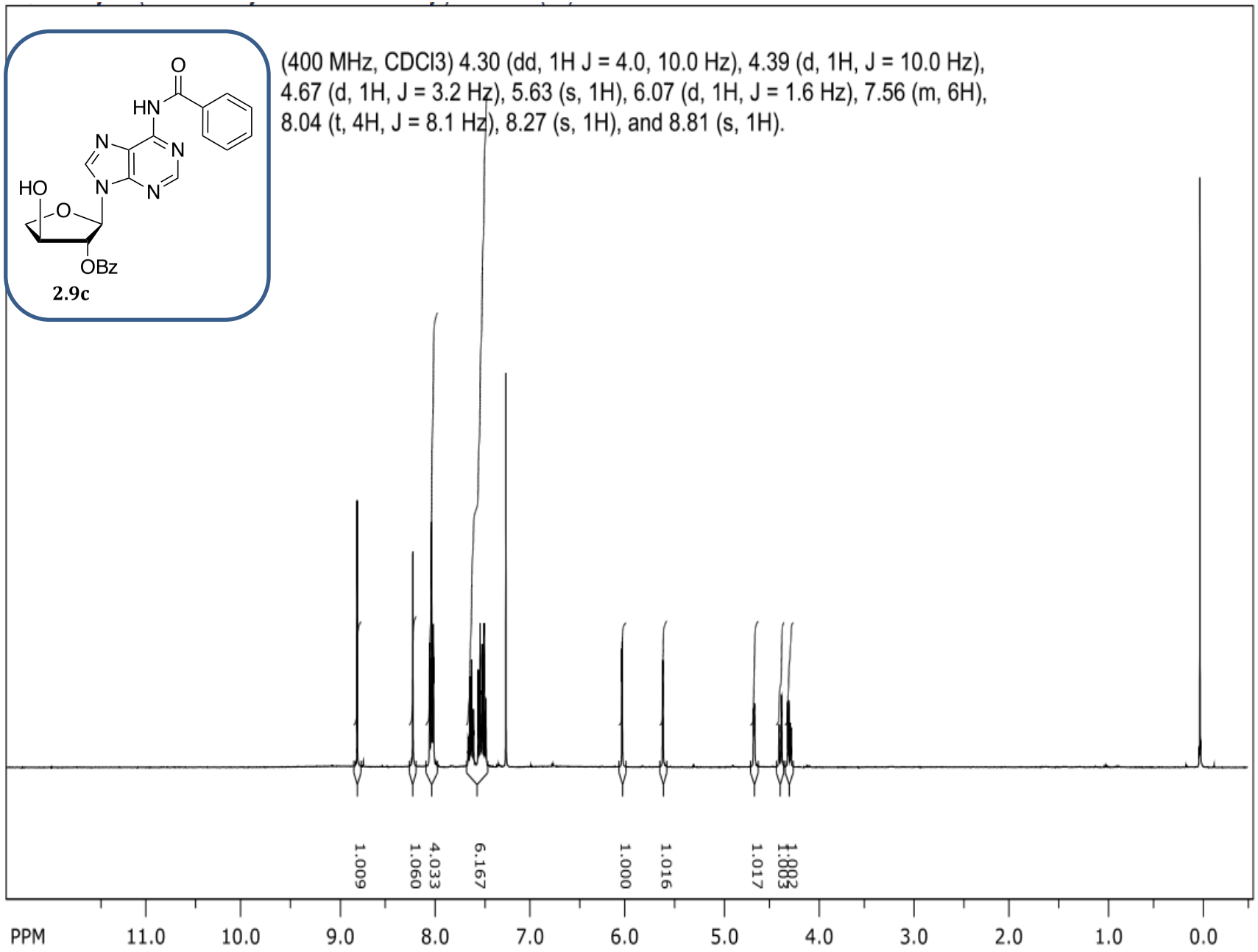
^{13}C NMR spectrum of compound **2.9a** (125 MHz, $CDCl_3$)



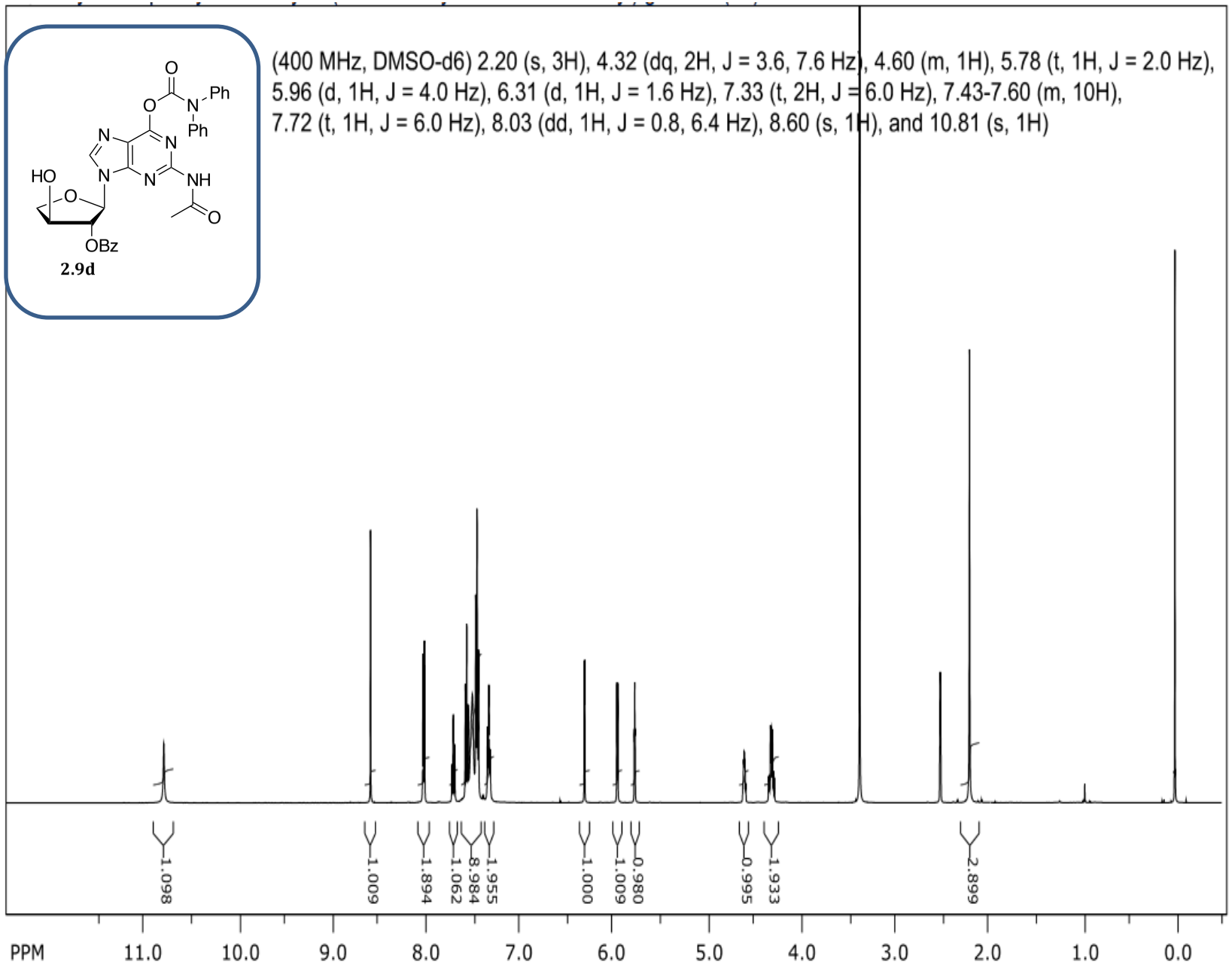
^{1}H NMR spectrum of compound **2.9b** (400 MHz, DMSO-d_6)



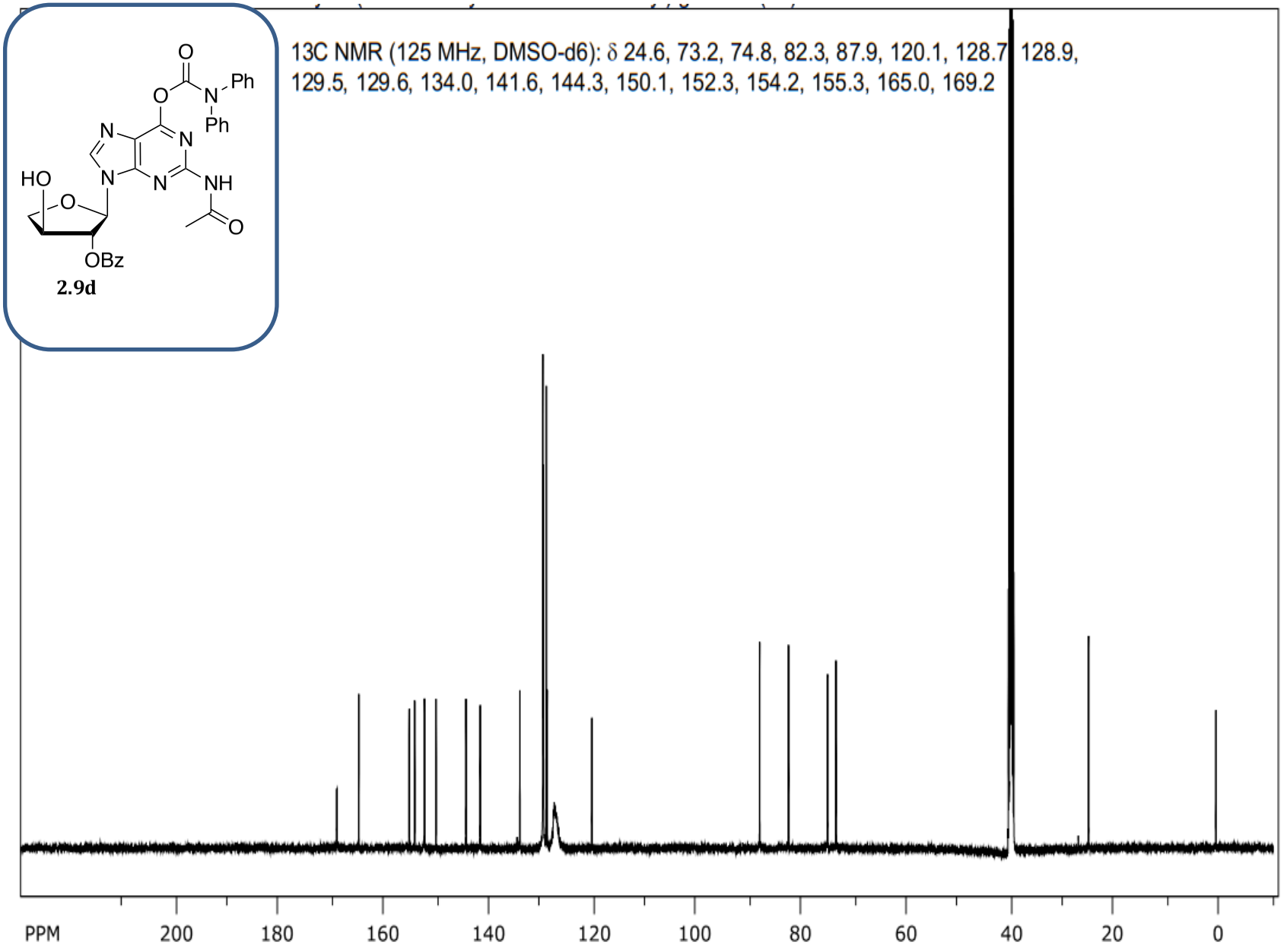
^{13}C NMR spectrum of compound **2.9b** (125 MHz, DMSO-d_6)



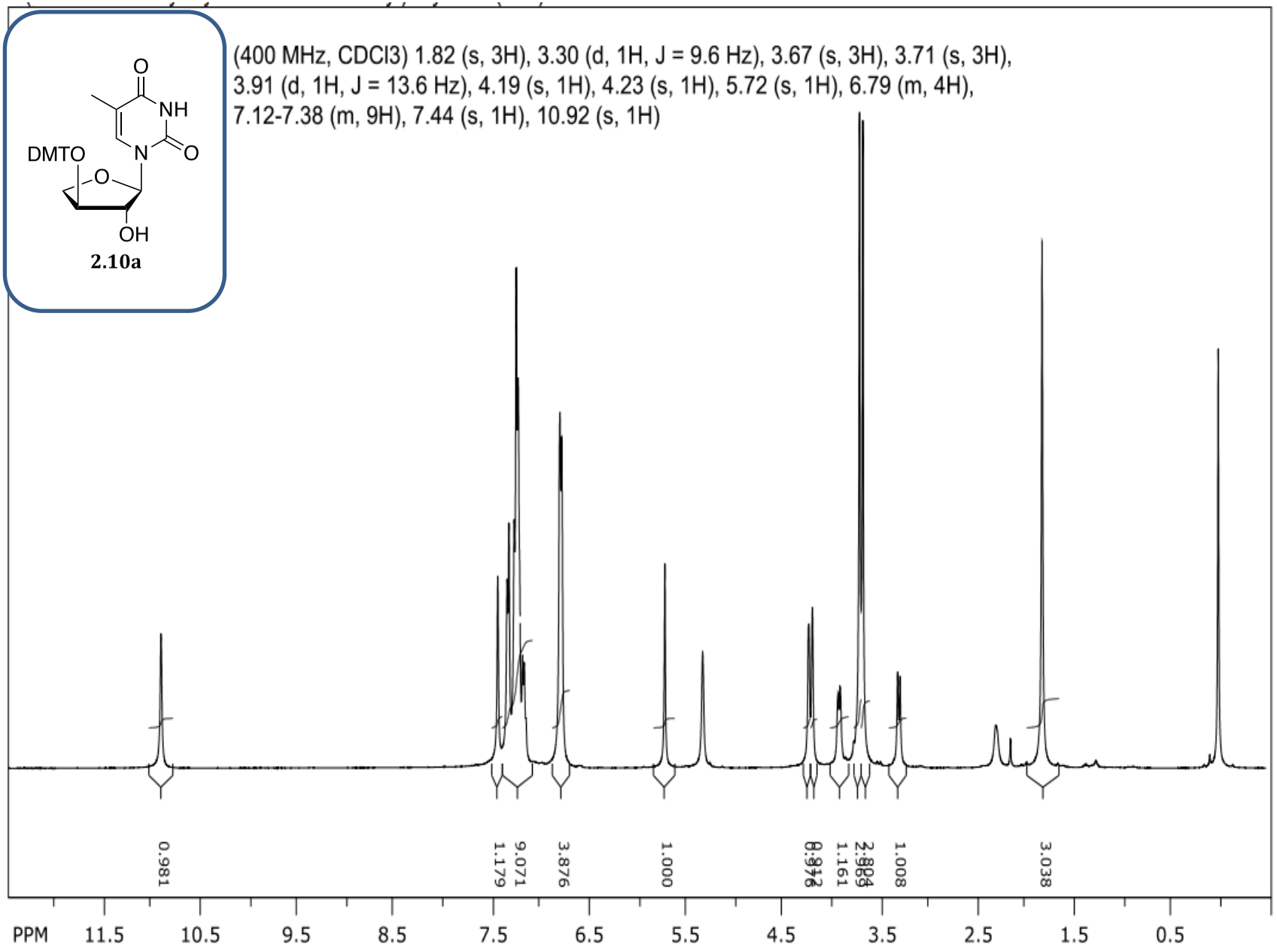
^{1}H NMR spectrum of compound **2.9c** (400 MHz, DMSO-d_6)



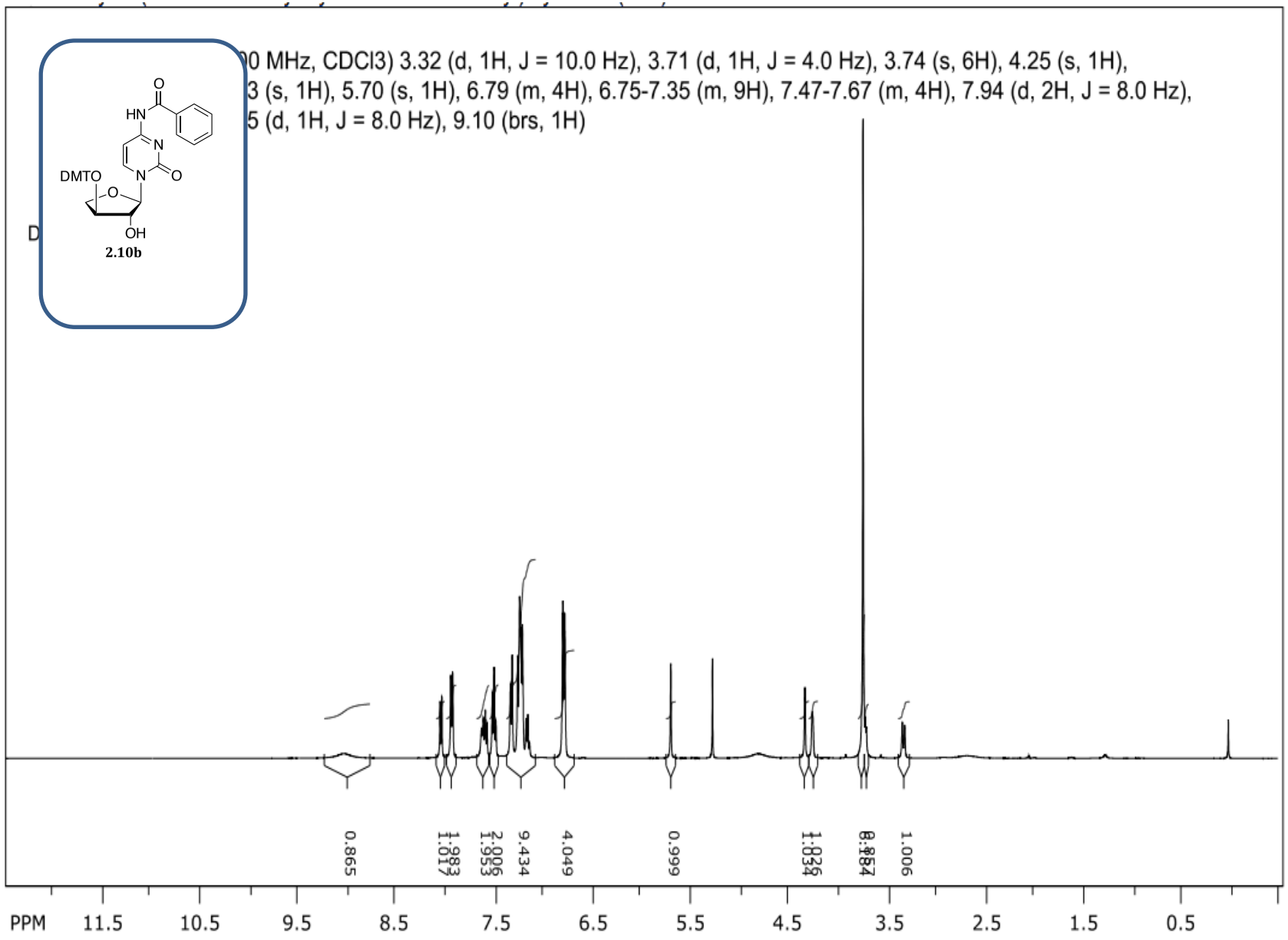
^{1}H NMR spectrum of compound **2.9cd**(400 MHz, DMSO-d_6)



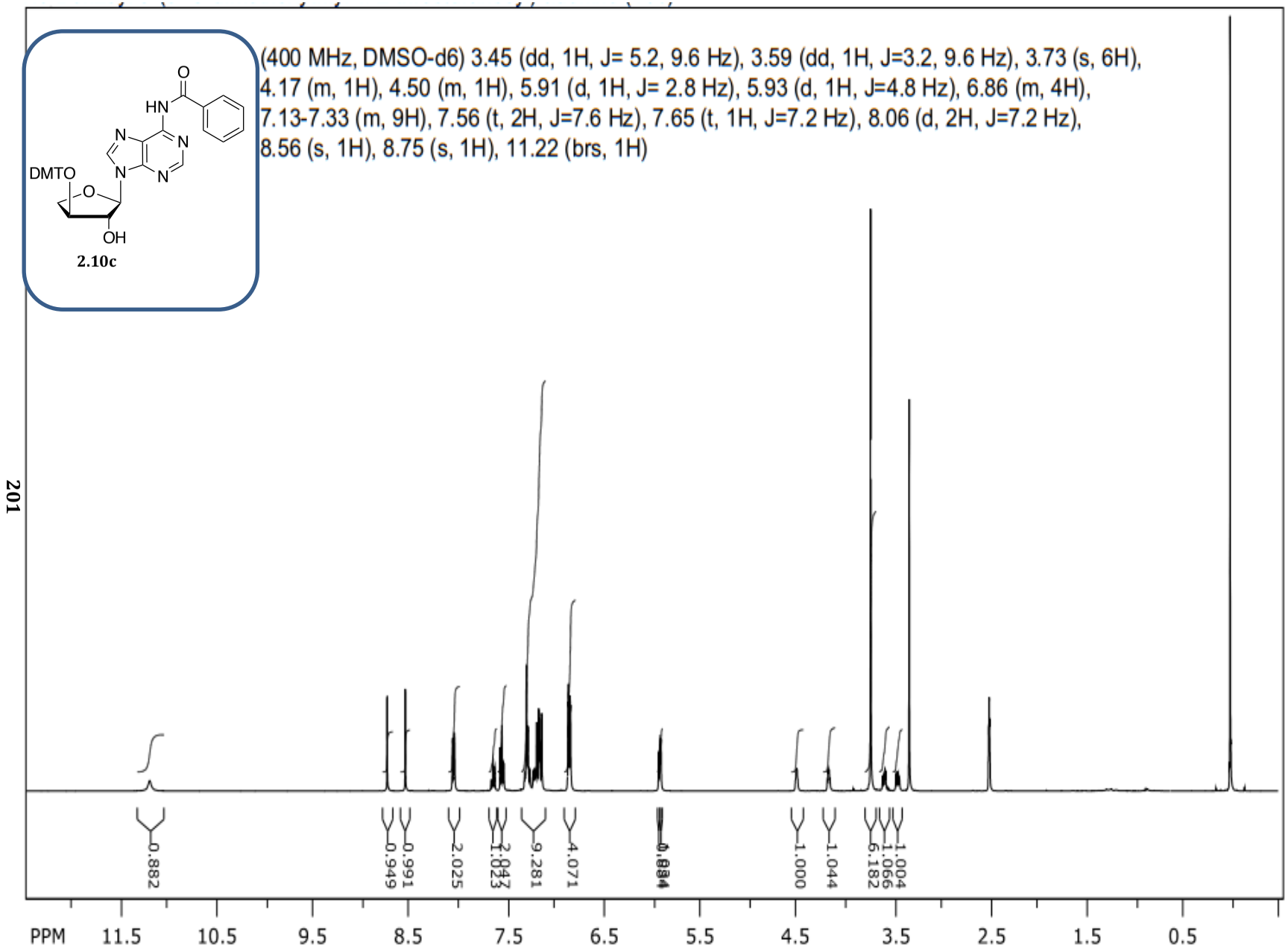
^{13}C NMR spectrum of compound **2.9d** (125 MHz, DMSO-d_6)



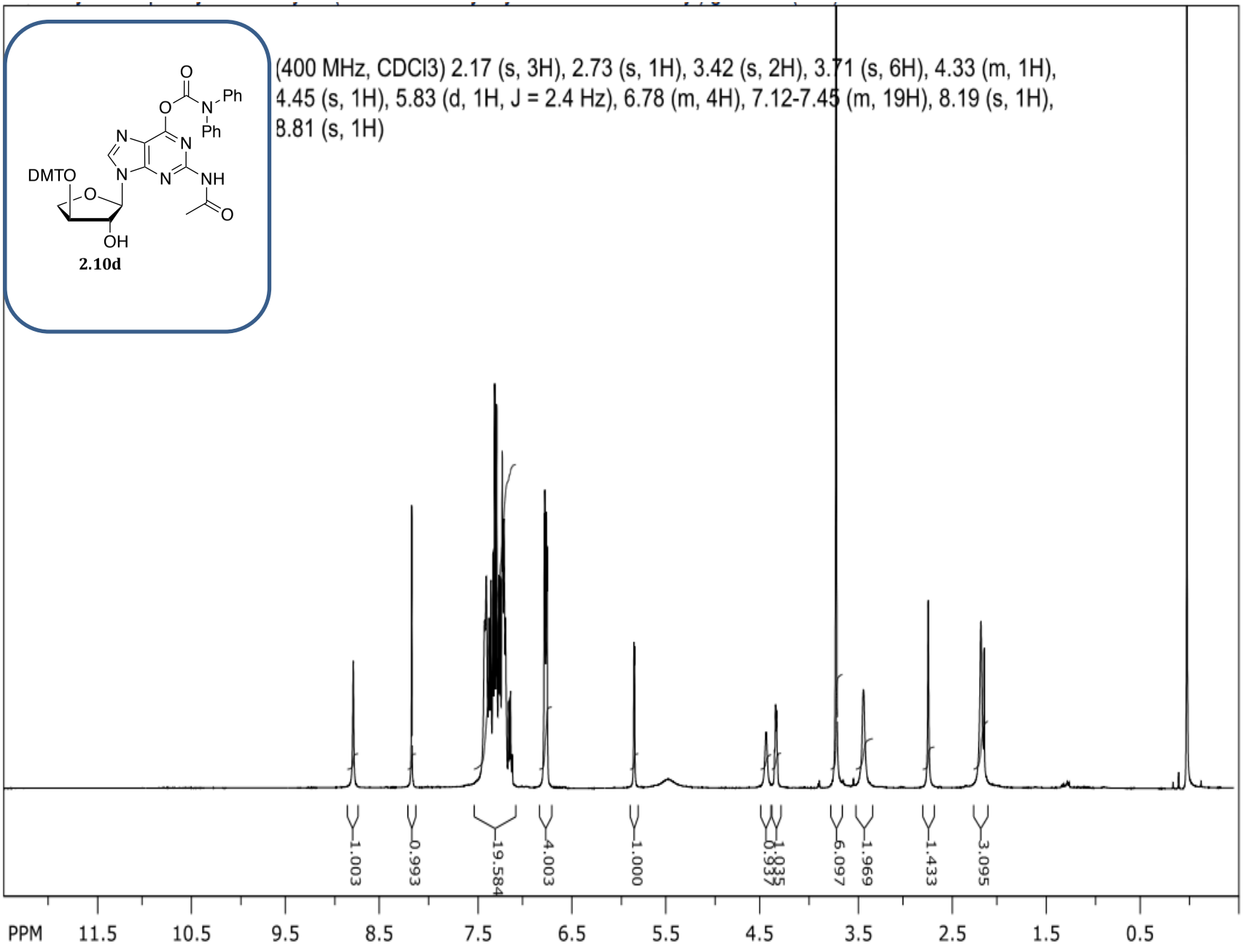
^{1}H NMR spectrum of compound **2.10a** (400 MHz, $CDCl_3$)



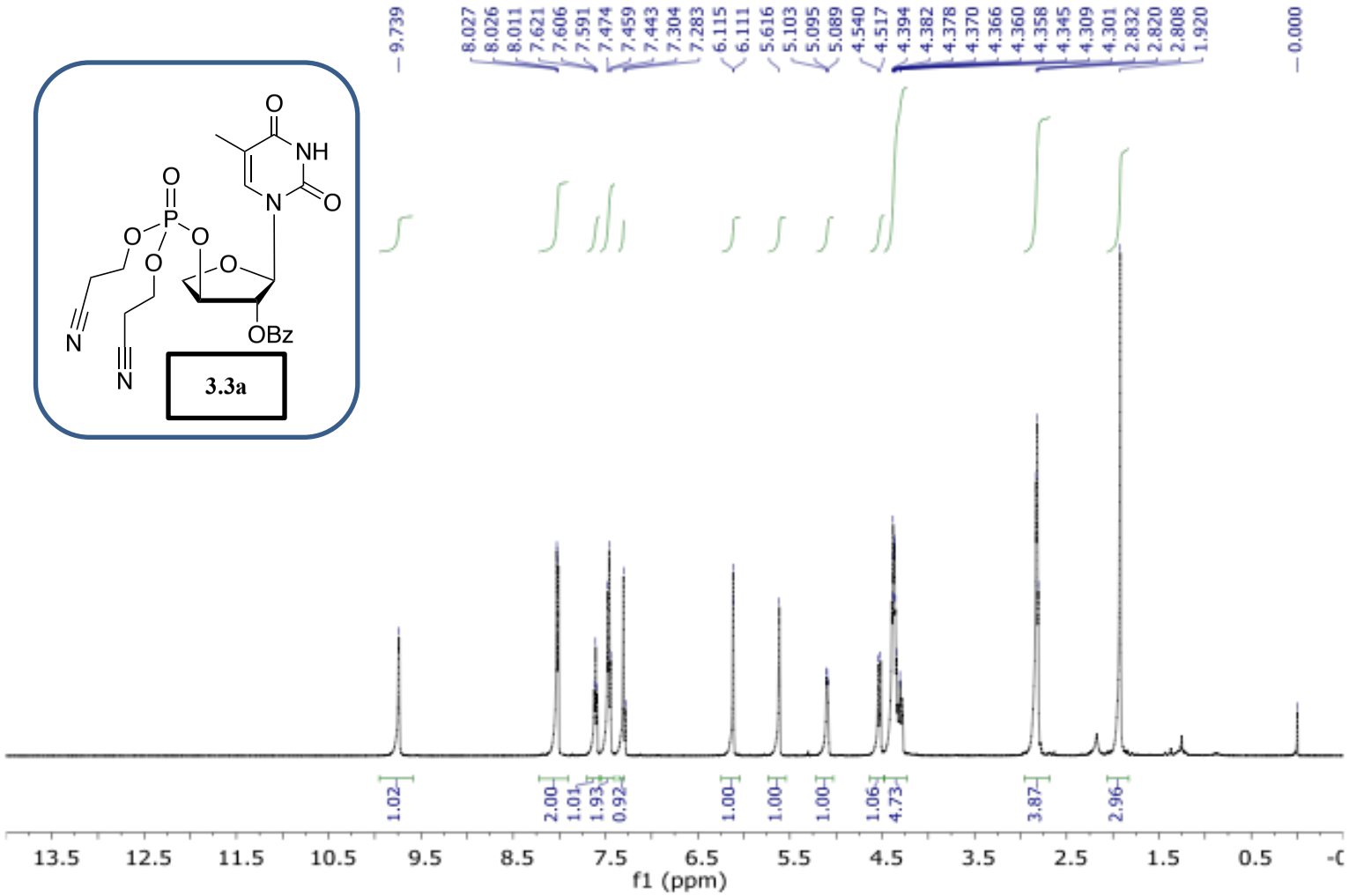
^{1}H NMR spectrum of compound **2.10b** (400 MHz, $CDCl_3$)



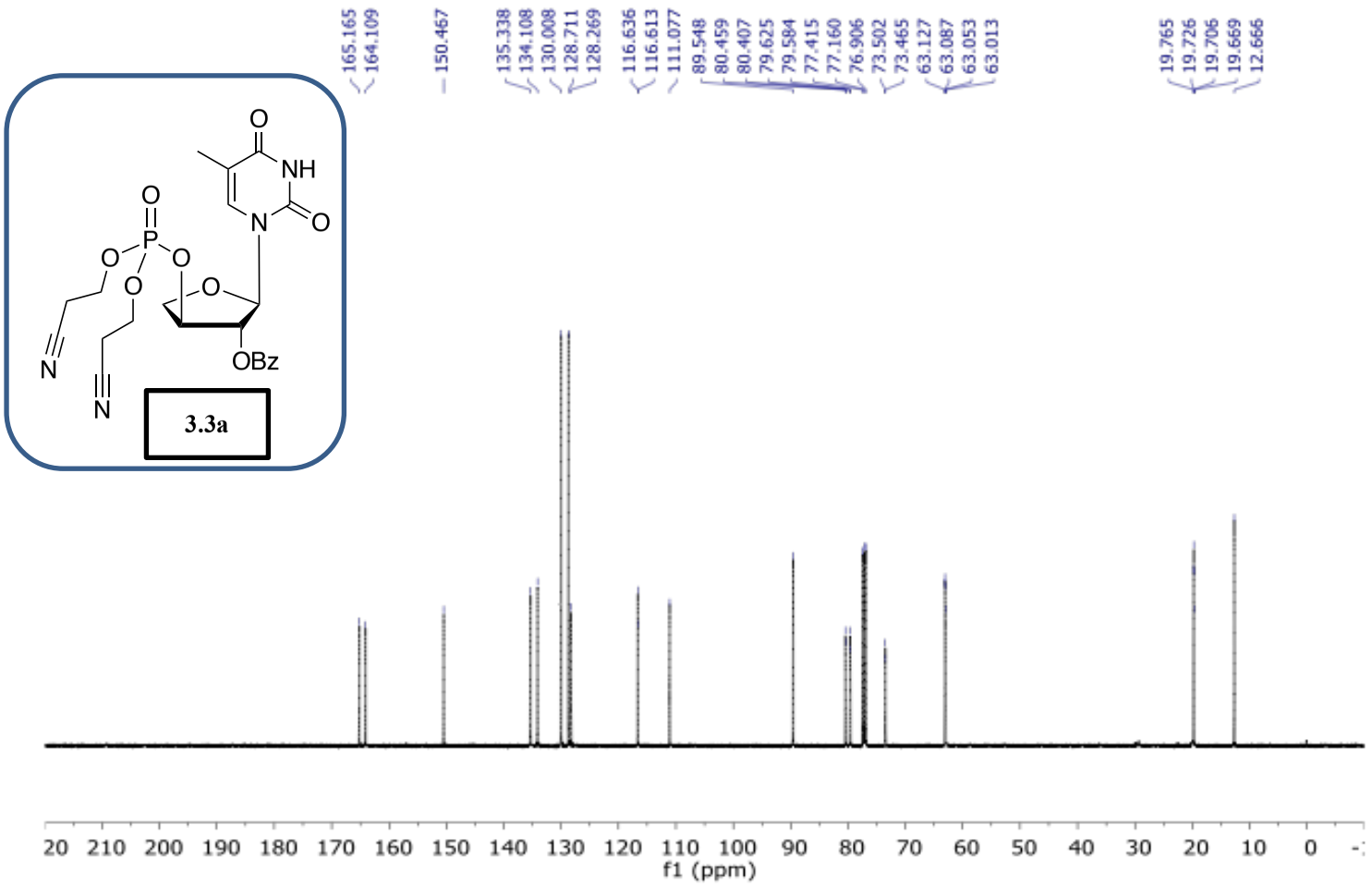
^{1}H NMR spectrum of compound **2.10c** (400 MHz,DMSO-0d_6)



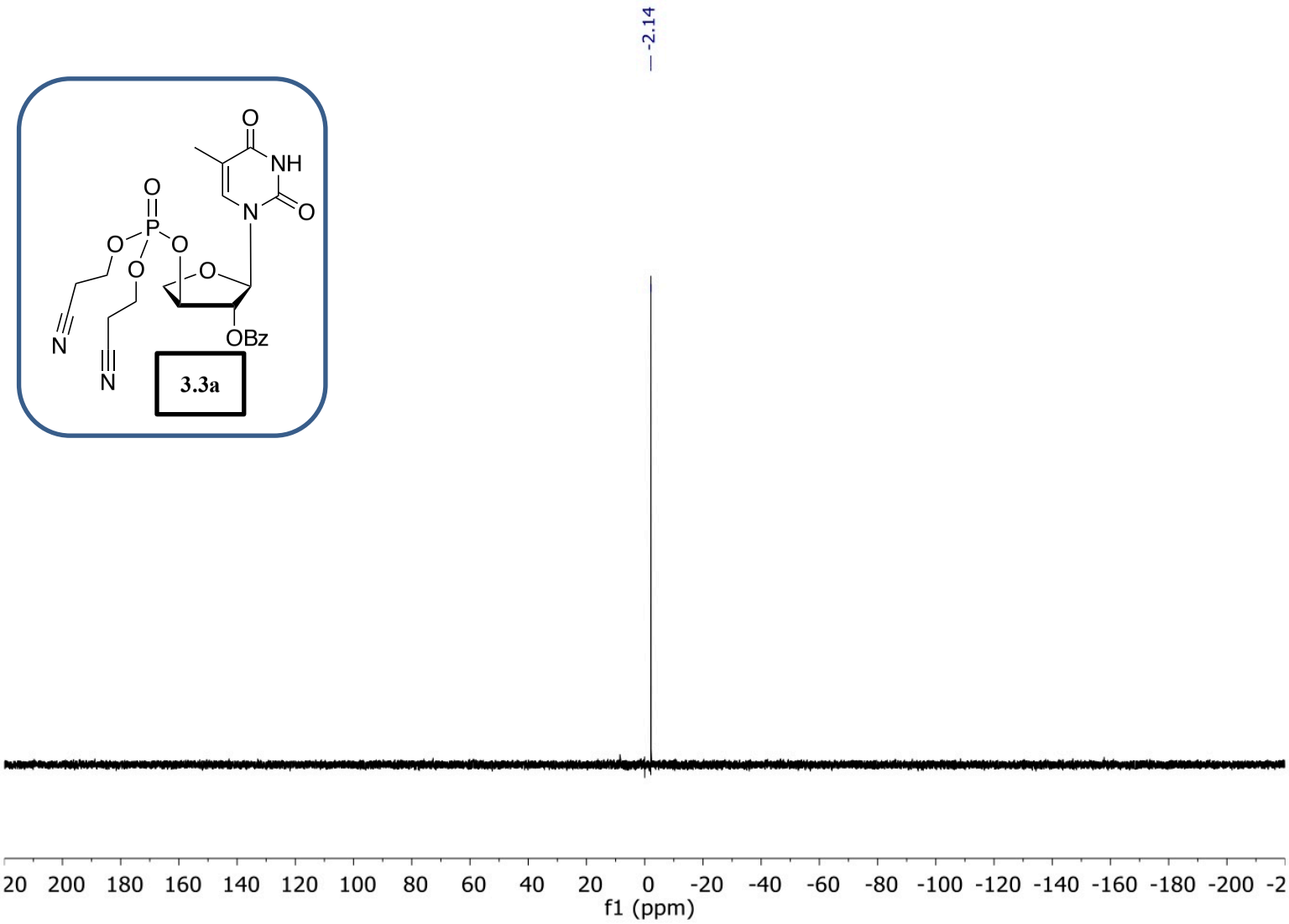
^{1}H NMR spectrum of compound **3.10d** (400 MHz,DMSO-0d_6)



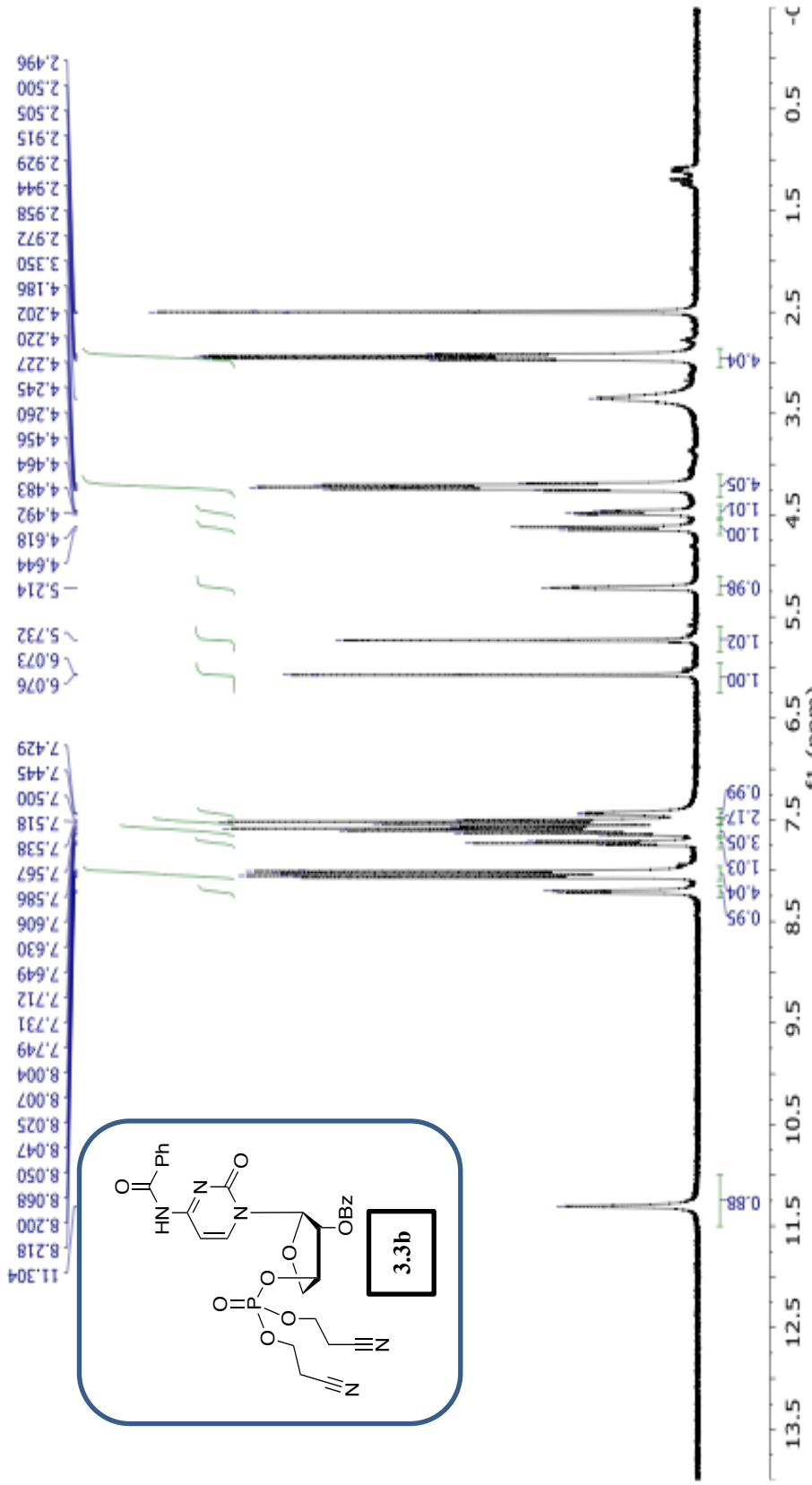
^{1}H NMR spectrum of compound **3.3a** (500 MHz,$CDCl_3$)



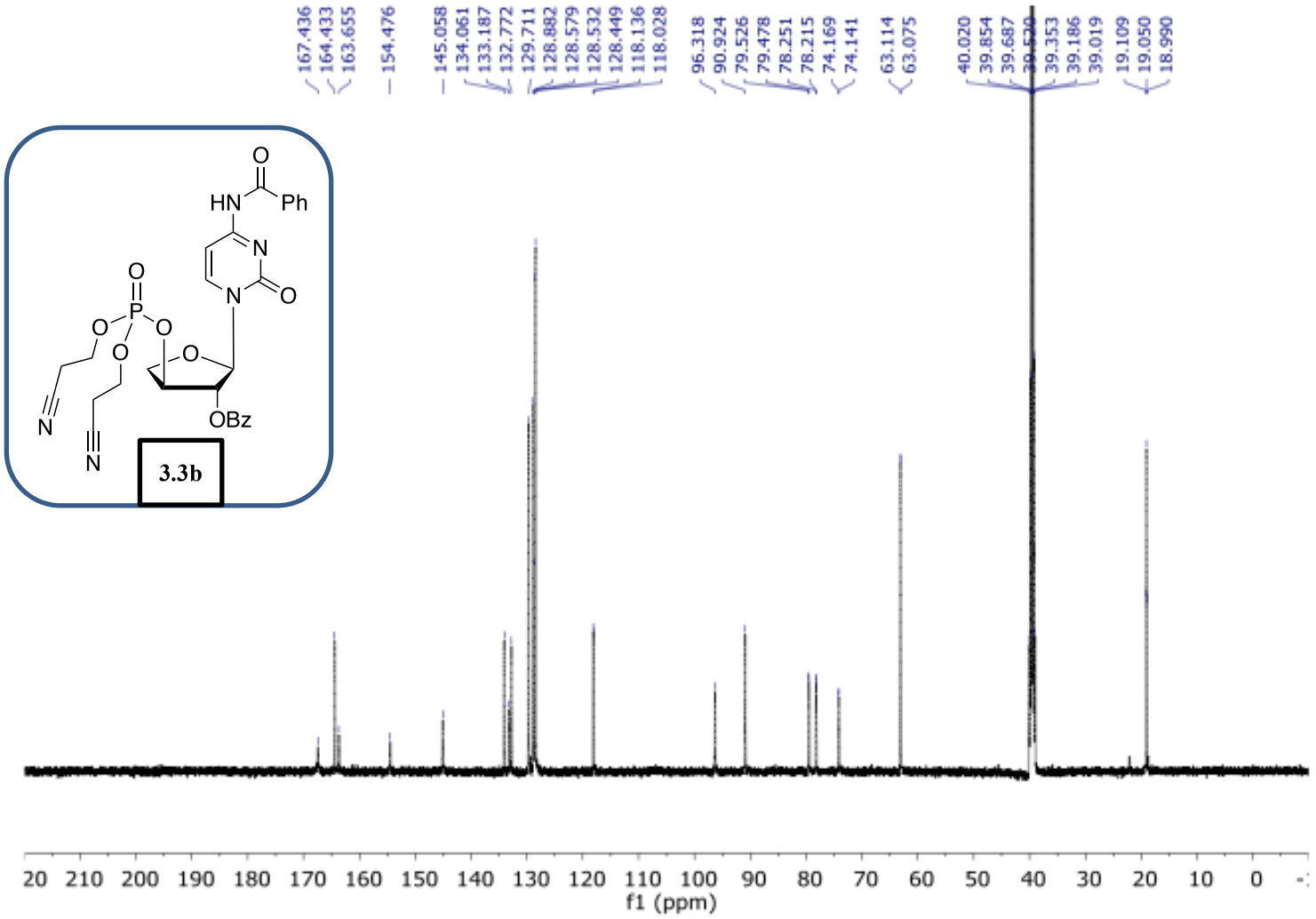
^{13}C NMR spectrum of compound **3.3a** (500 MHz,$CDCl_3$)



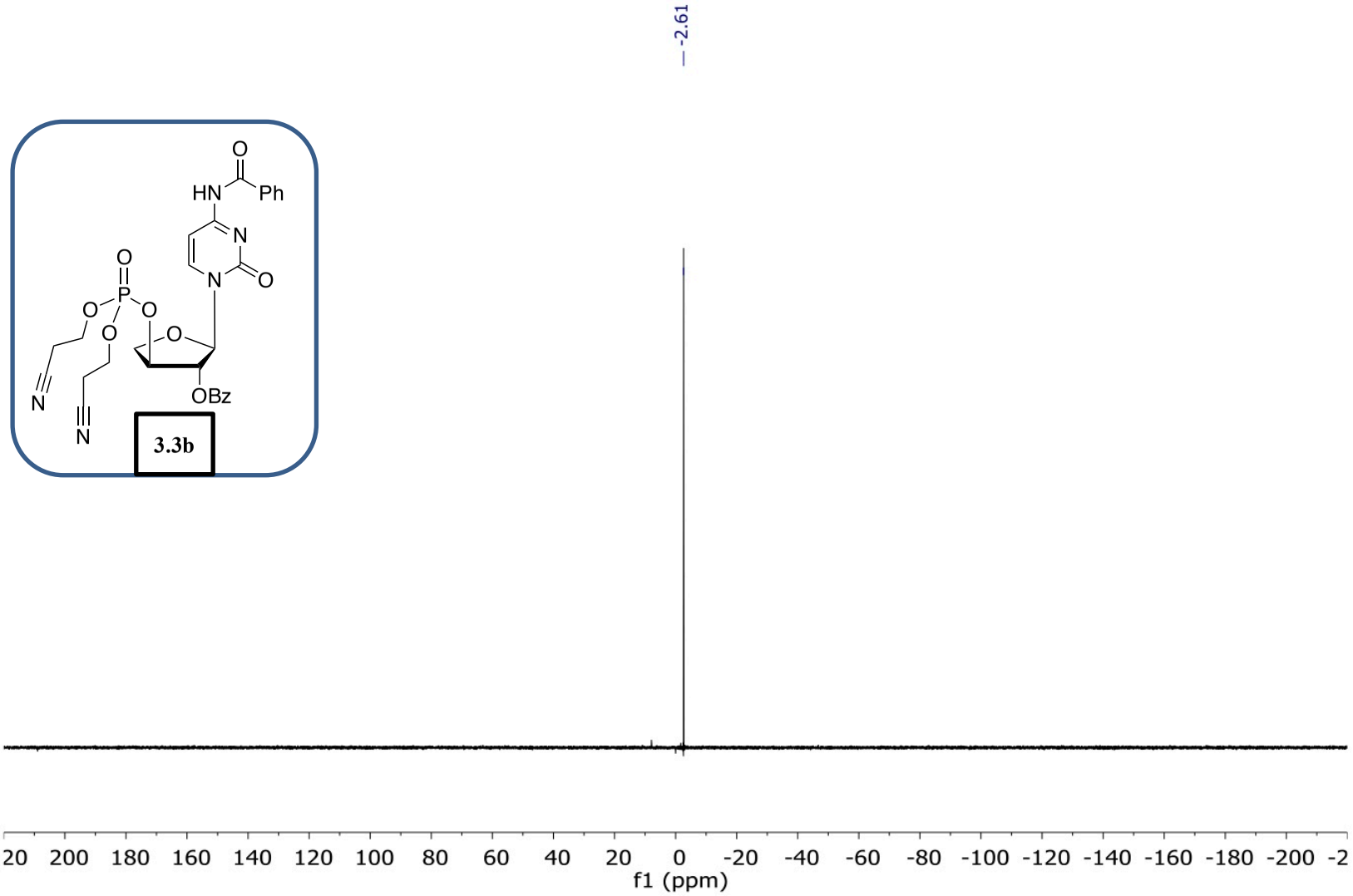
^{31}P NMR spectrum of compound **3.3a** (500 MHz,$CDCl_3$)



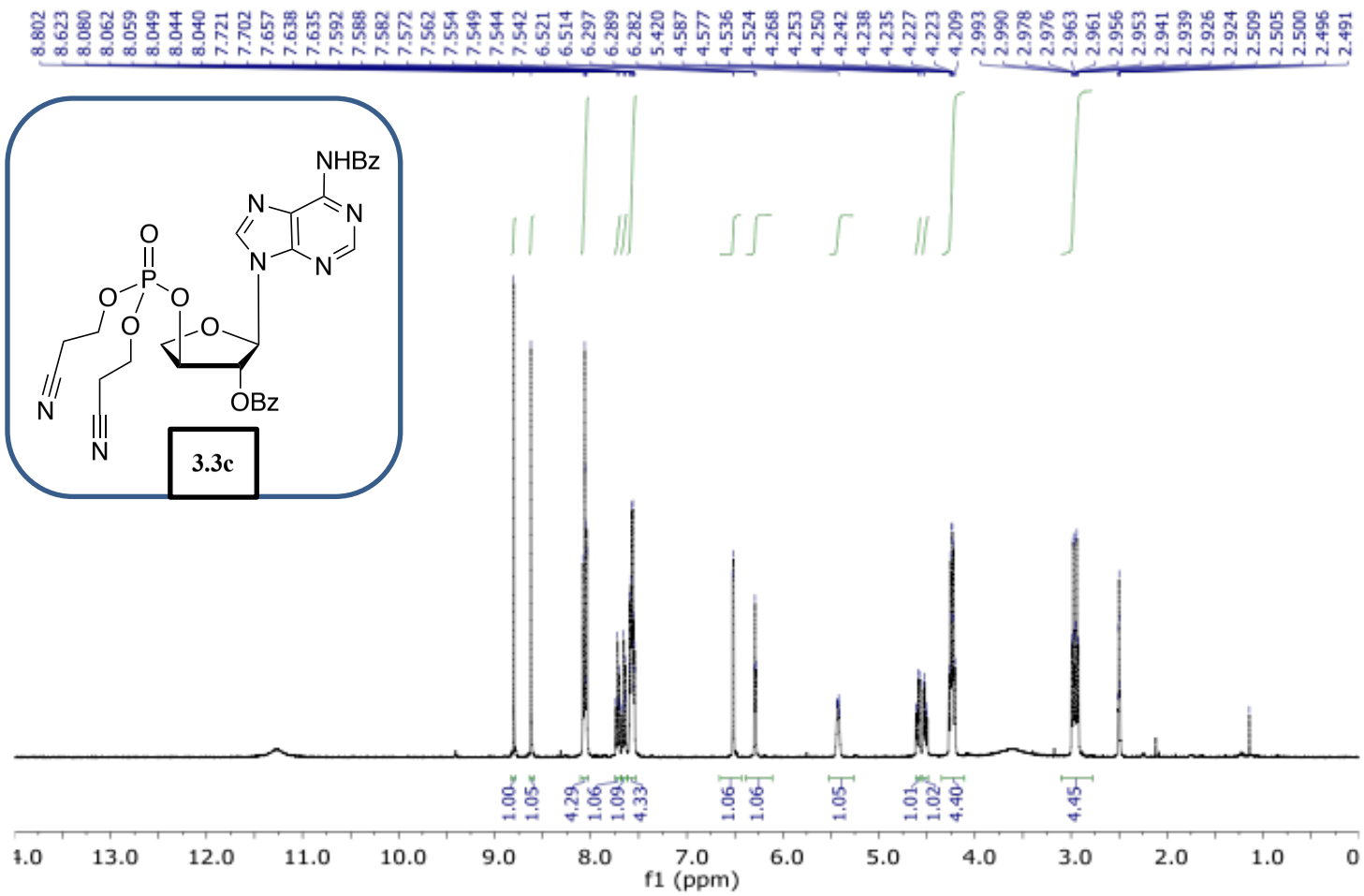
^{1}H NMR spectrum of compound **3.3b** (500 MHz,DMSO-d^{6})



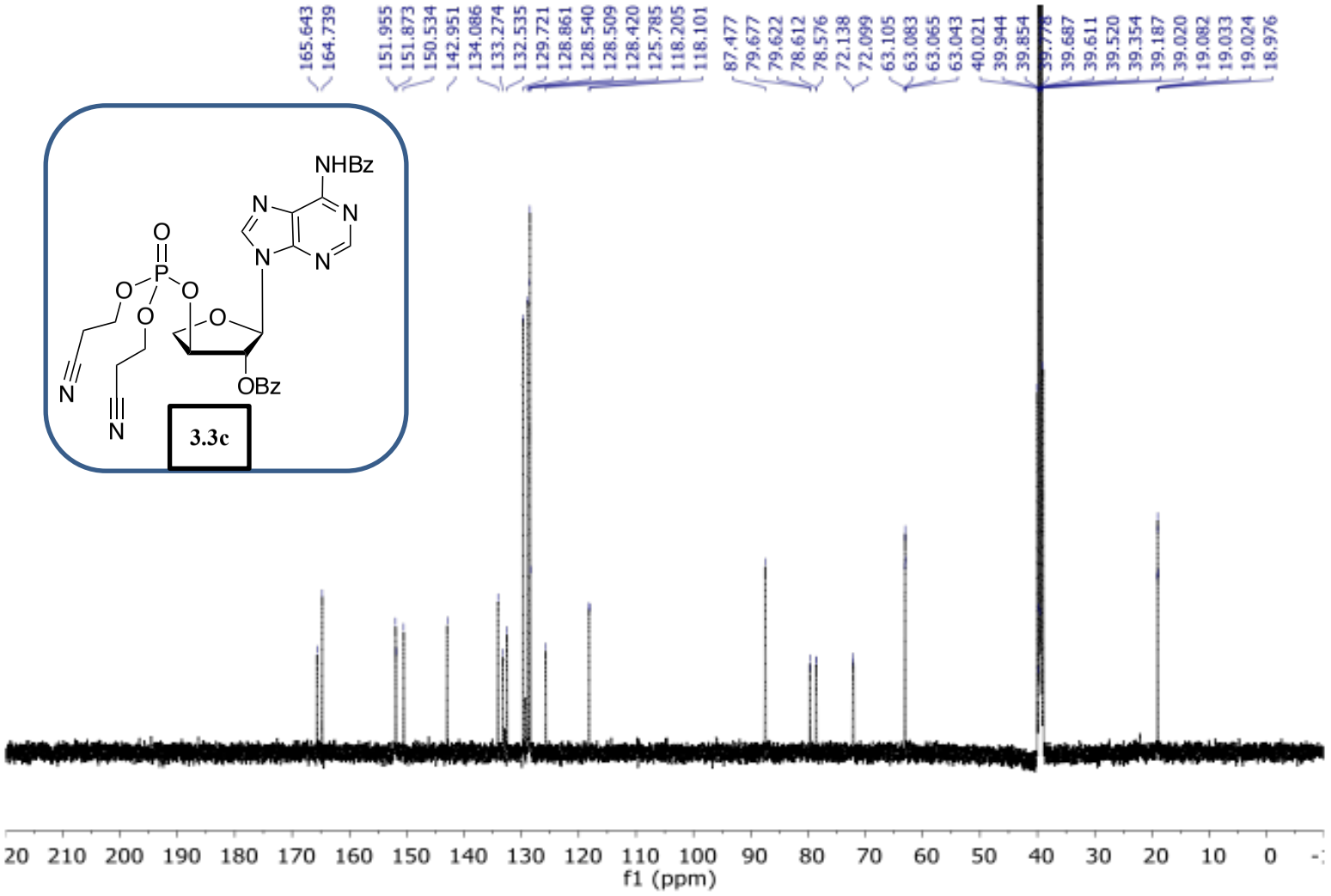
^{13}C NMR spectrum of compound **3.3b** (500 MHz,DMSO-*d*^{6})



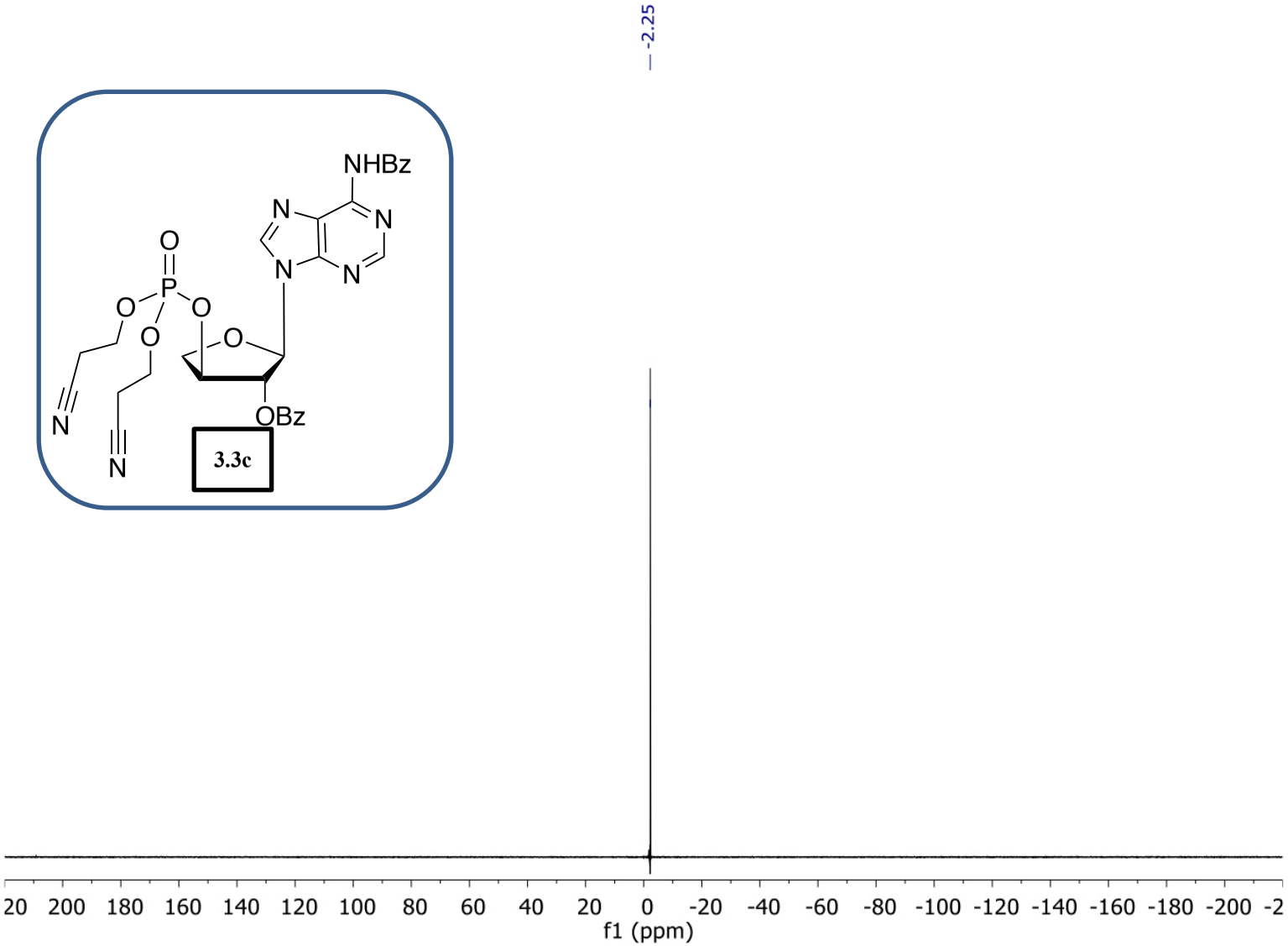
^{31}P NMR spectrum of compound **3.3b** (500 MHz,DMSO-d^6)



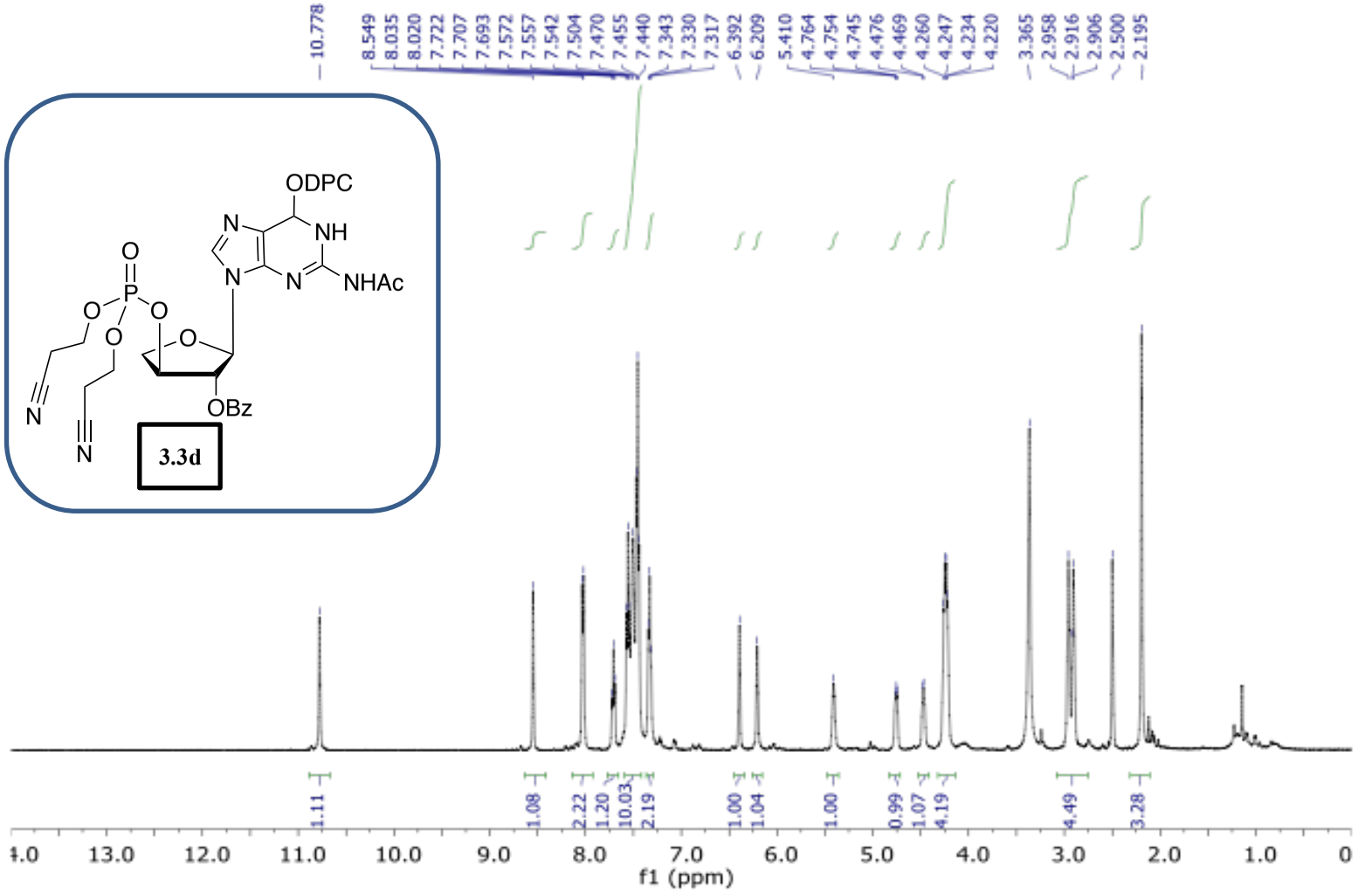
^{1}H NMR spectrum of compound **3.3c** (500 MHz,DMSO-d^{6})



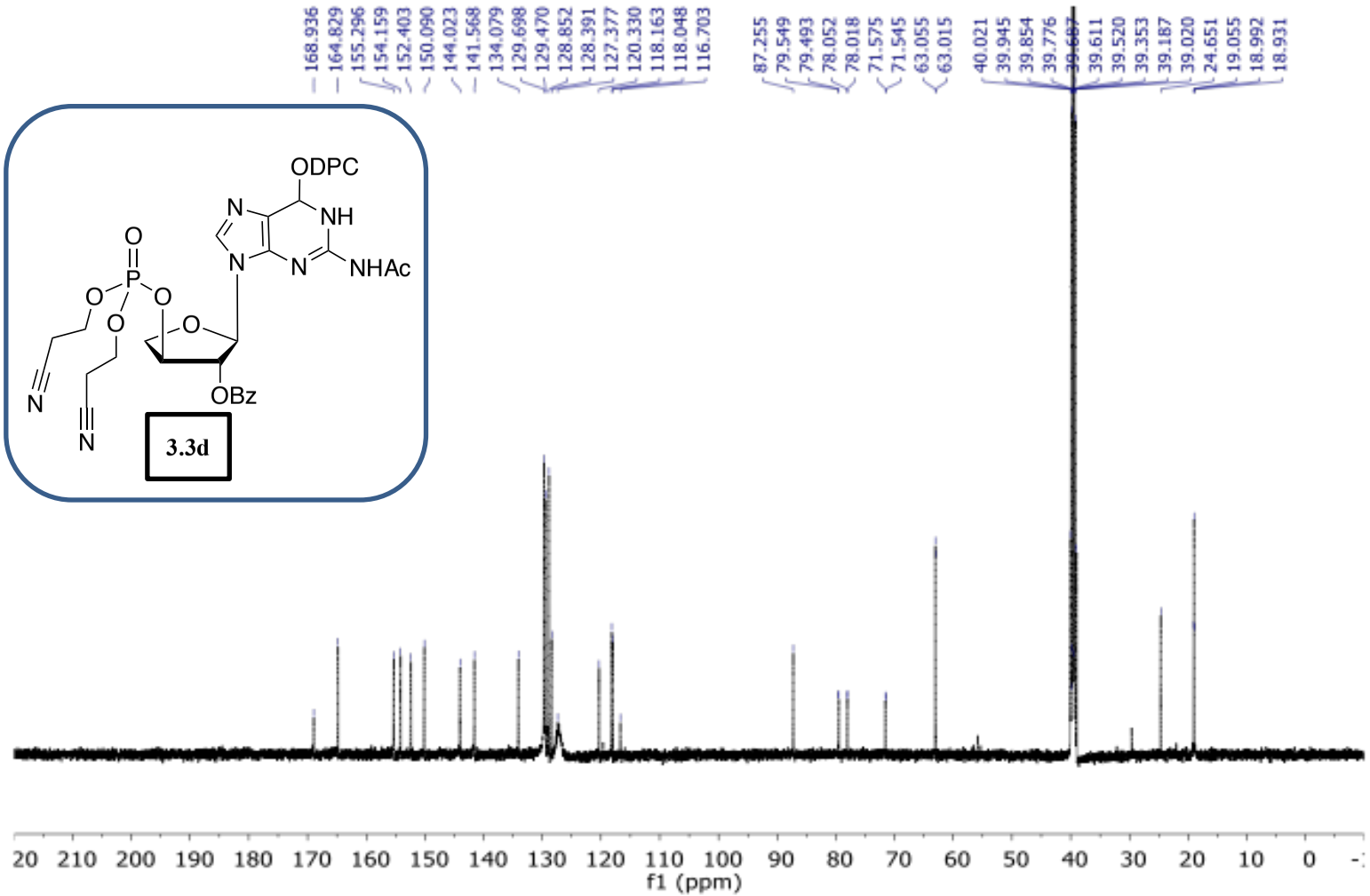
^{13}C NMR spectrum of compound **3.3c** (500 MHz,DMSO-d^6)



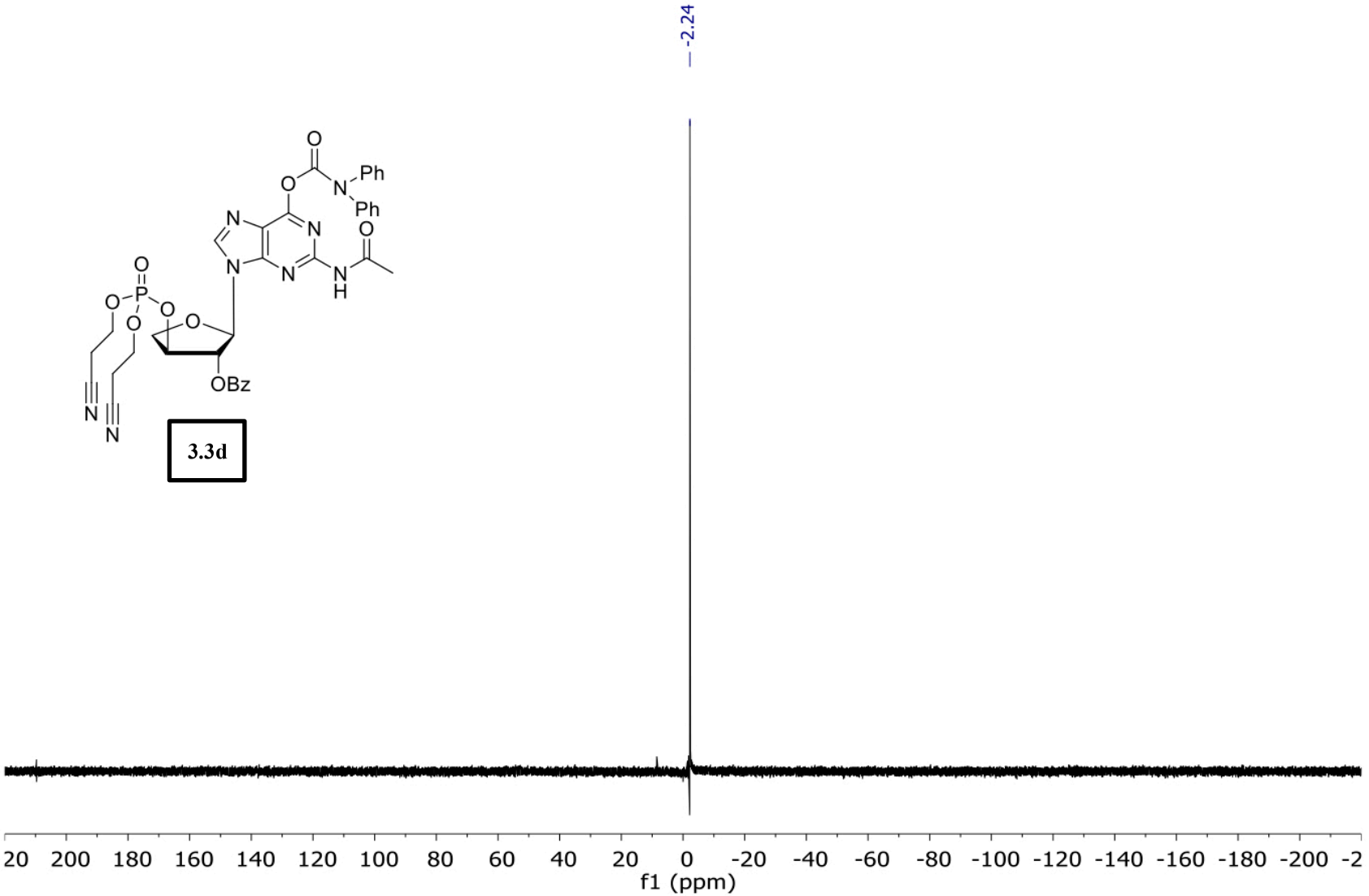
[31] P NMR spectrum of compound **3.3c** (500 MHz,DMSO-d^6)



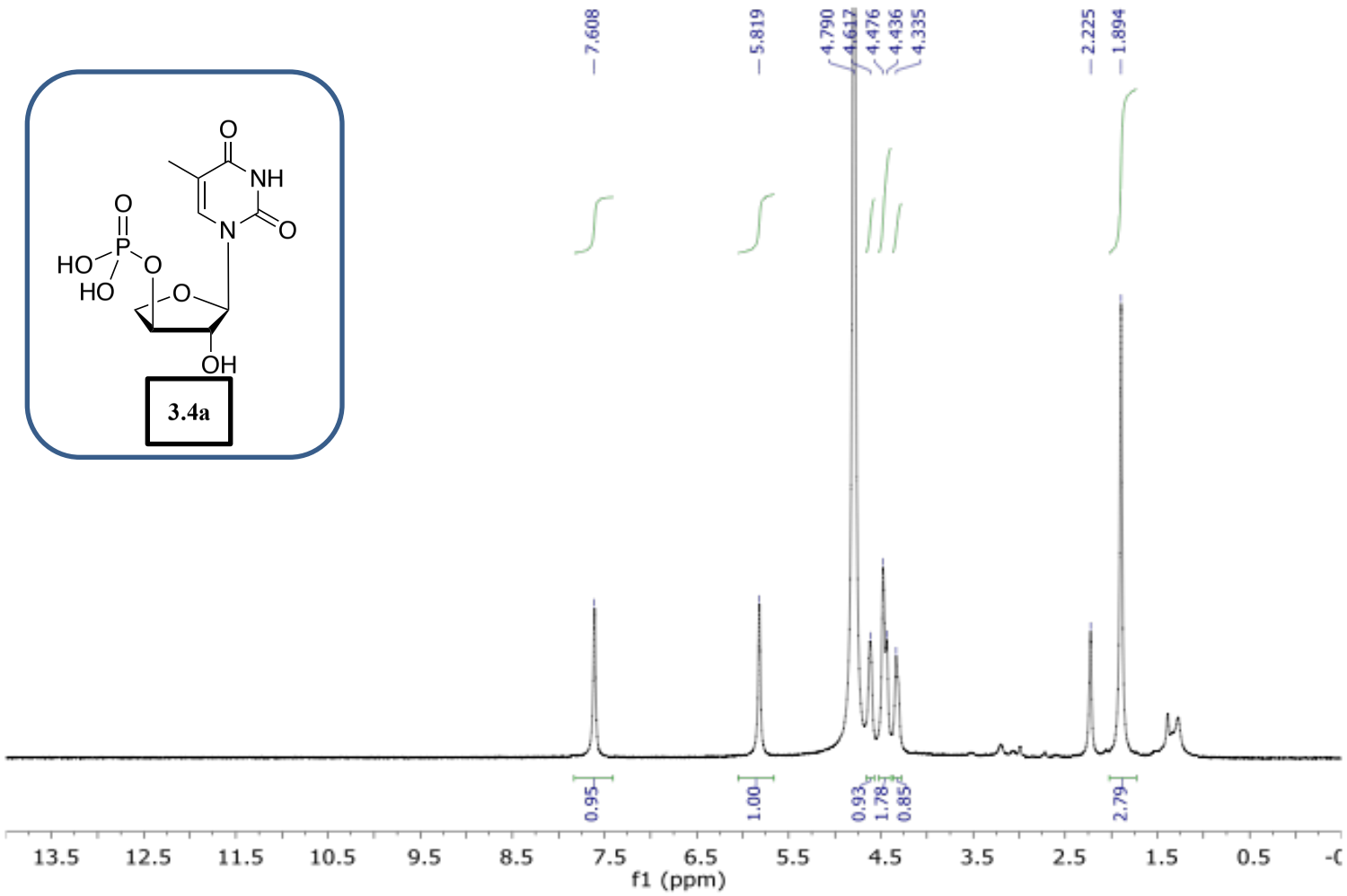
^{1}H NMR spectrum of compound **3.3d** (500 MHz,DMSO-d^{6})



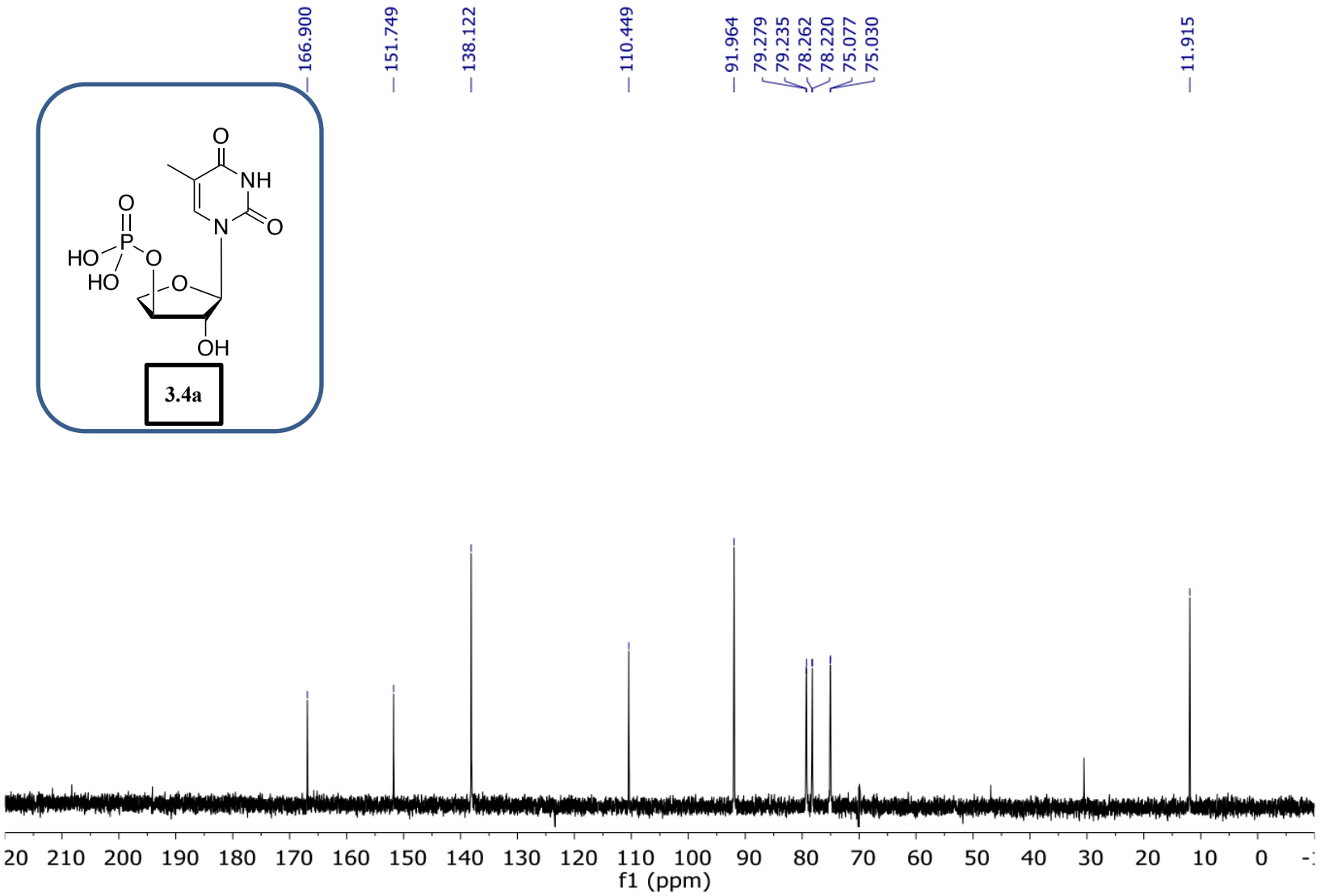
^{13}C NMR spectrum of compound **3.3d** (500 MHz,DMSO-d^6)



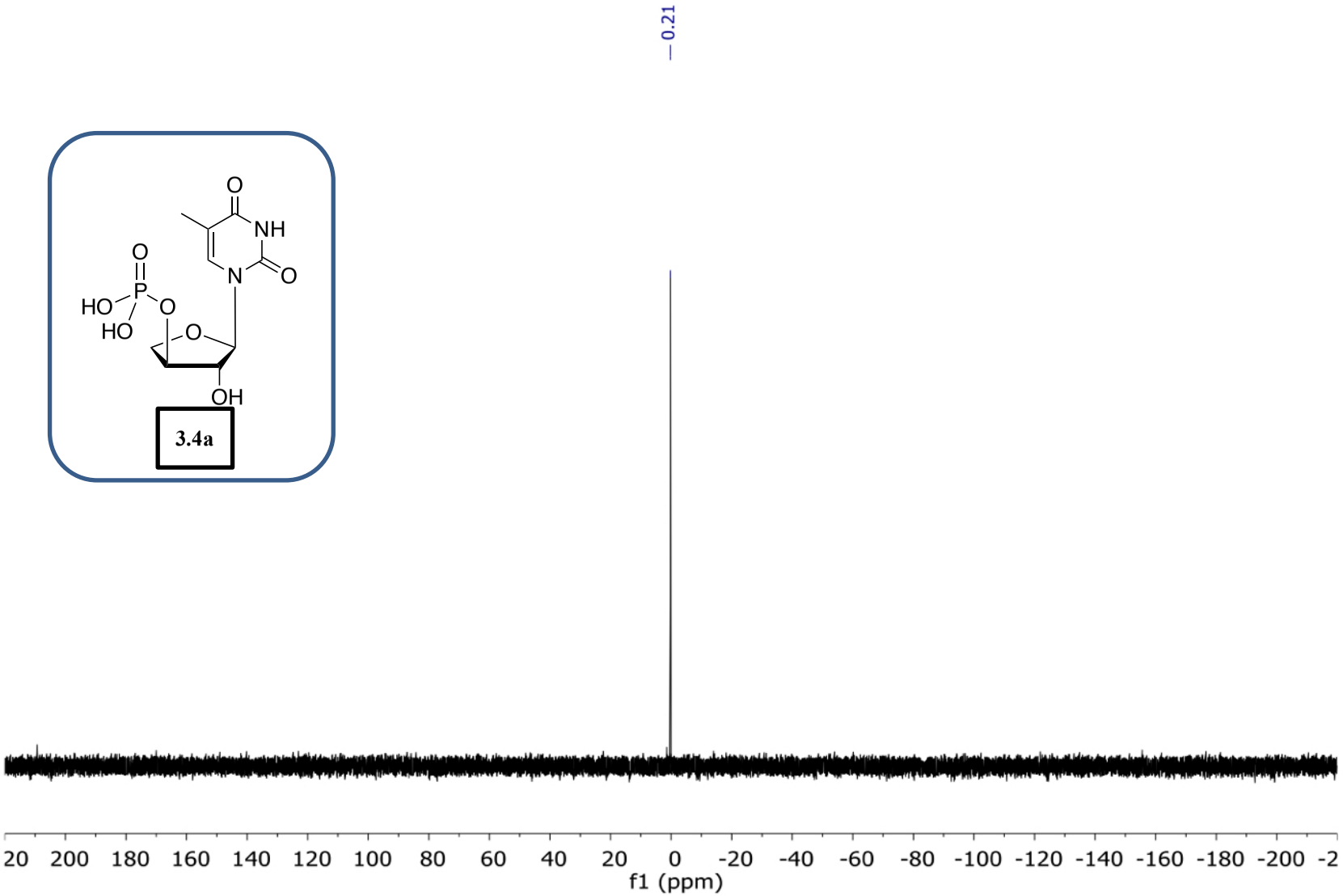
^{31}P NMR spectrum of compound **3.3d** (500 MHz,DMSO-d^6)



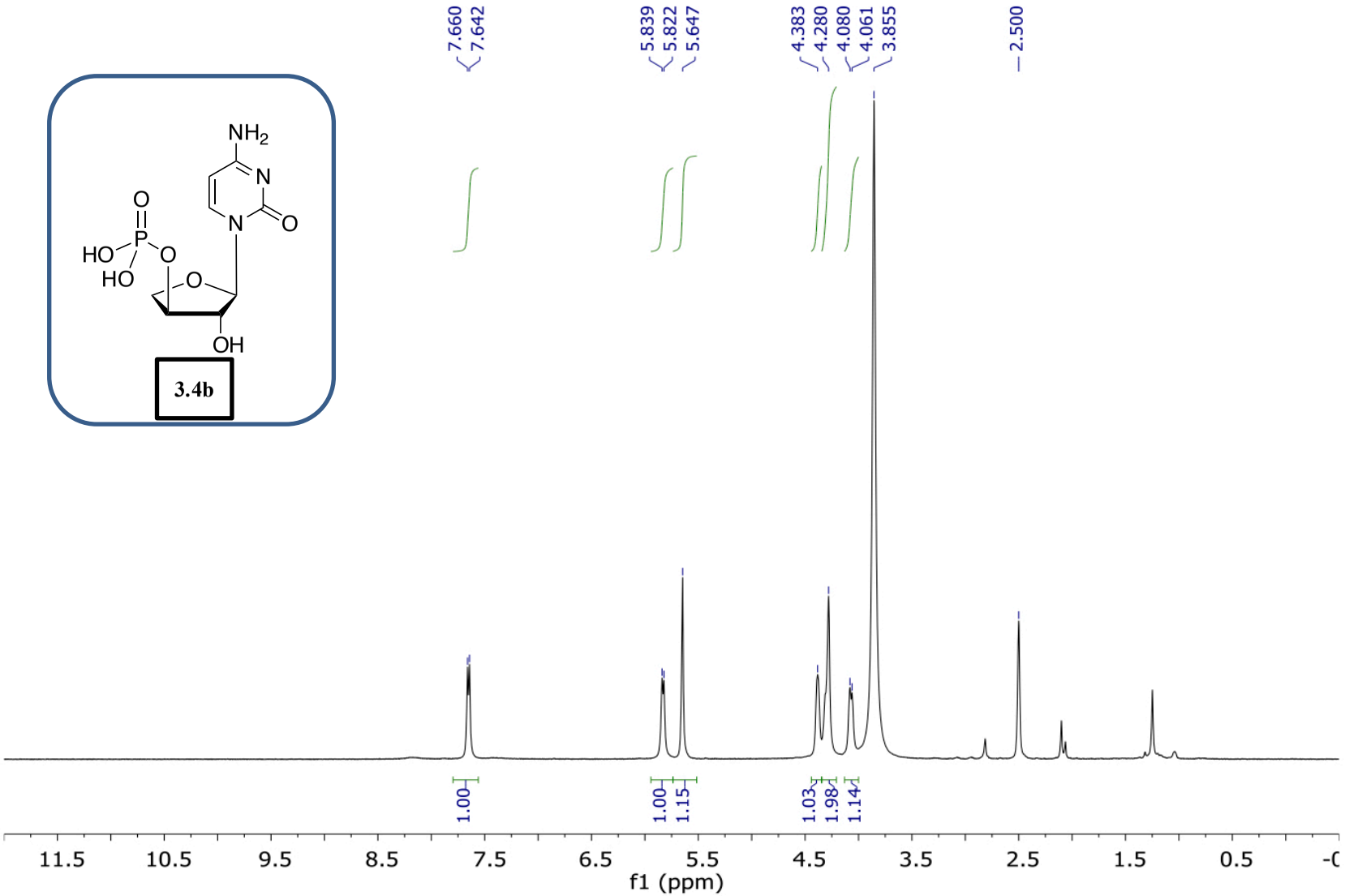
^{1}H NMR spectrum of compound **3.4a** (500 MHz,DMSO-d^{6})



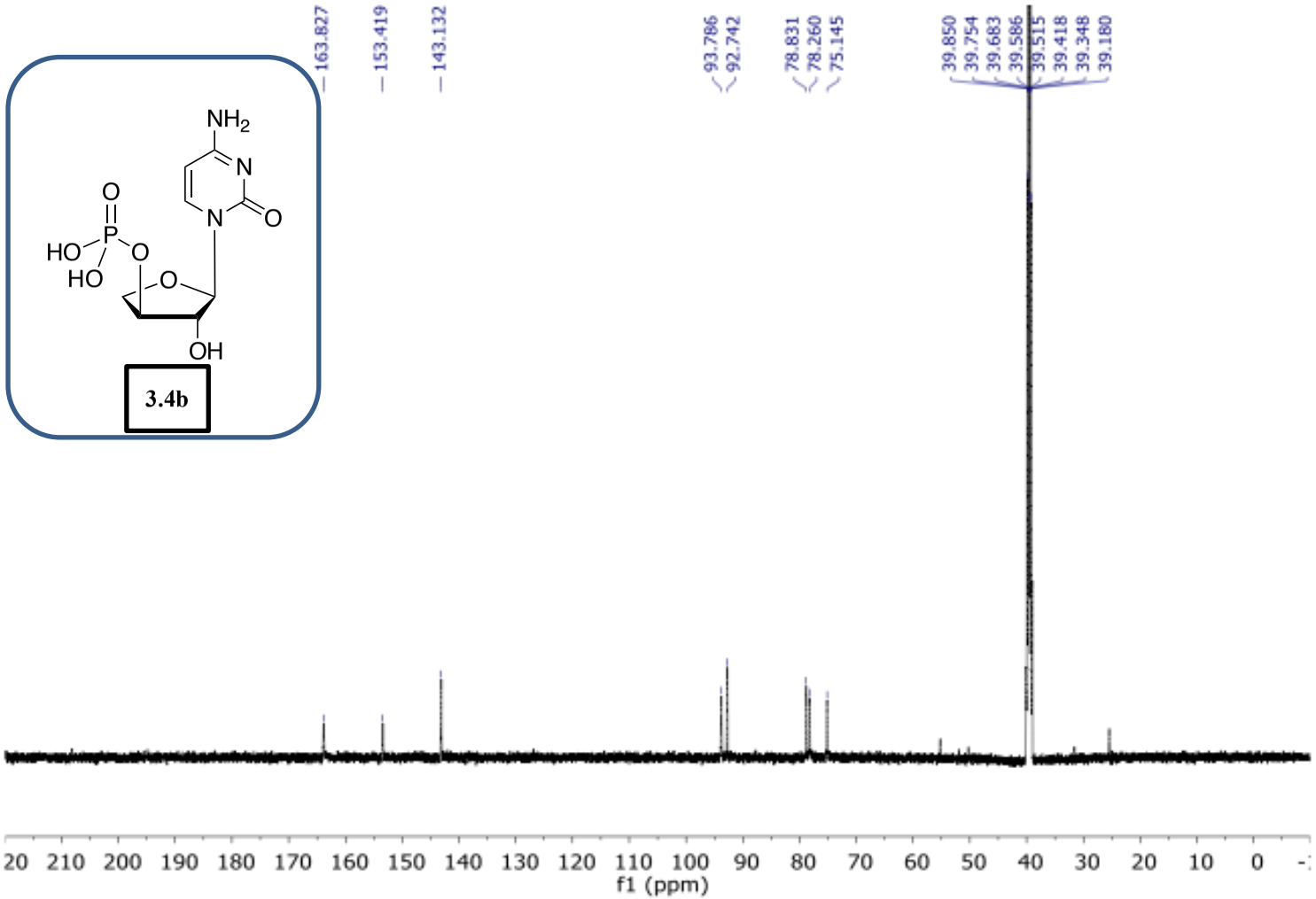
^{13}C NMR spectrum of compound **3.4a** (500 MHz,DMSO-d^6)



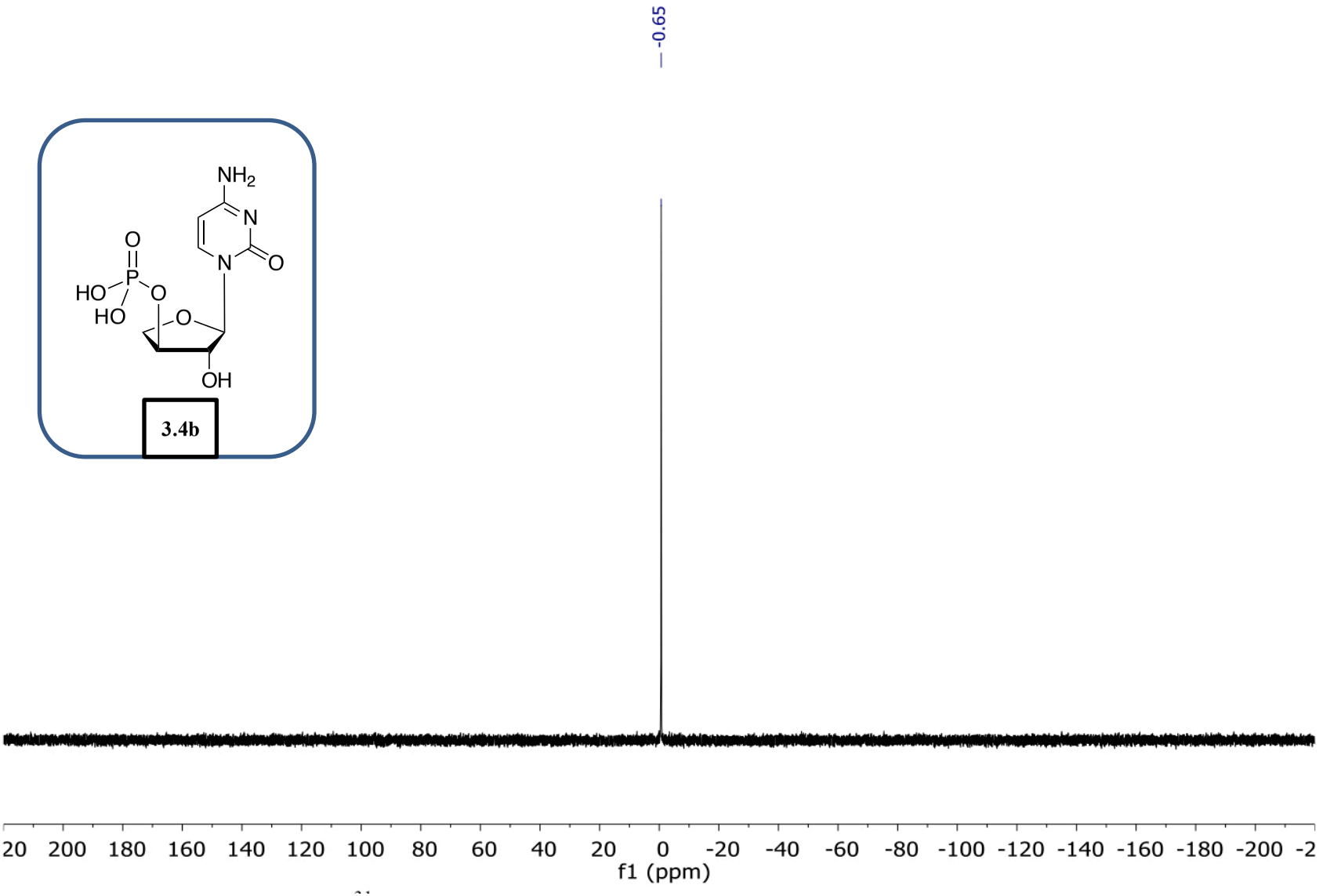
^{31}P NMR spectrum of compound **3.4a** (500 MHz,DMSO-d^6)



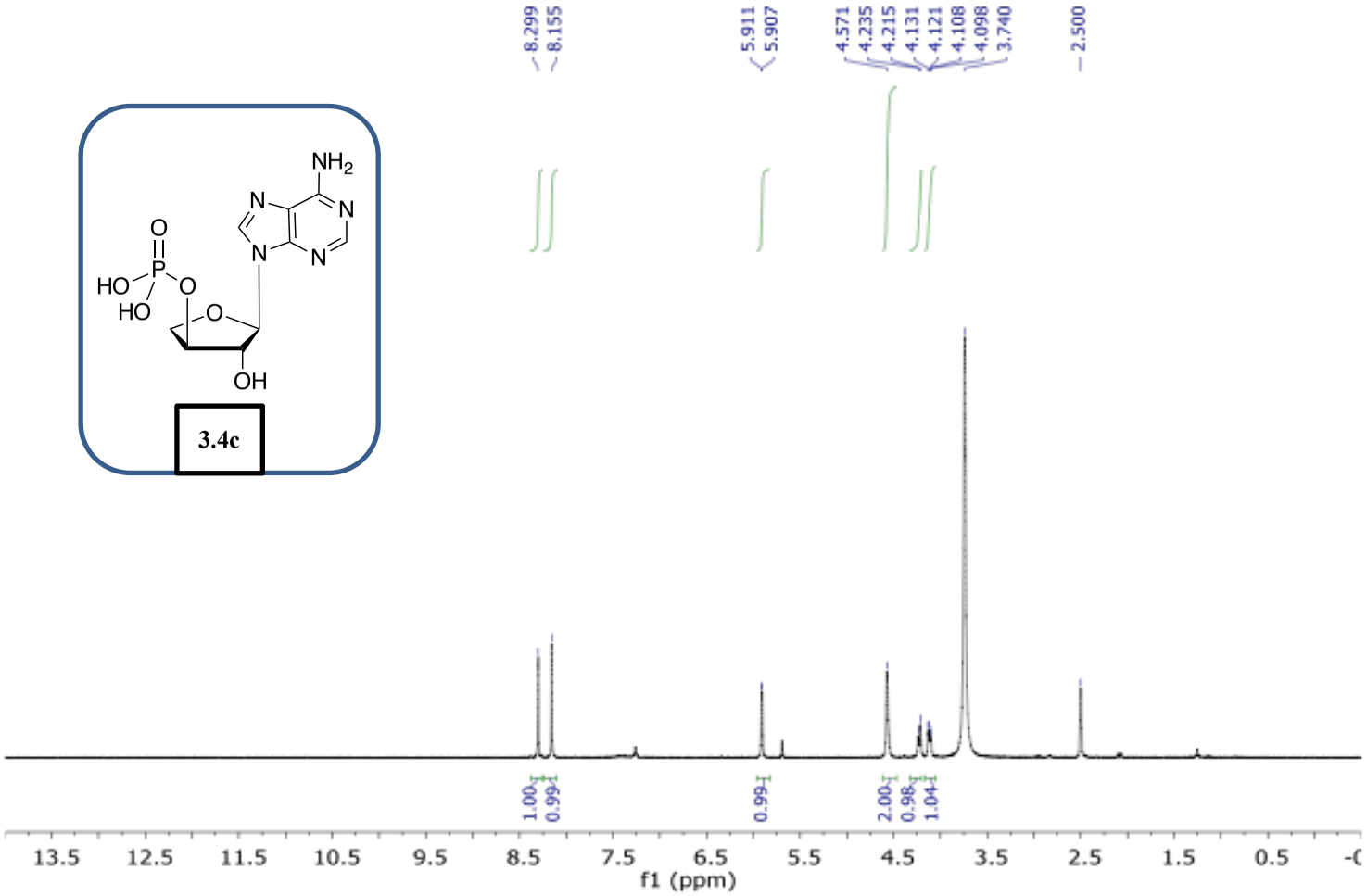
^{1}H NMR spectrum of compound **3.4b** (500 MHz,DMSO-d^{6})



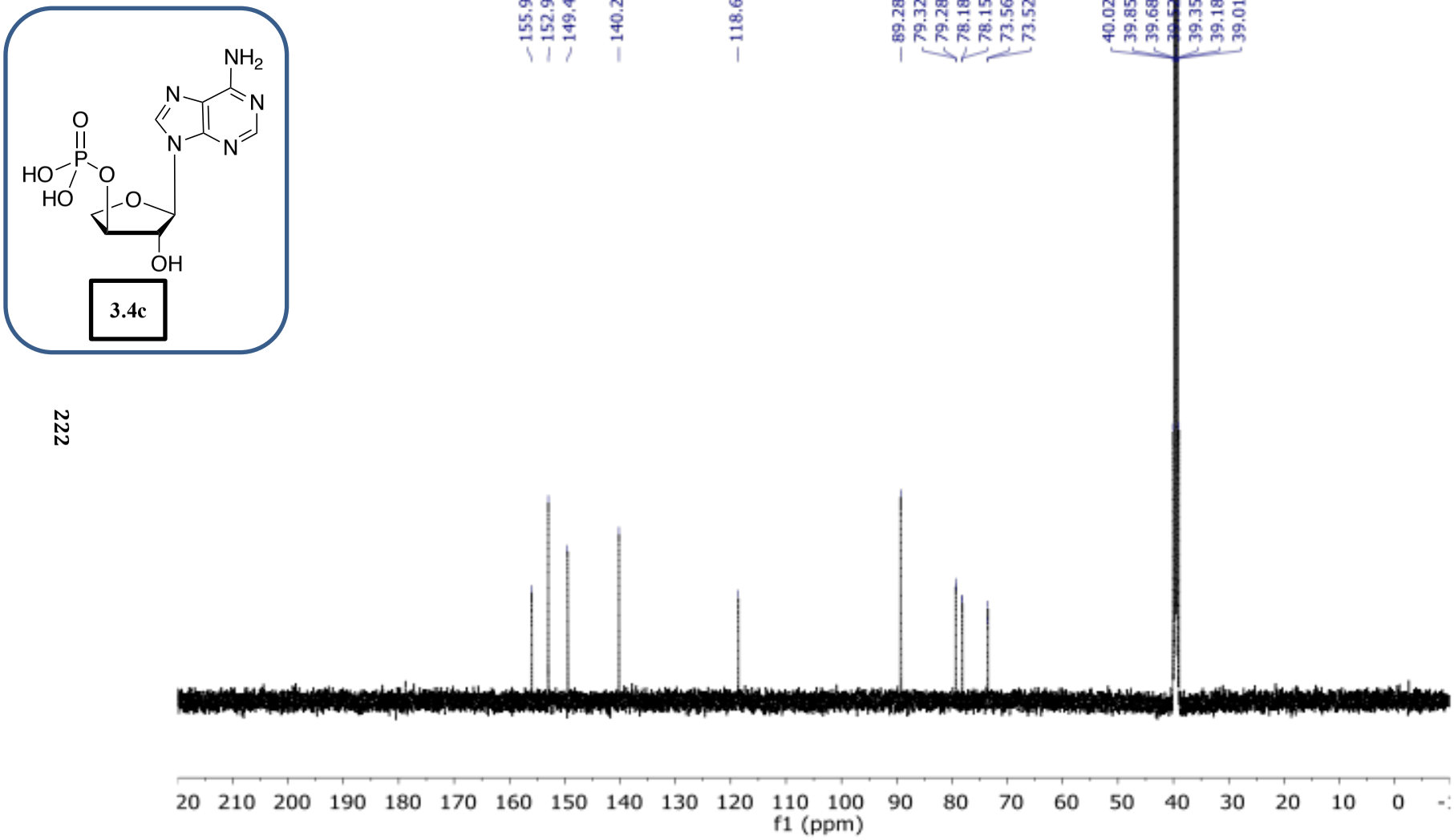
^{13}C NMR spectrum of compound **3.4b** (500 MHz,DMSO-d^6)



^{31}P NMR spectrum of compound **3.4b** (500 MHz,DMSO-*d*^{6})

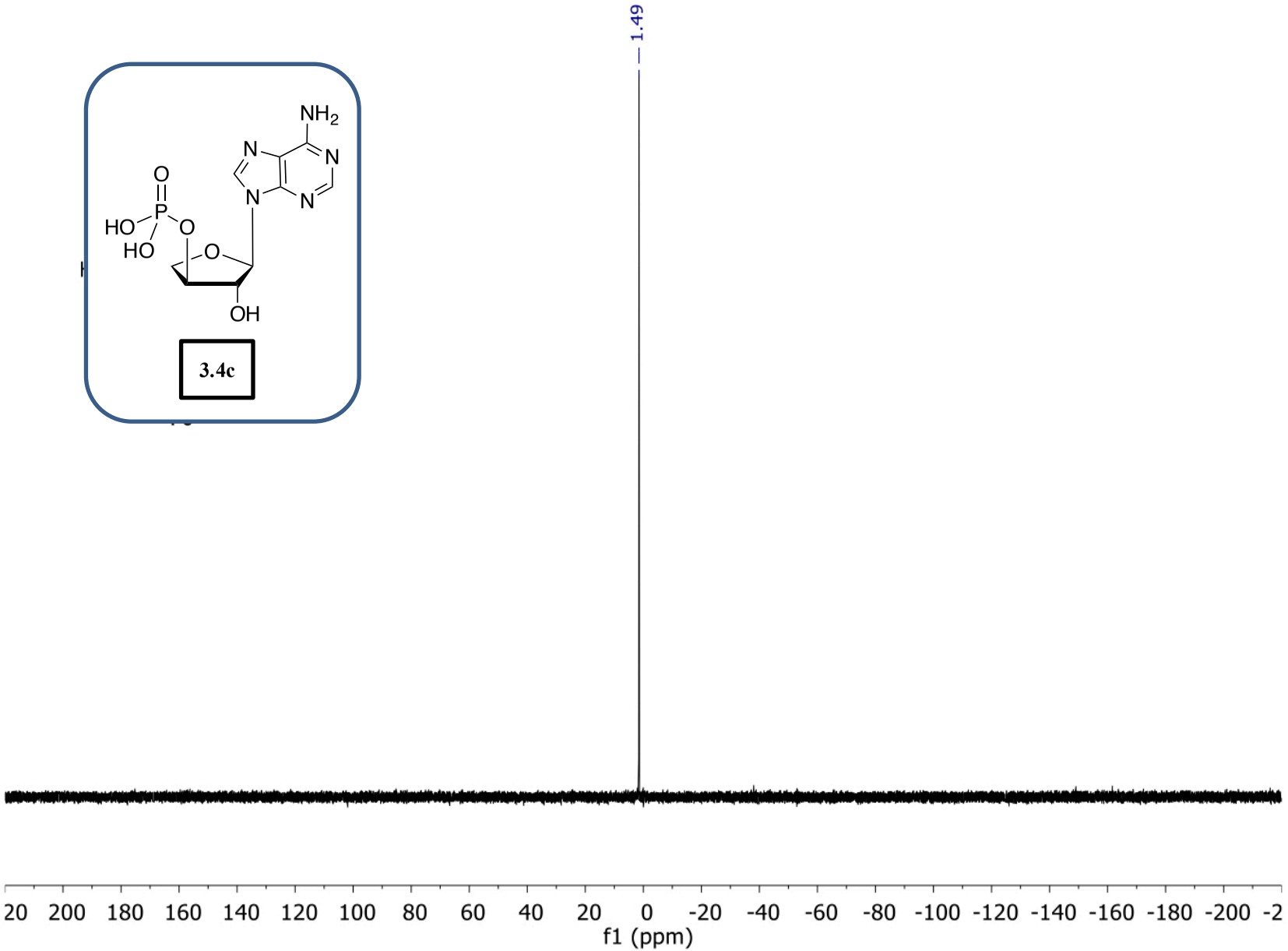


^{1}H NMR spectrum of compound **3.4c** (500 MHz,DMSO-d^{6})

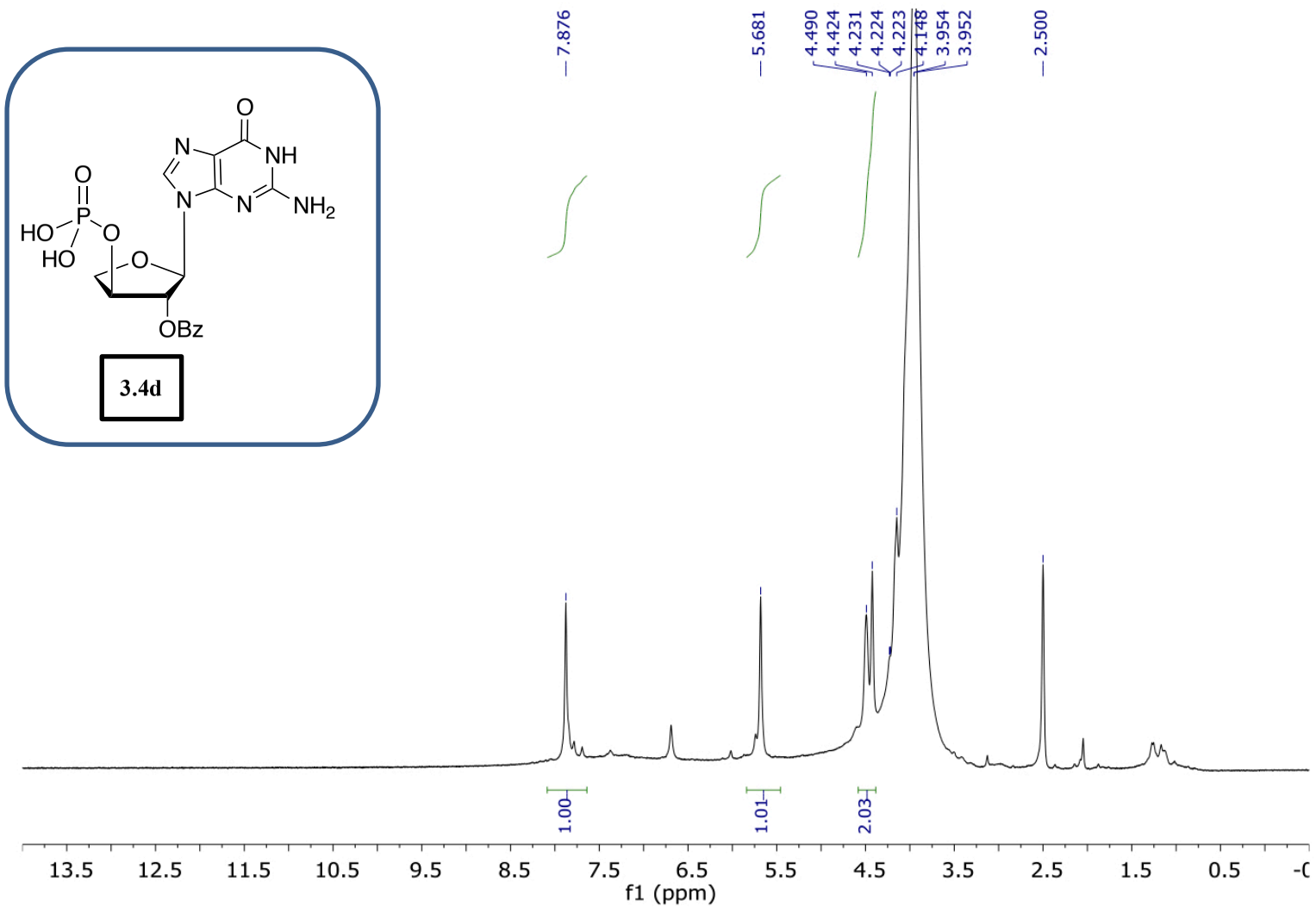


222

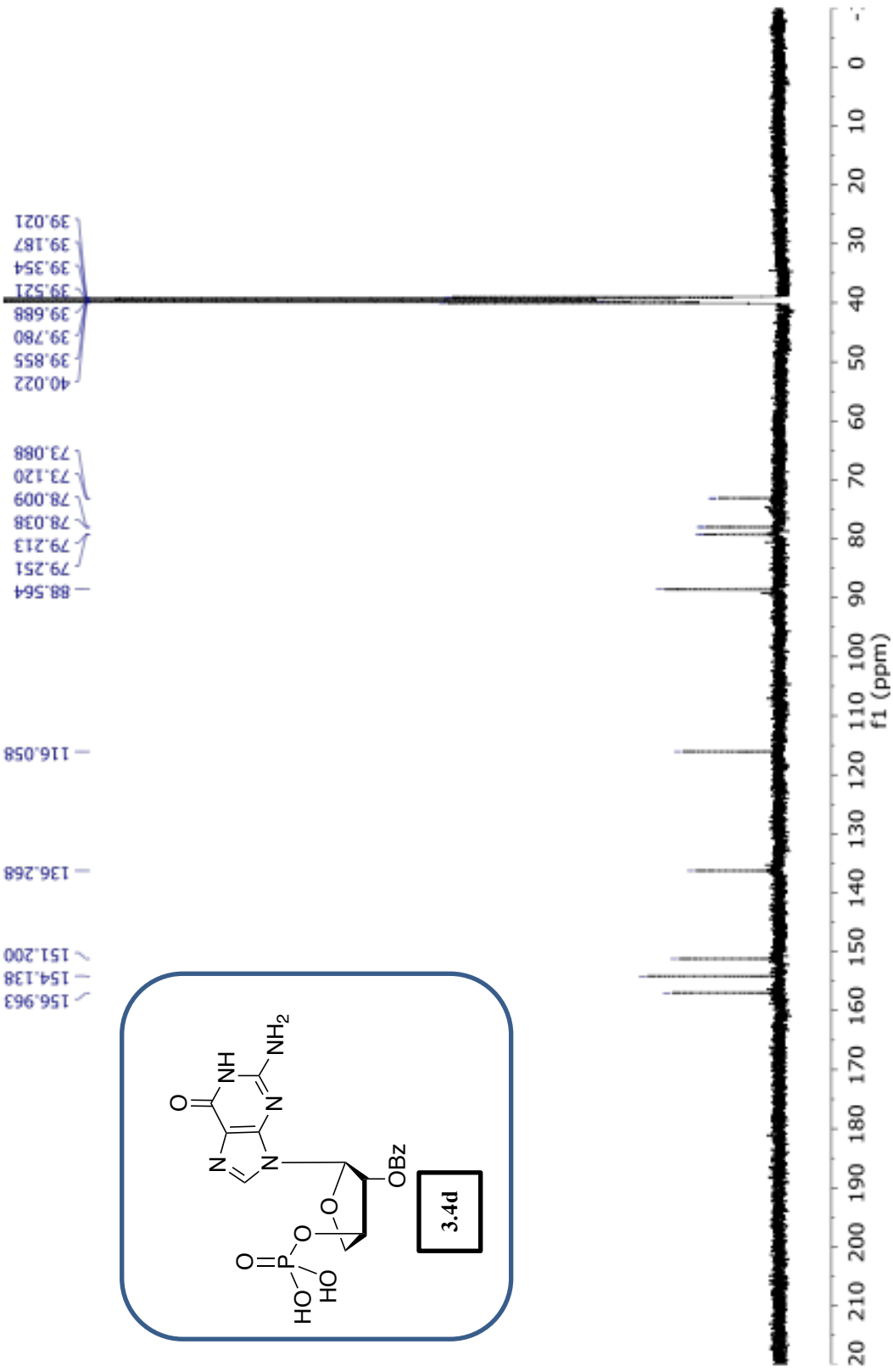
^{13}C NMR spectrum of compound **3.4c** (500 MHz,DMSO-d^6)



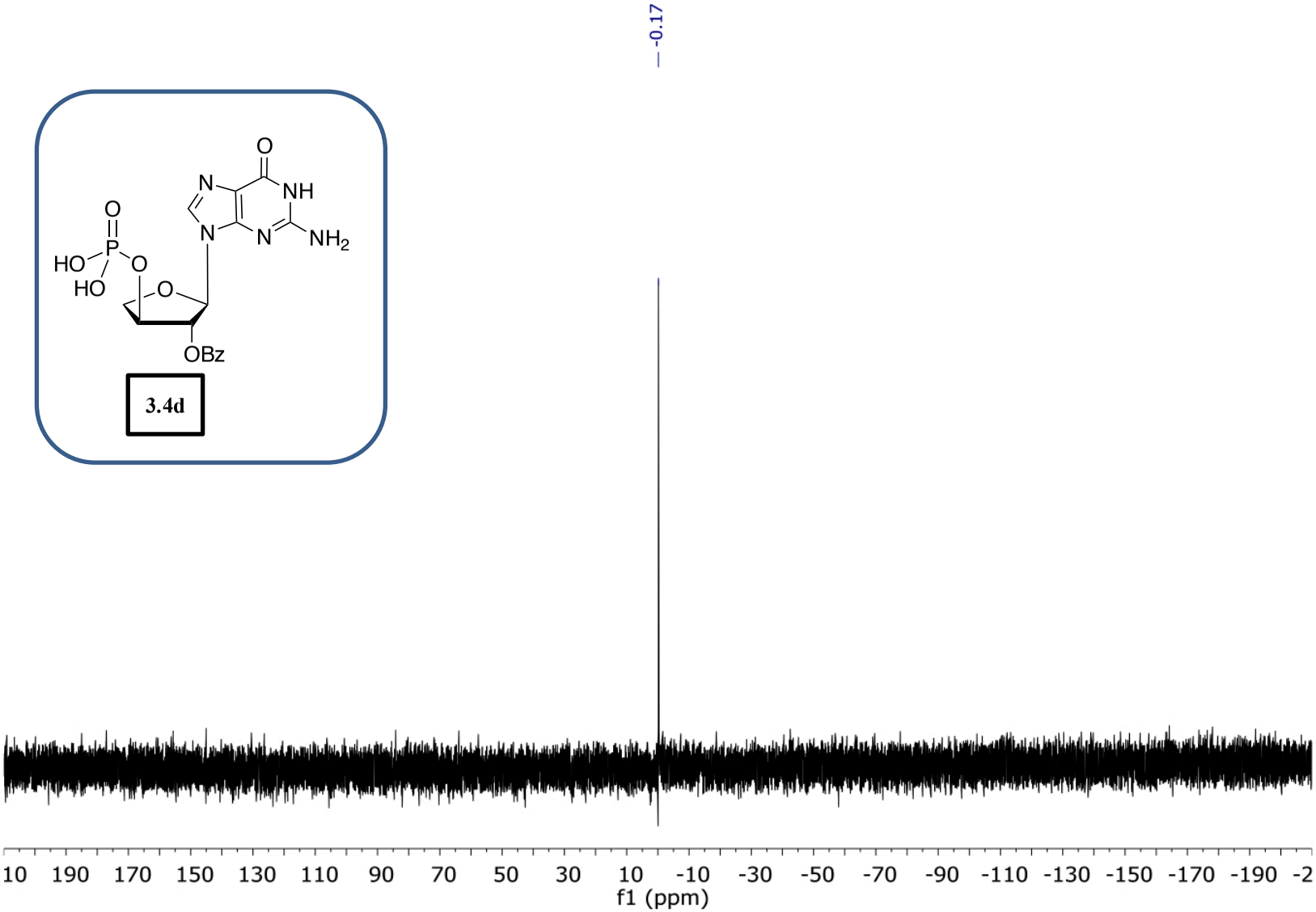
^{31}P NMR spectrum of compound **3.4c** (500 MHz,DMSO-d^6)



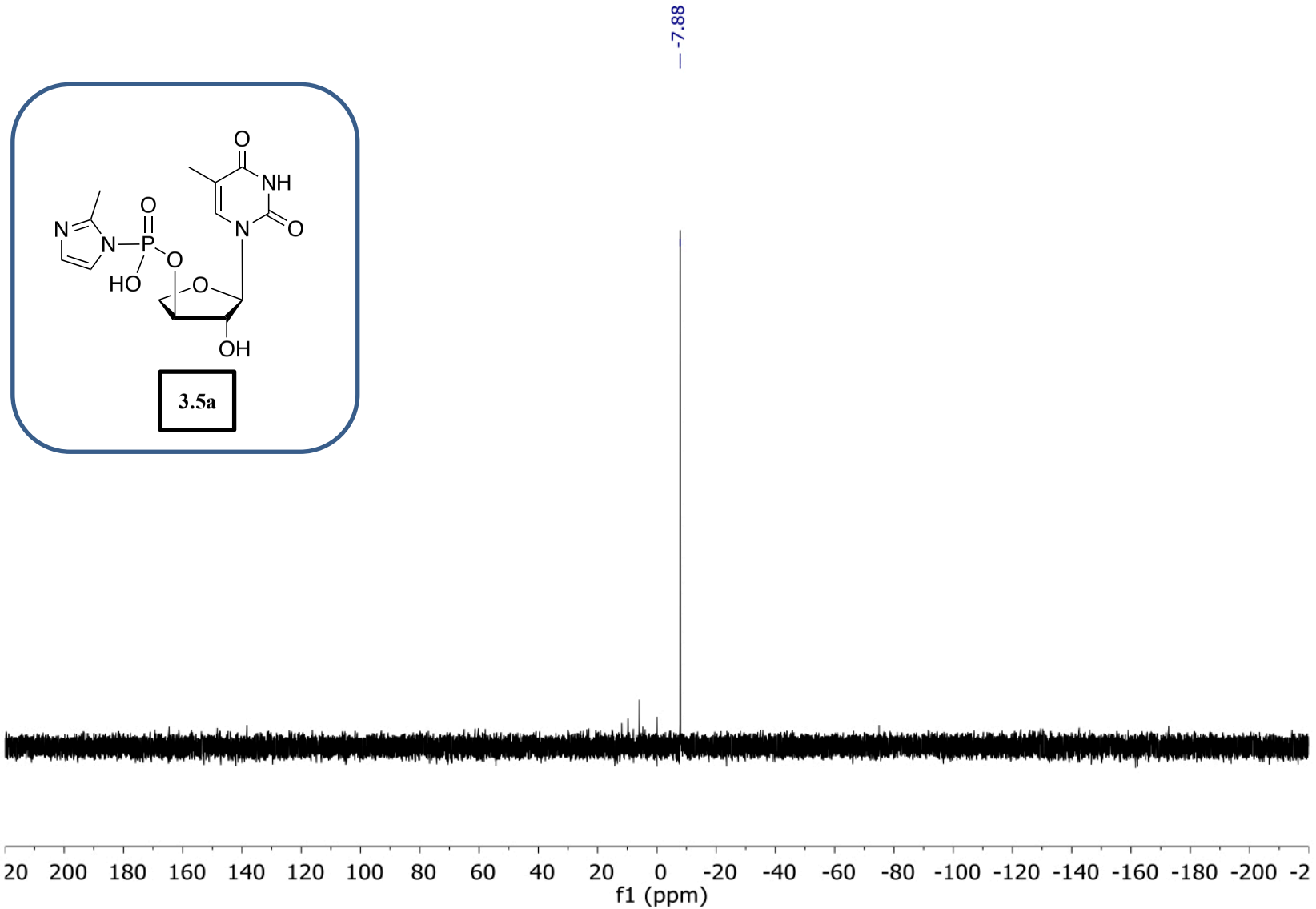
[1] HNMR spectrum of compound **3.4d** (500 MHz,DMSO-d^6)



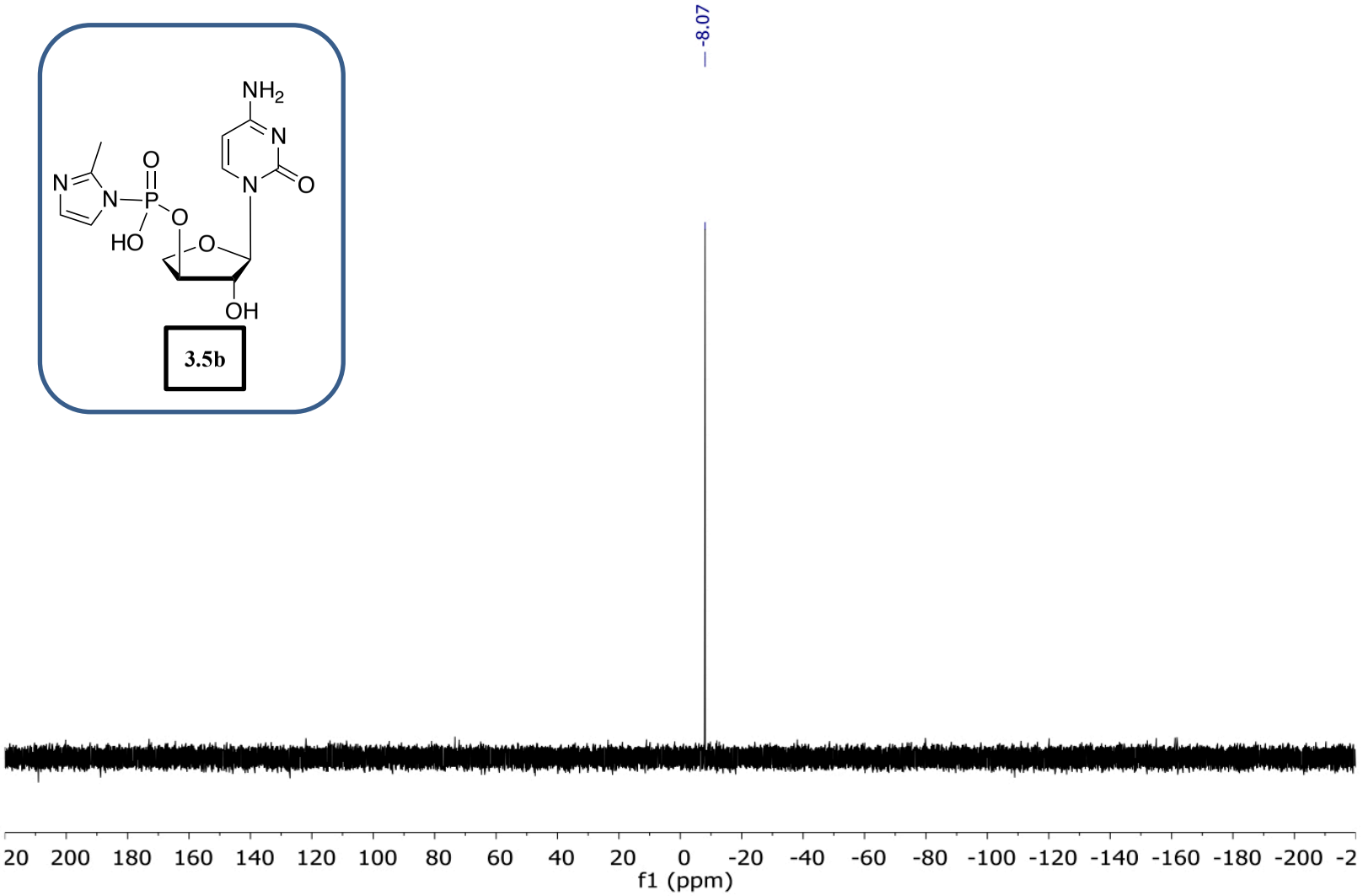
^{31}C NMR spectrum of compound **3.4d** (500 MHz,DMSO-d^6)



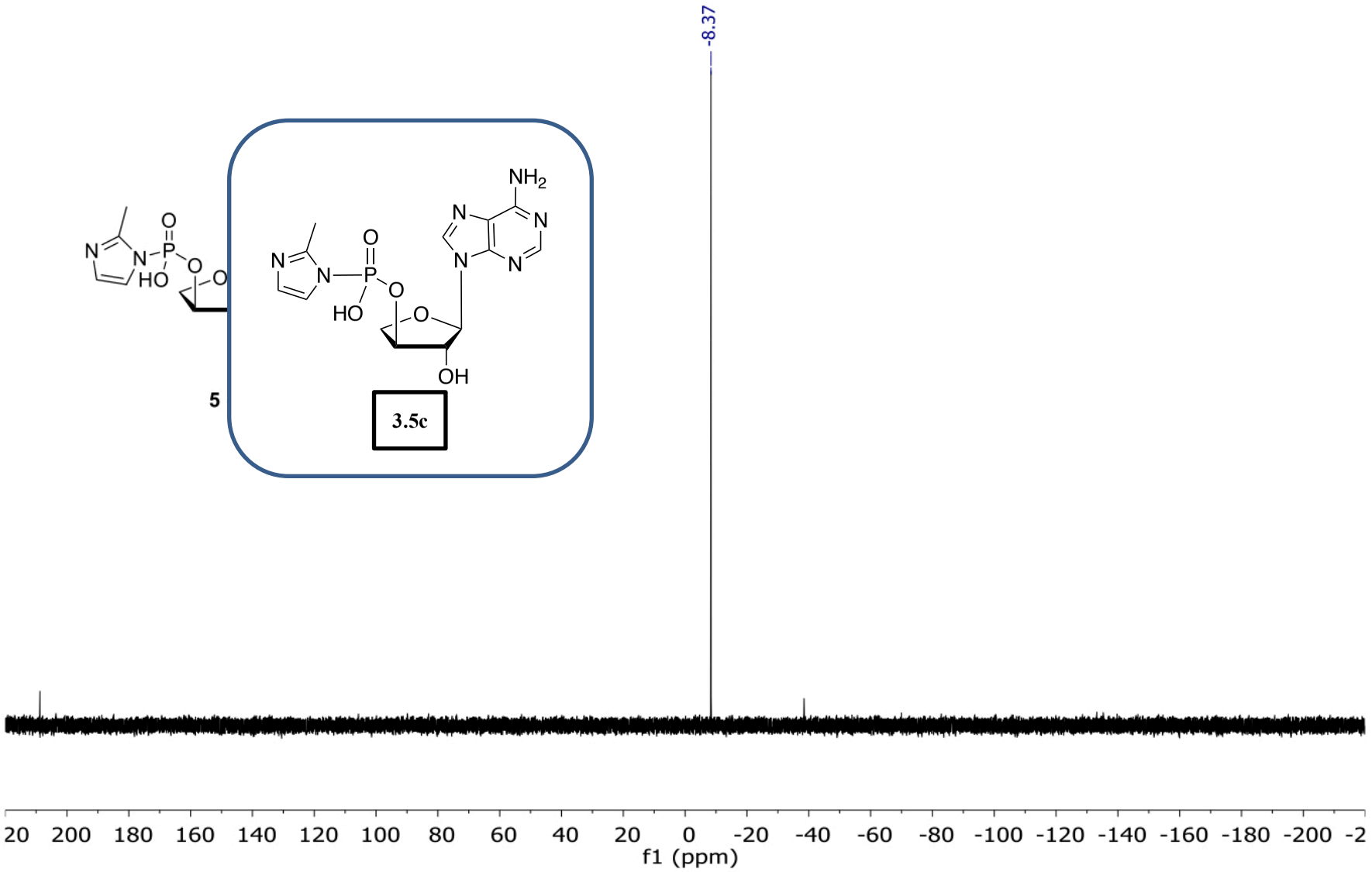
^{31}PNMR spectrum of compound **3.4d** (500 MHz,DMSO-d^6)



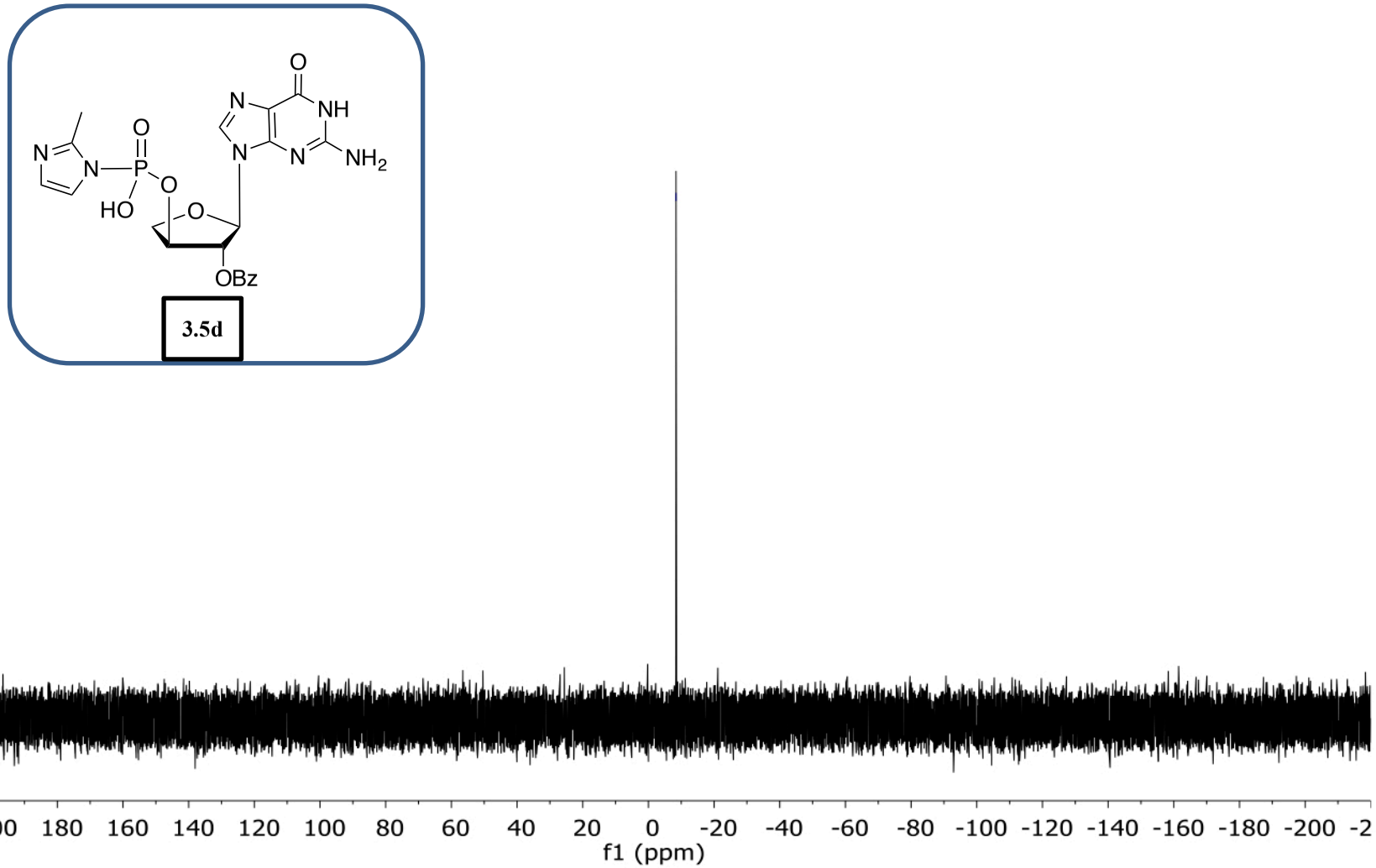
^{31}PNMR spectrum of compound **3.5a** (500 MHz,DMSO-d^6)



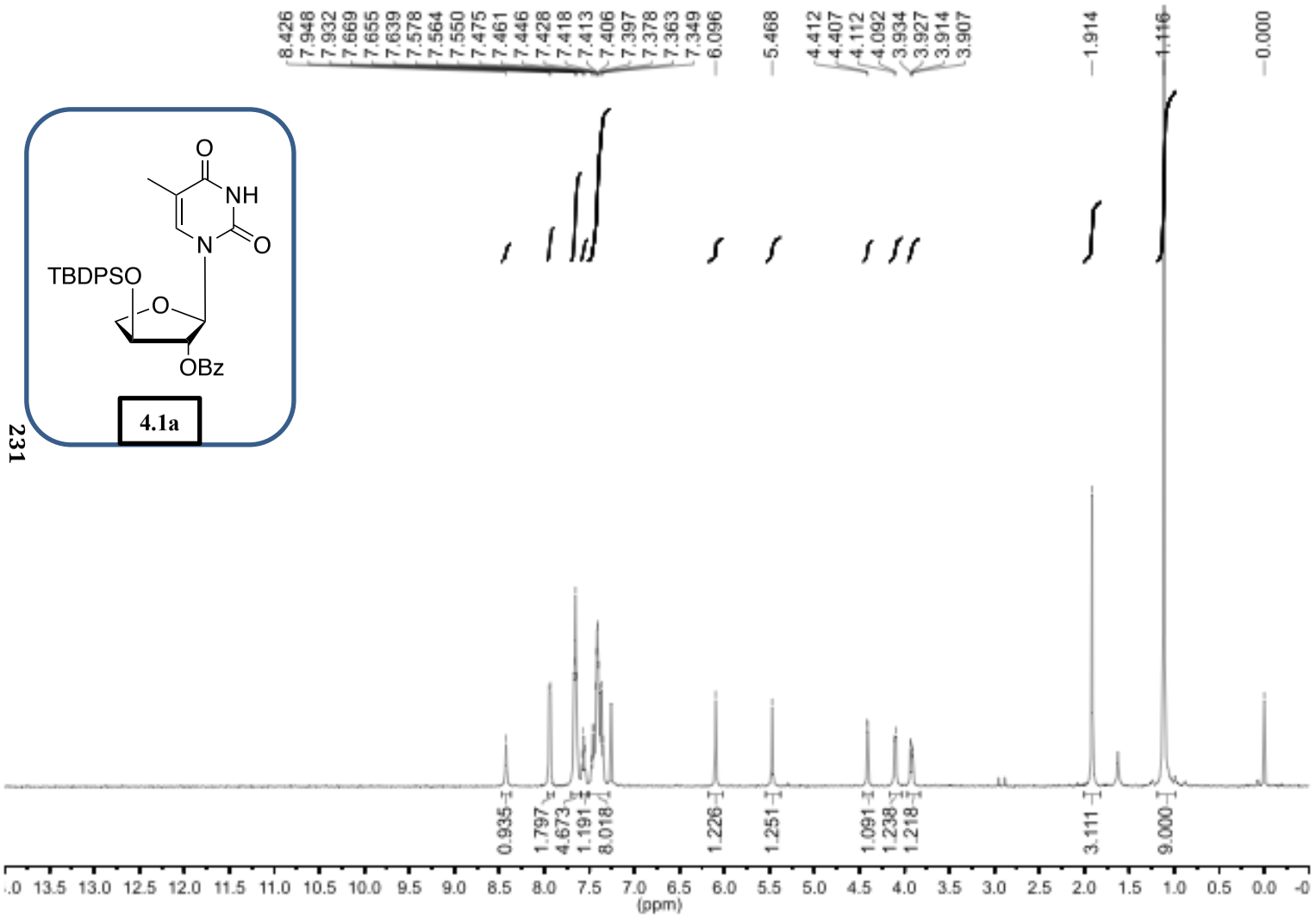
^{31}PNMR spectrum of compound **3.5b** (500 MHz,DMSO-d^6)



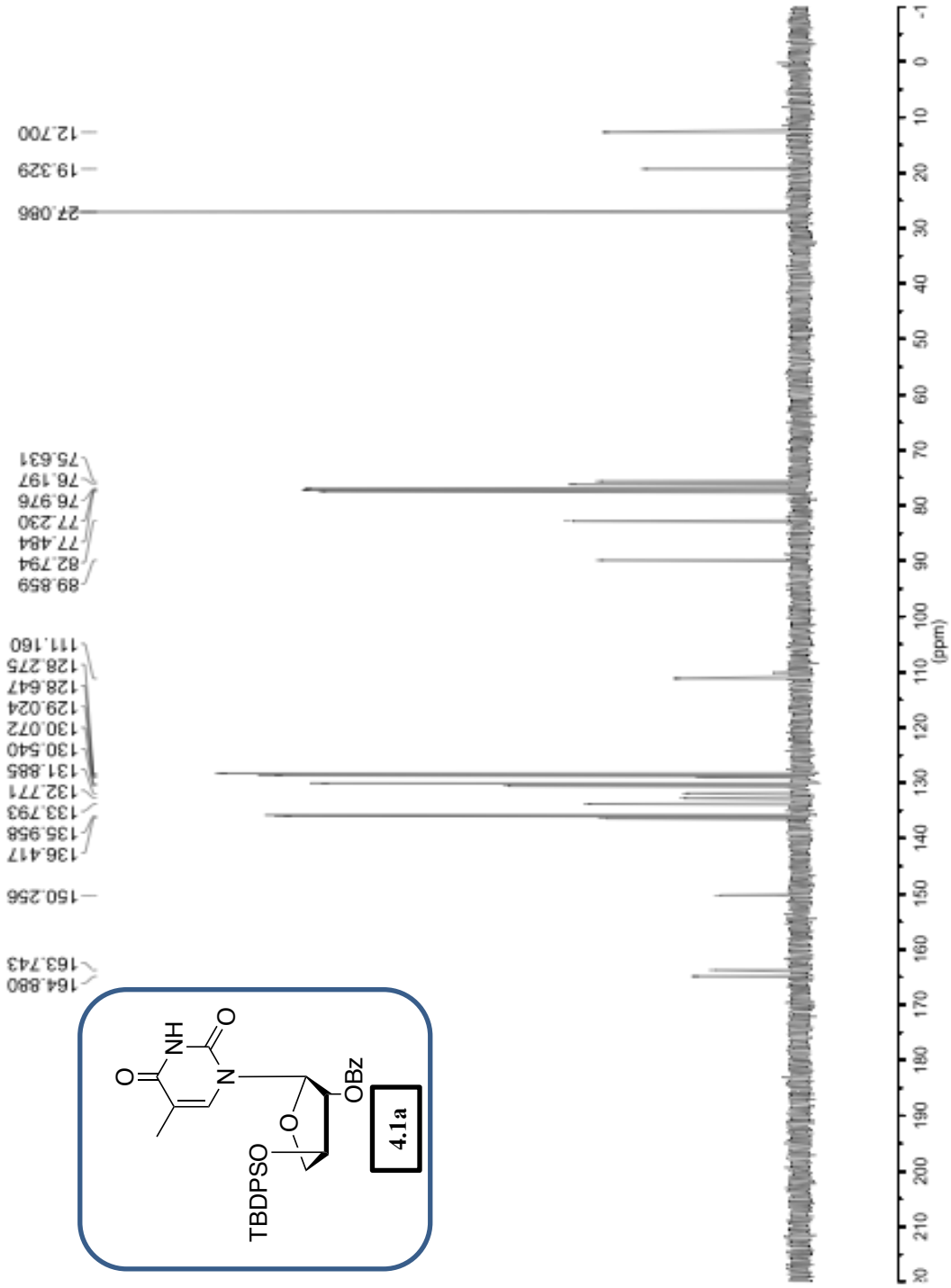
^{31}PNMR spectrum of compound **3.5c** (500 MHz,DMSO-d^6)



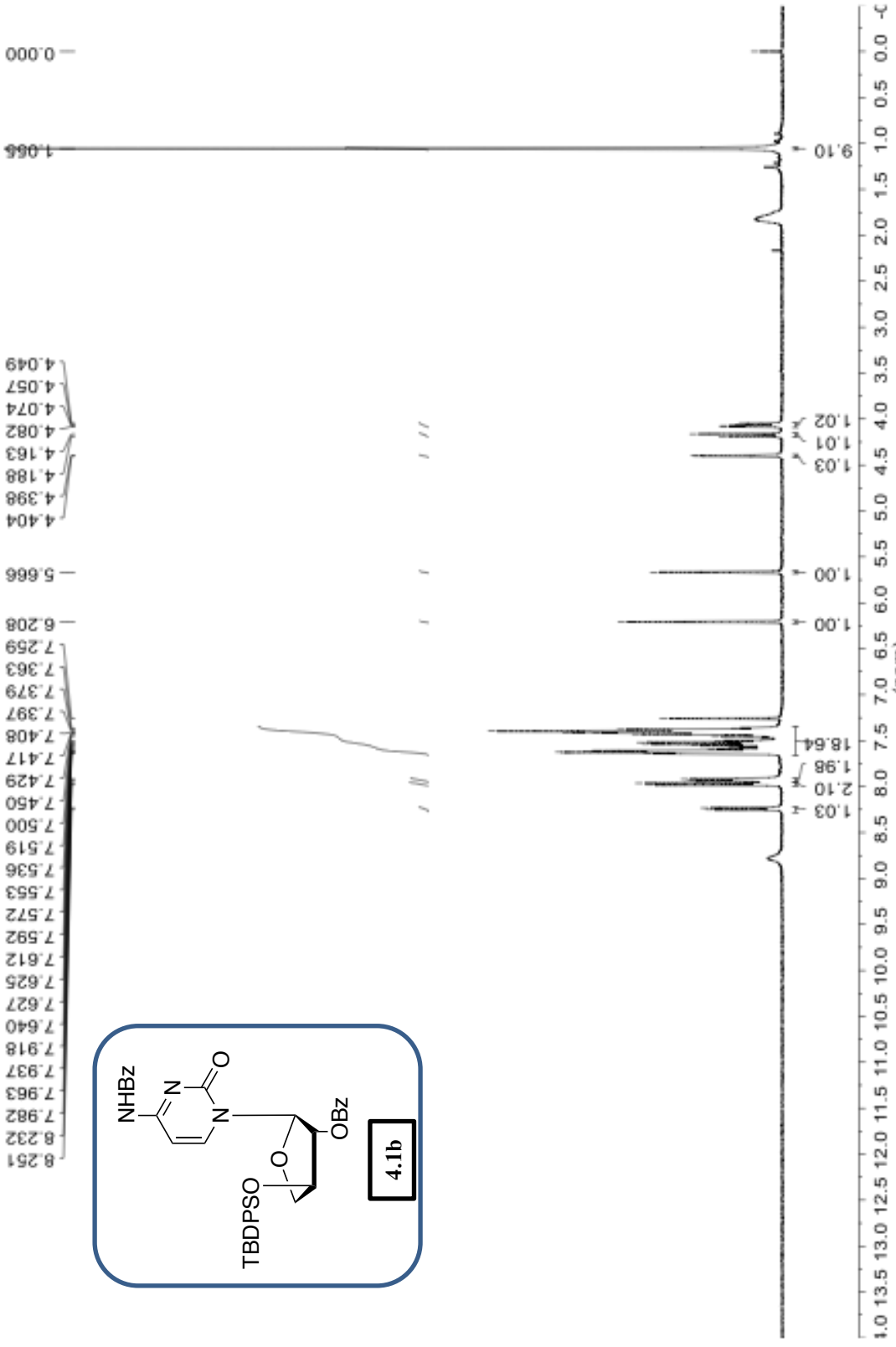
^{31}PNMR spectrum of compound **3.5d** (500 MHz,DMSO-d^6)



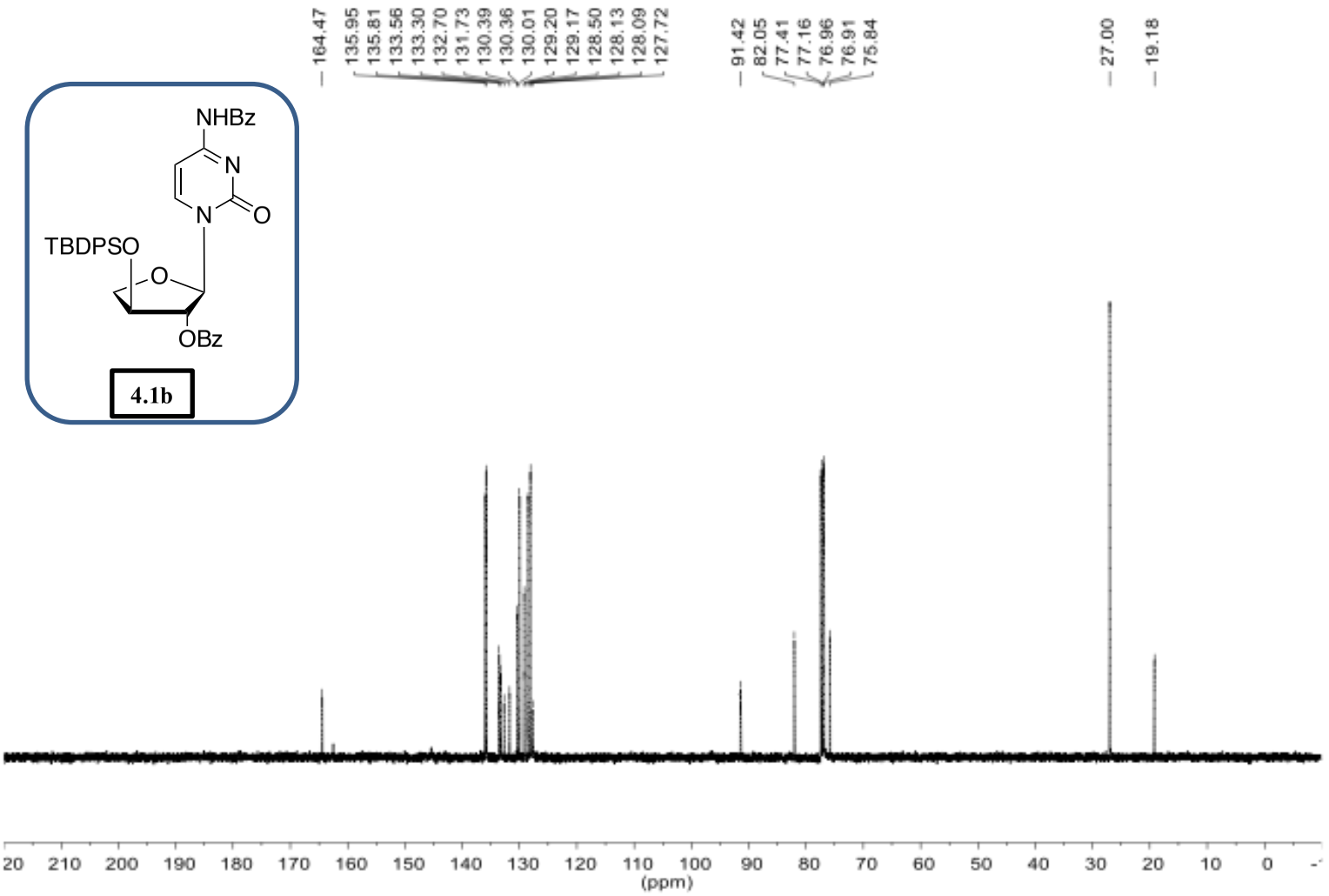
^{1}H NMR spectrum of compound **4.1a** (500 MHz,$CDCl_3$)



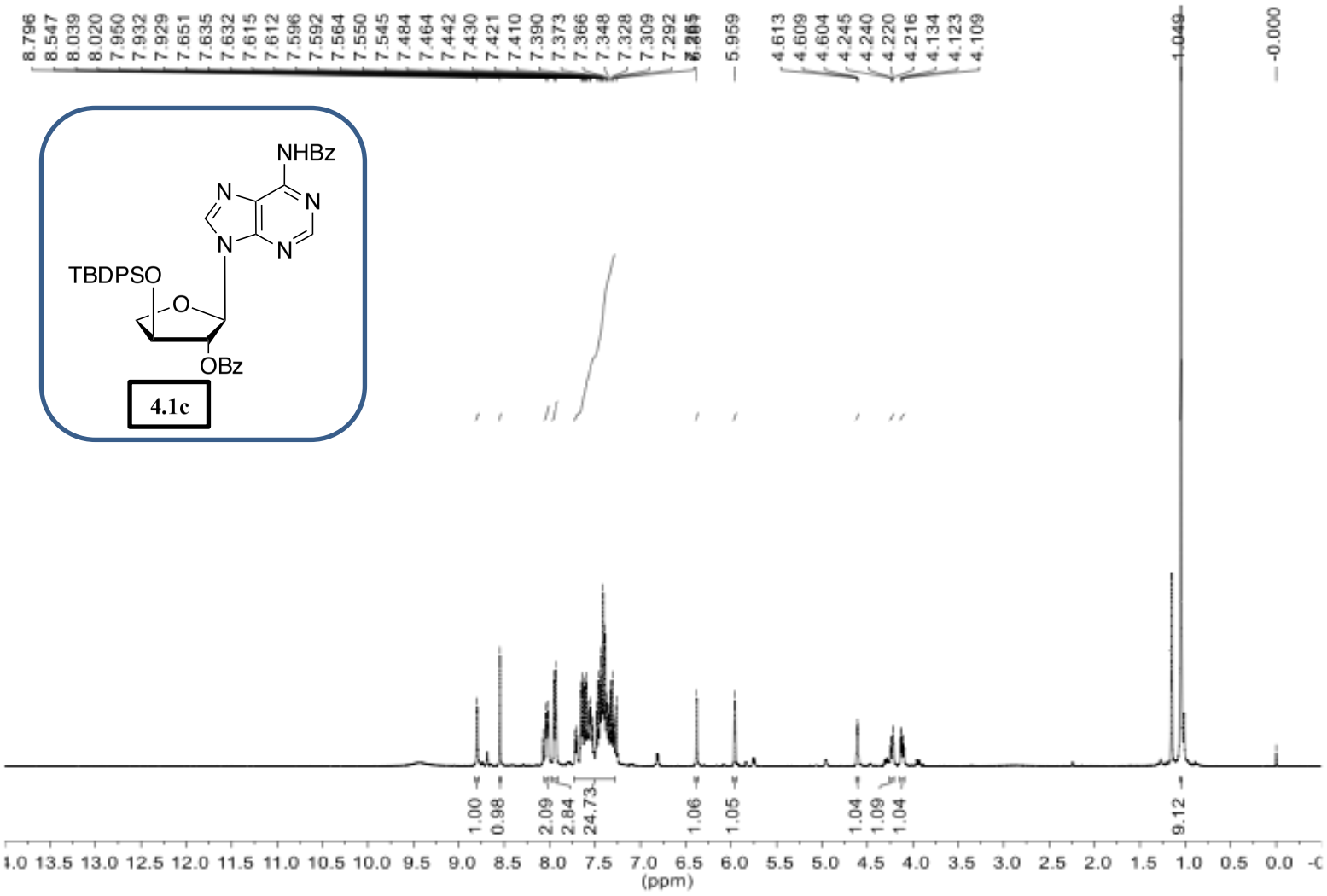
^{13}C NMR spectrum of compound **4.1a** (500 MHz,$CDCl_3$)



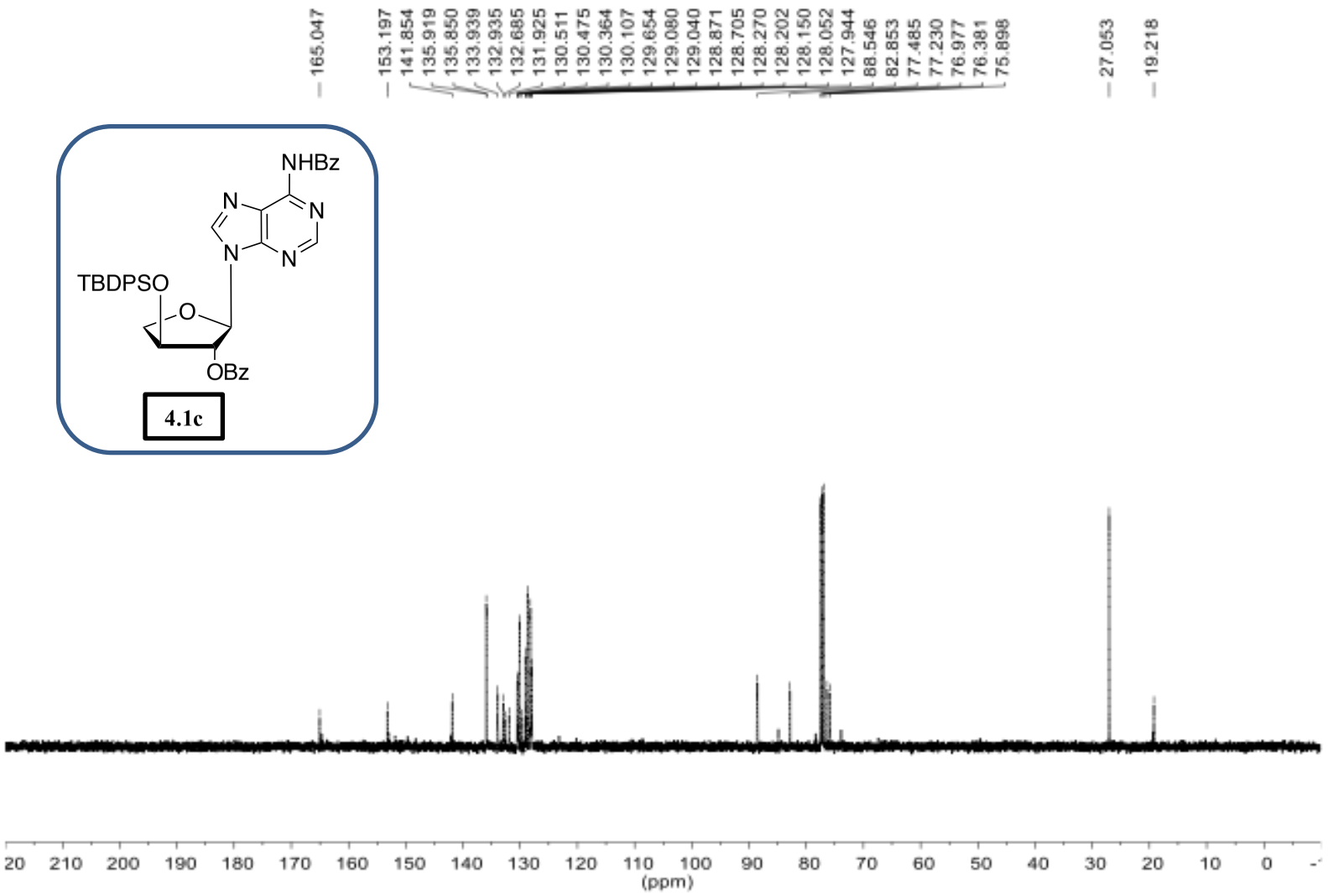
^{1}H NMR spectrum of compound **4.1b** (500 MHz,$CDCl_3$)



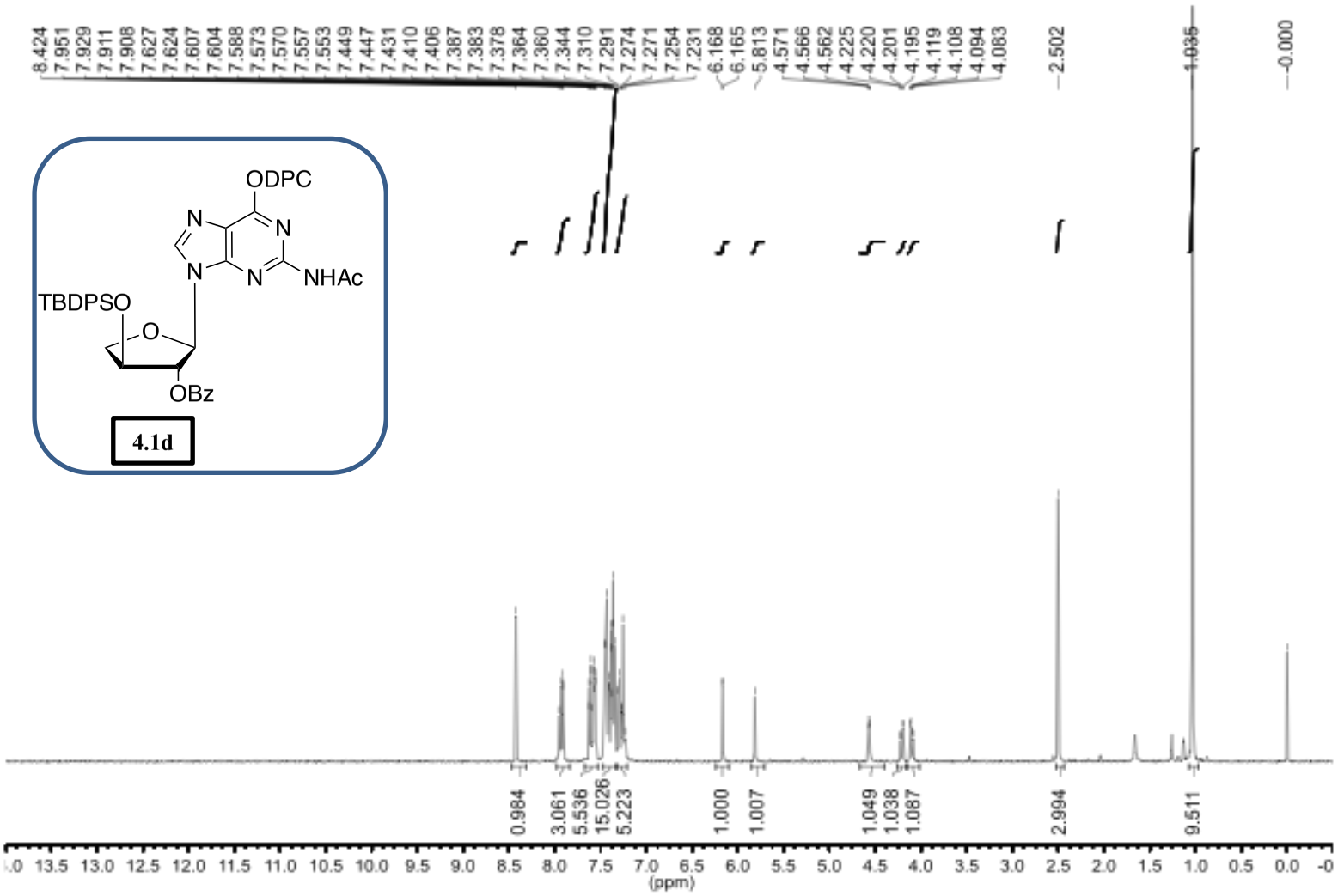
^{13}C NMR spectrum of compound **3.1b** (500 MHz,$CDCl_3$)



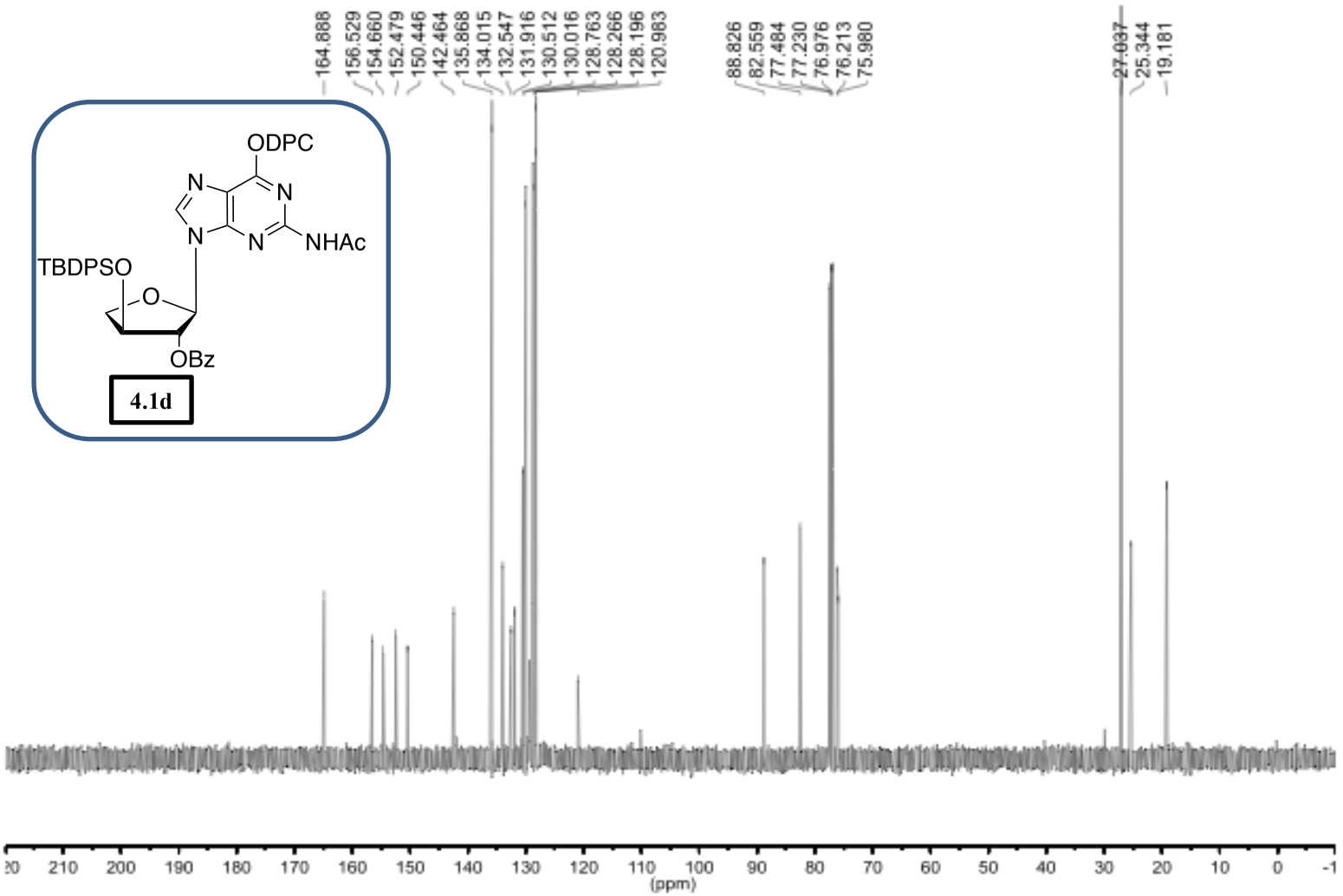
^{1}H NMR spectrum of compound **4.1c** (500 MHz,$CDCl_3$)



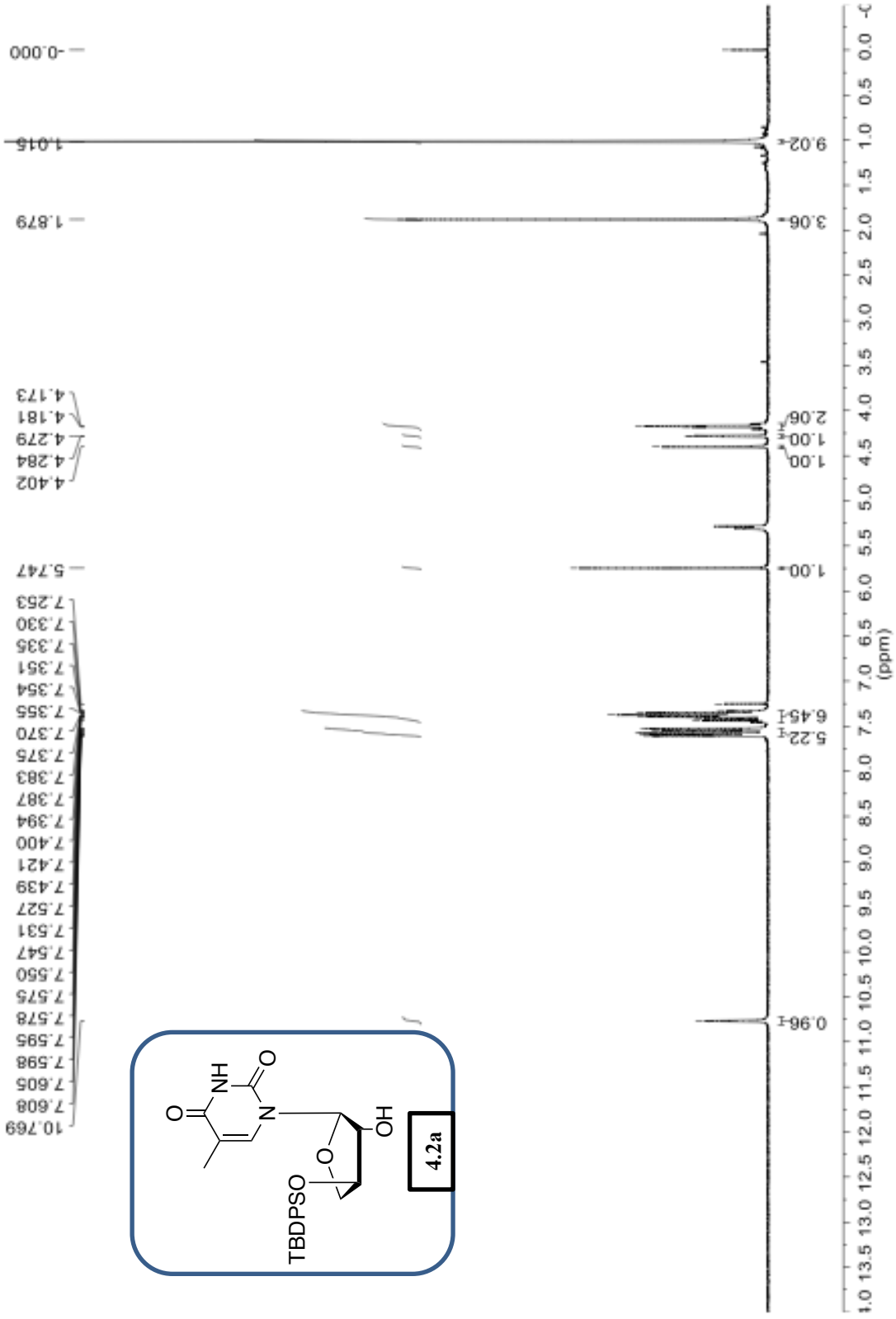
^{13}C NMR spectrum of compound **4.1c** (500 MHz,$CDCl_3$)



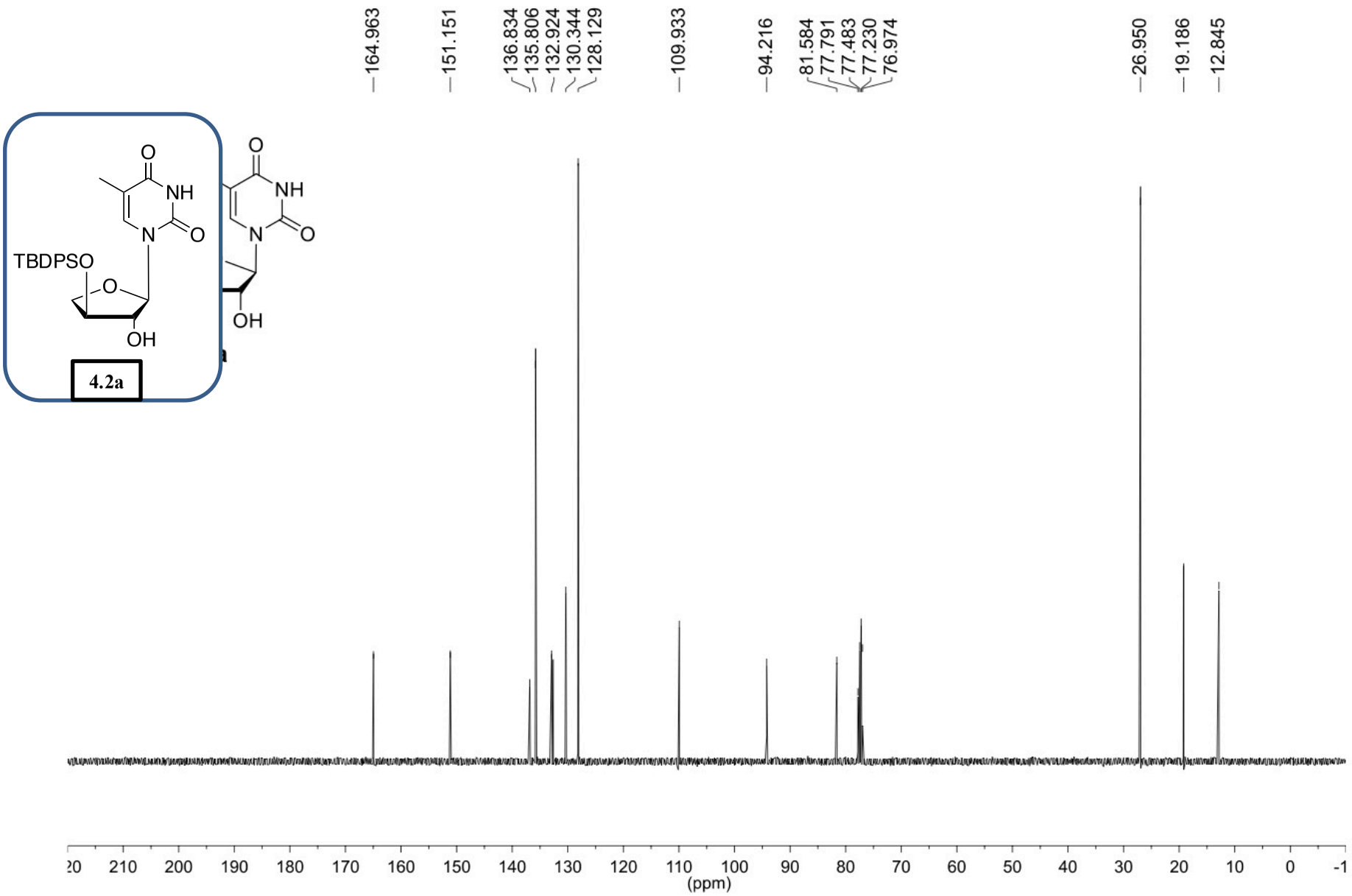
^{1}H NMR spectrum of compound **4.1d** (500 MHz,$CDCl_3$)



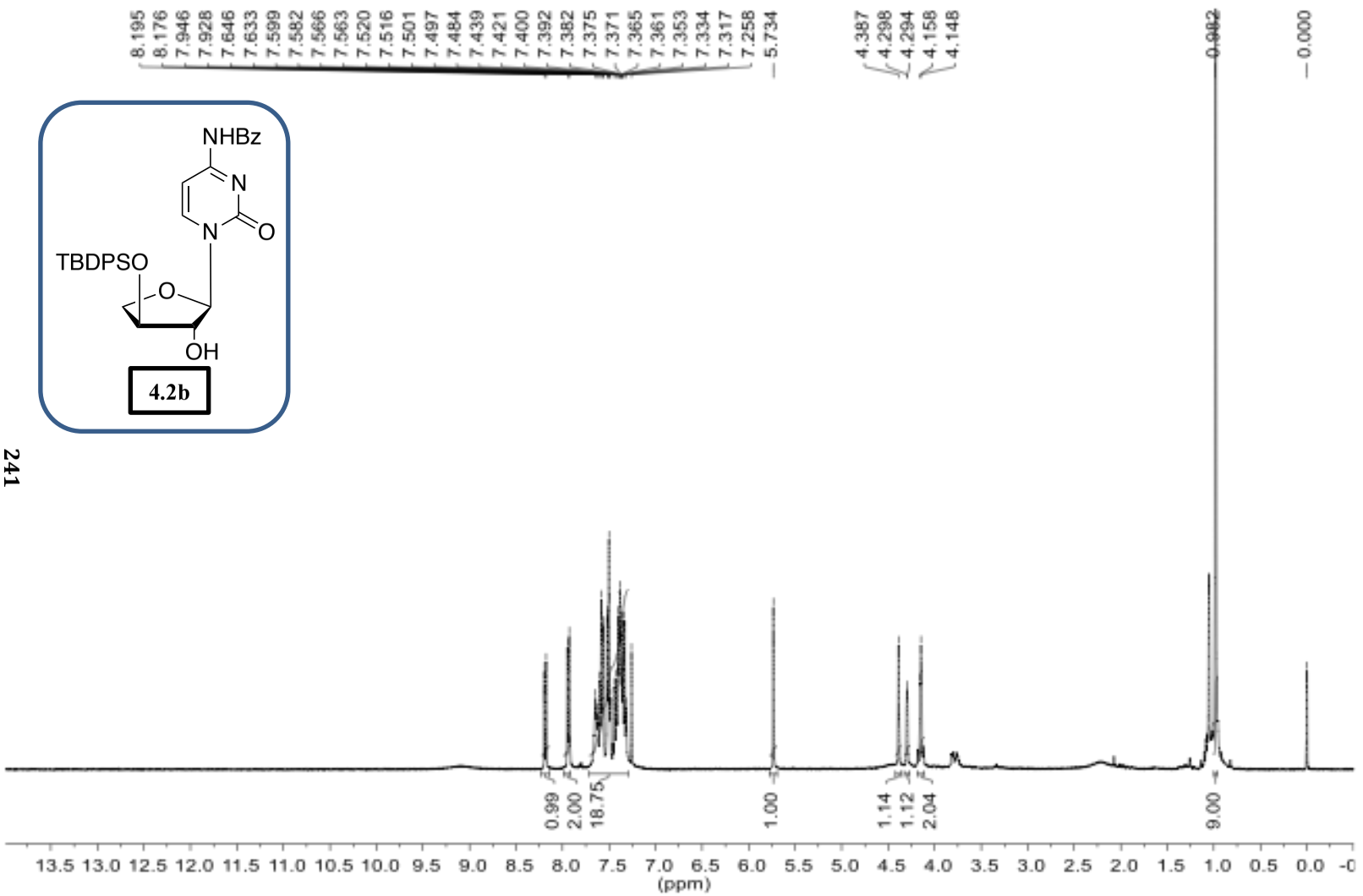
^{13}C NMR spectrum of compound **4.1d** (500 MHz,$CDCl_3$)



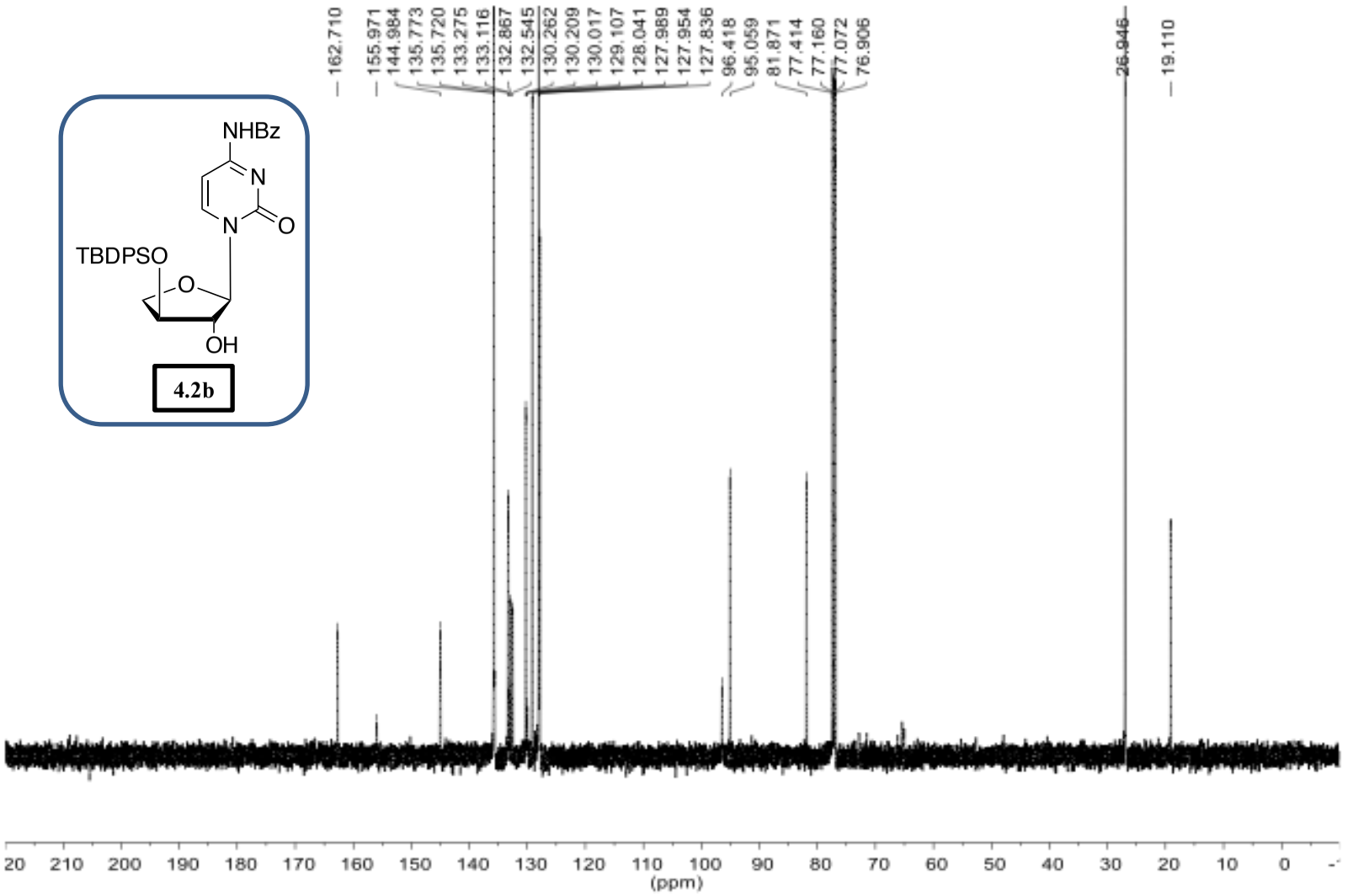
^{1}H NMR spectrum of compound **4.2a** (500 MHz,$CDCl_3$)



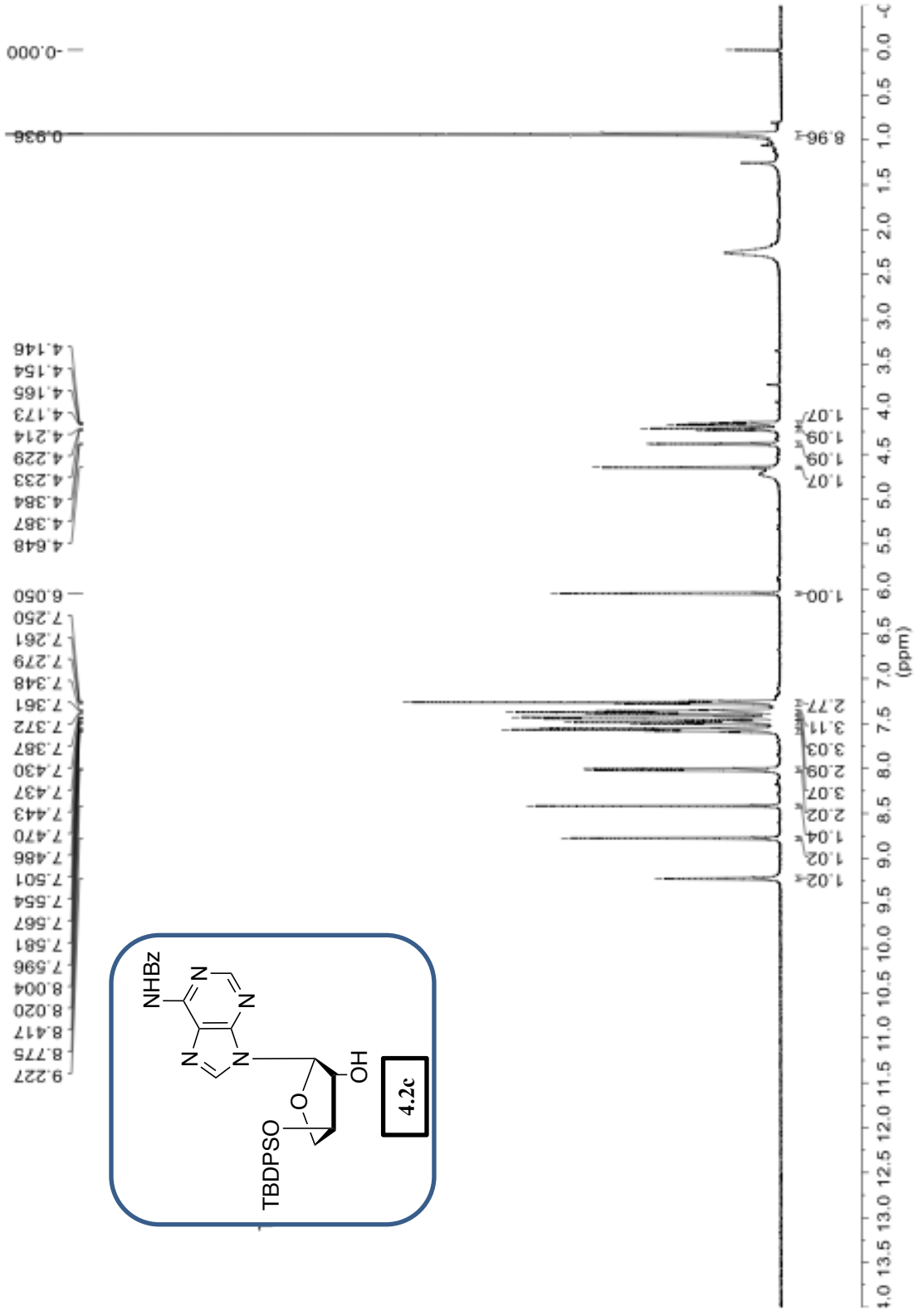
^{13}C NMR spectrum of compound **4.2a** (500 MHz,$CDCl_3$)



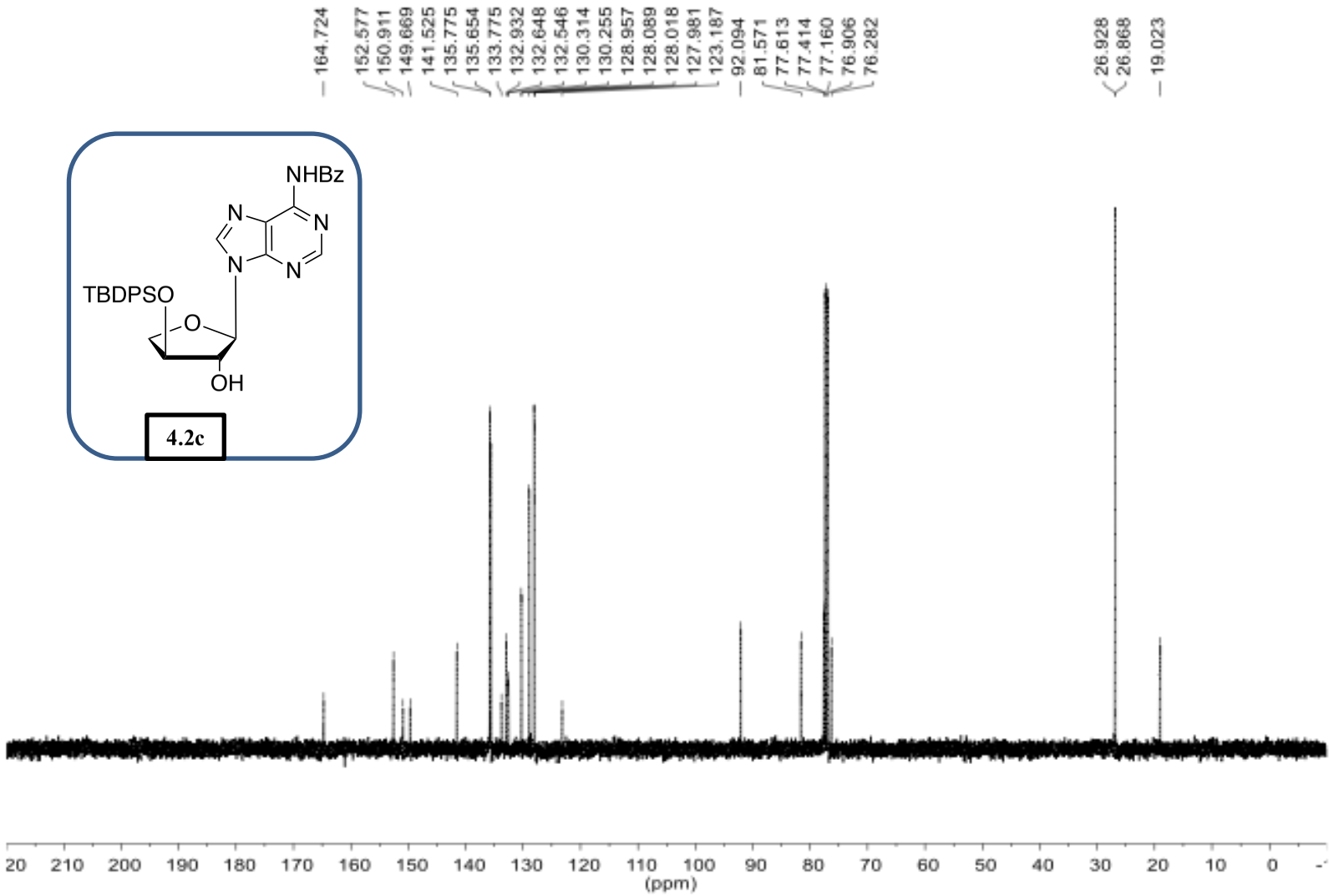
^{1}H NMR spectrum of compound **4.2b** (500 MHz,$CDCl_3$)



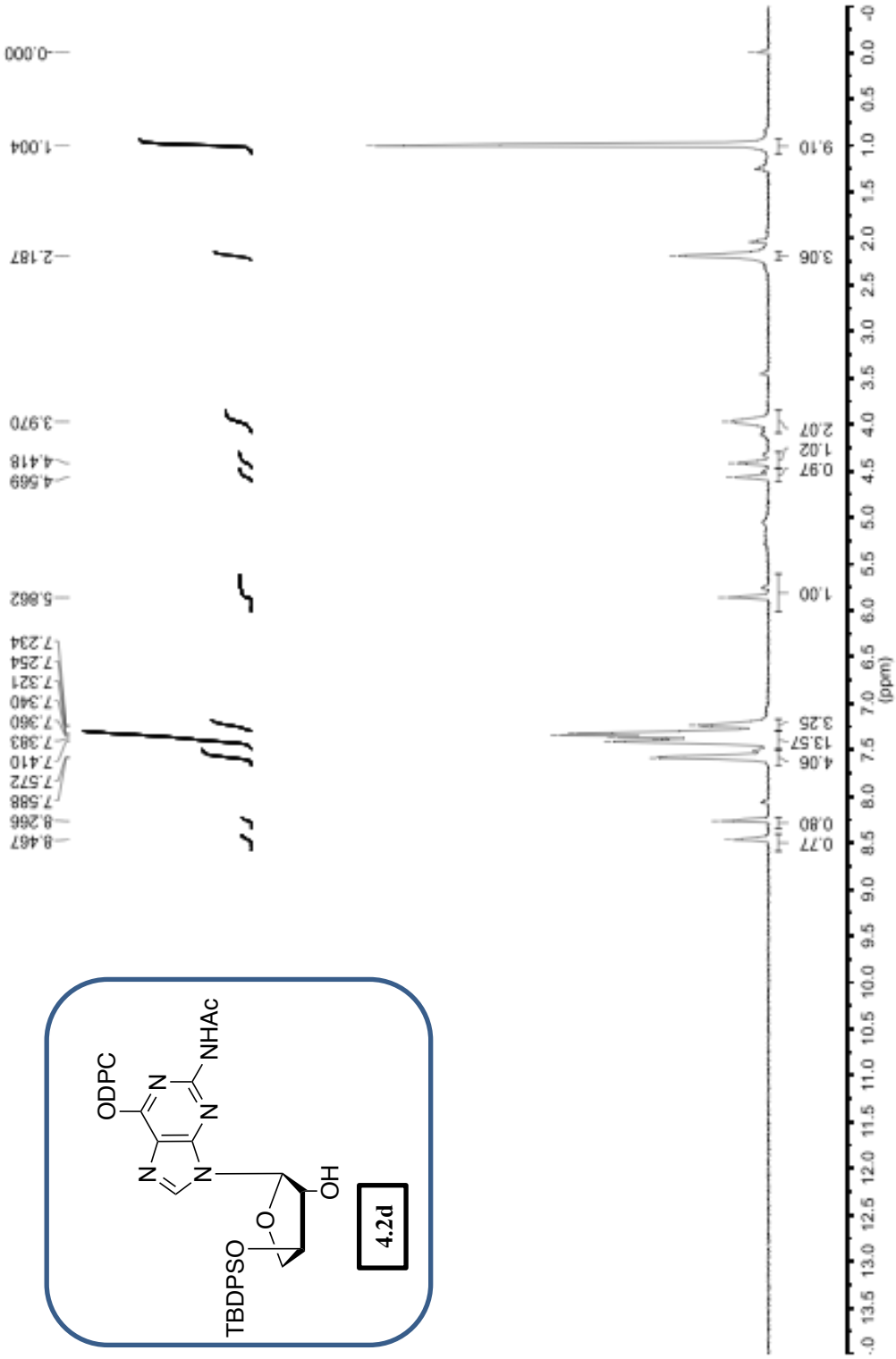
^{13}C NMR spectrum of compound **4.2b** (500 MHz,$CDCl_3$)



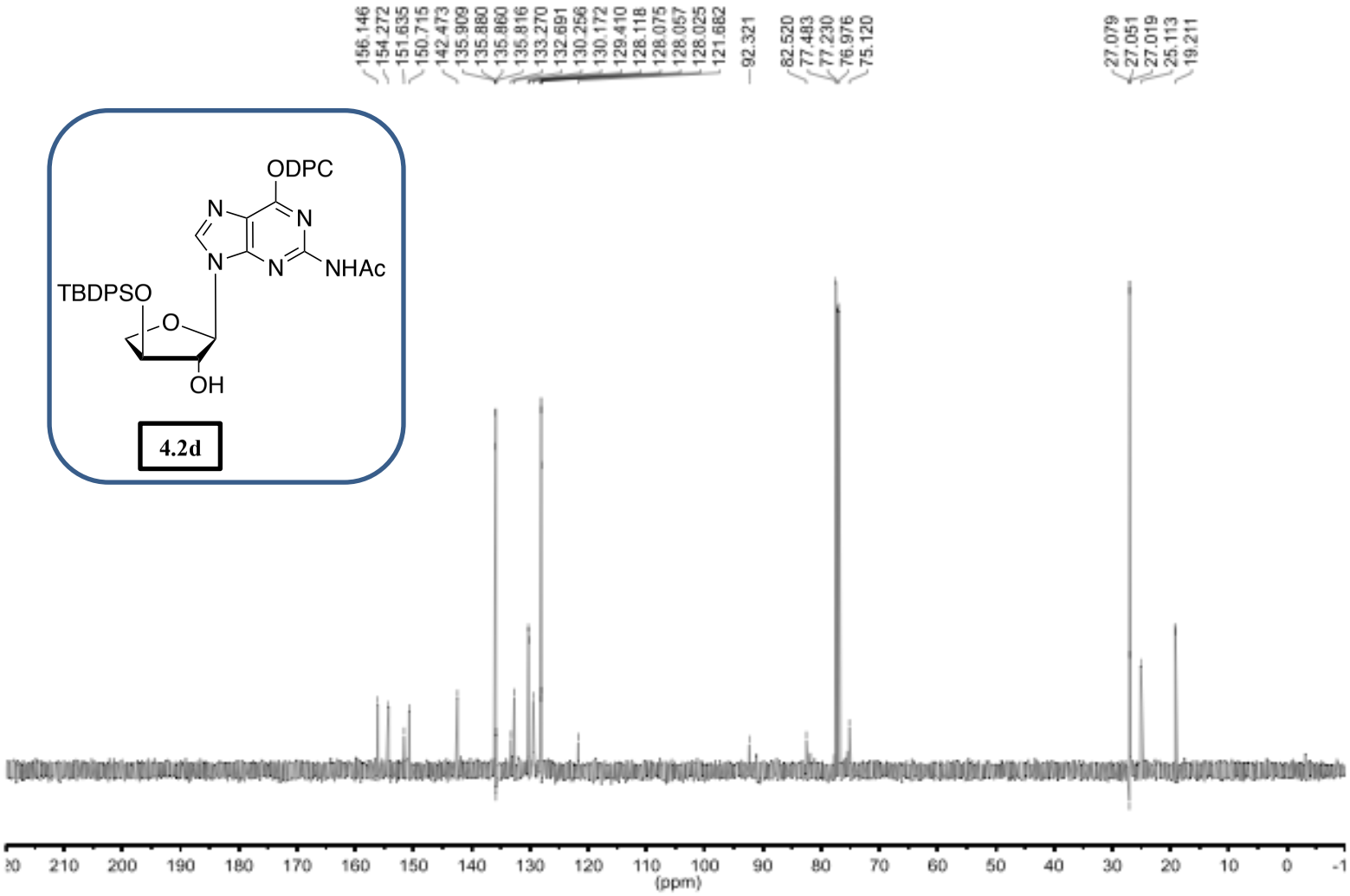
^{1}H NMR spectrum of compound **4.2c** (500 MHz,$CDCl_3$)



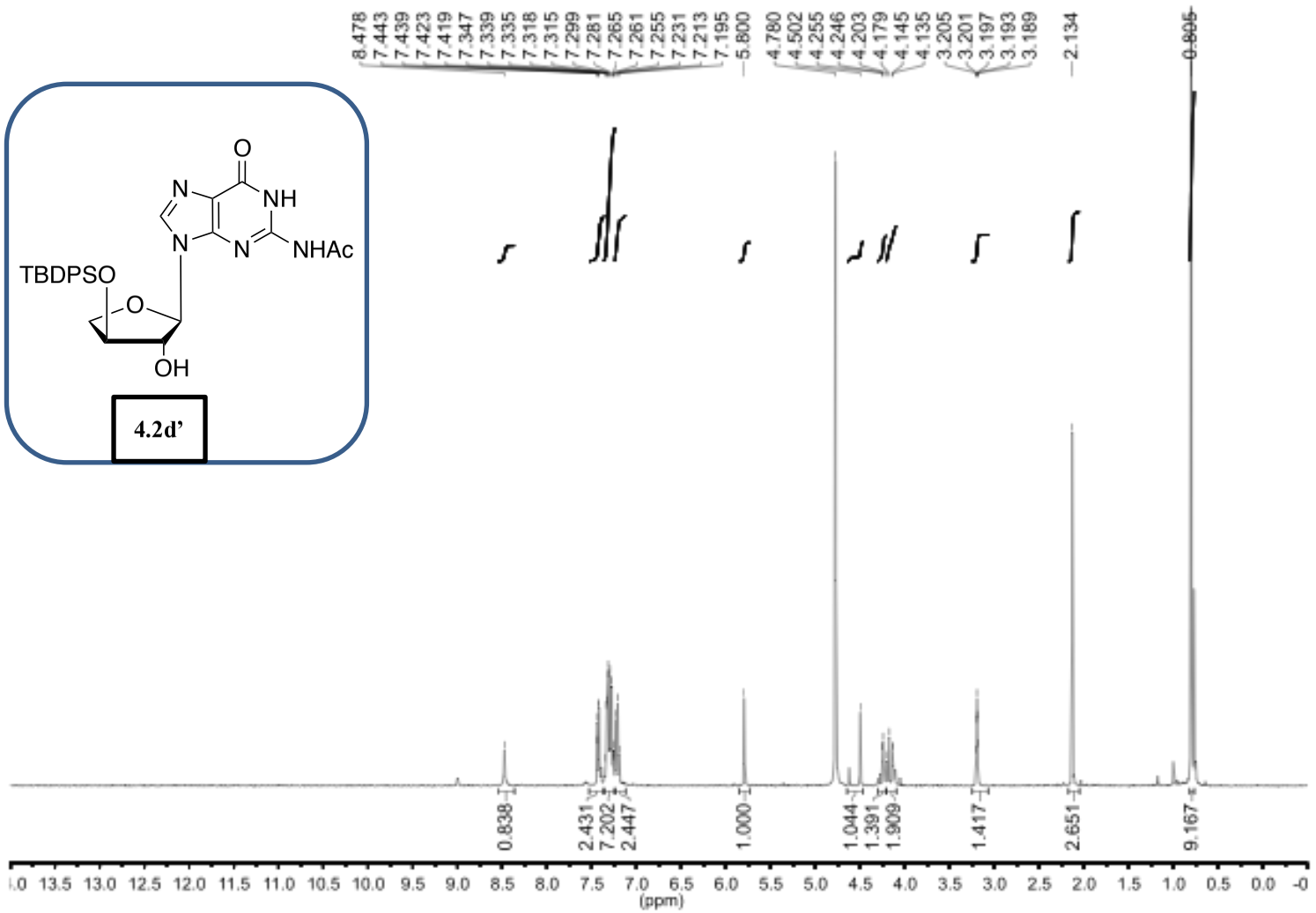
^{13}C NMR spectrum of compound **4.2c** (500 MHz,$CDCl_3$)



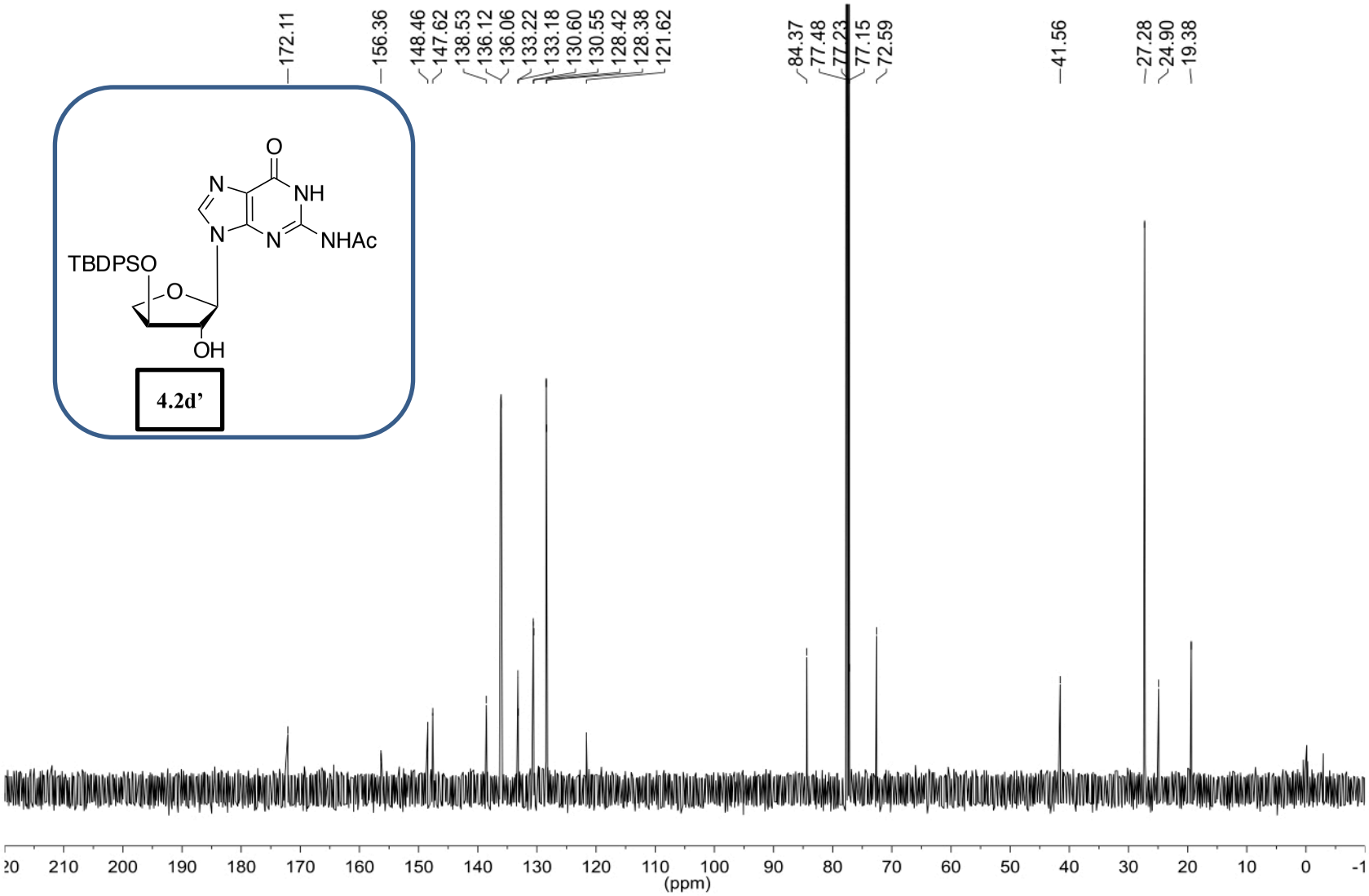
^{1}H NMR spectrum of compound **4.2d** (500 MHz,$CDCl_3$)



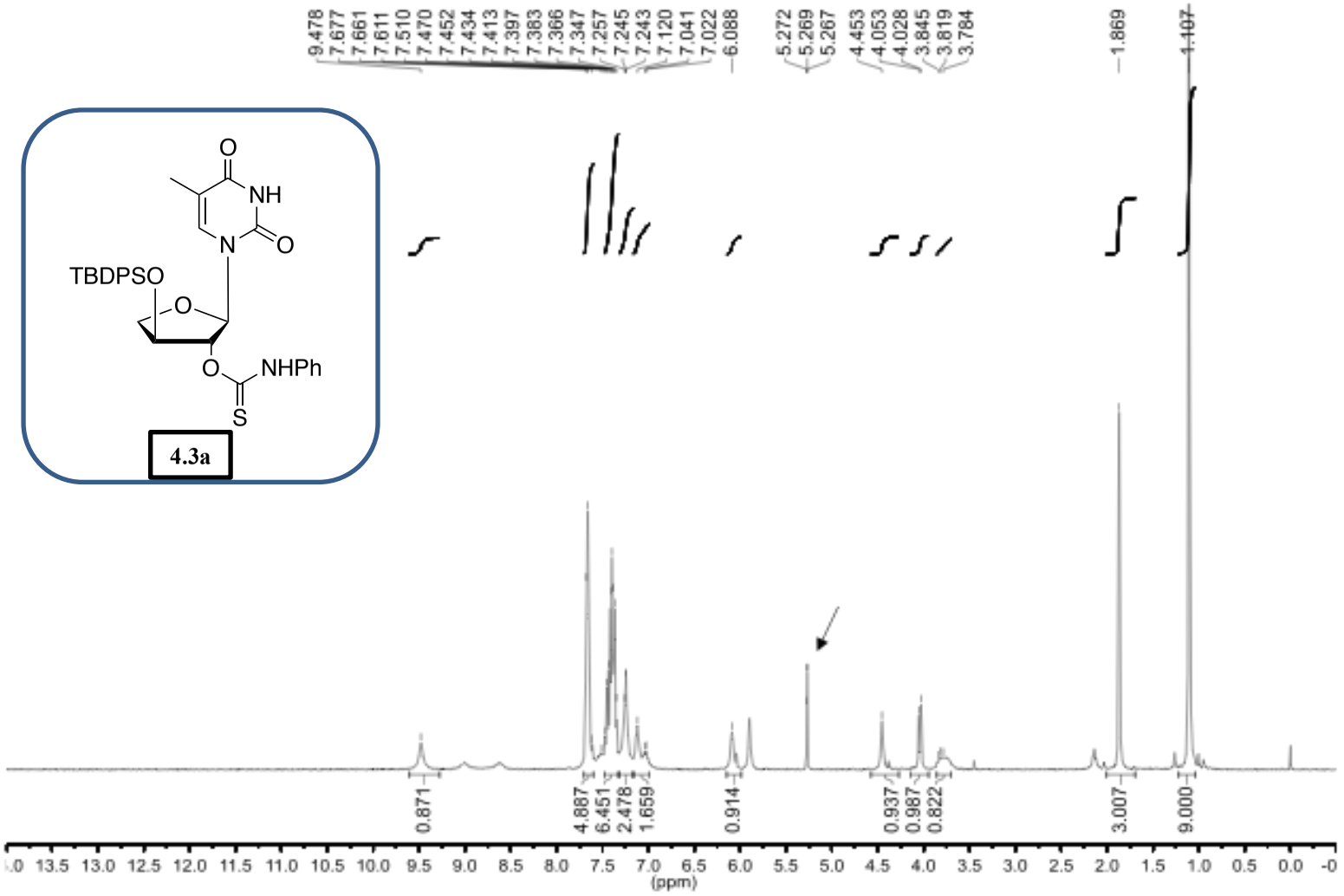
^{13}C NMR spectrum of compound **4.2d** (500 MHz,$CDCl_3$)



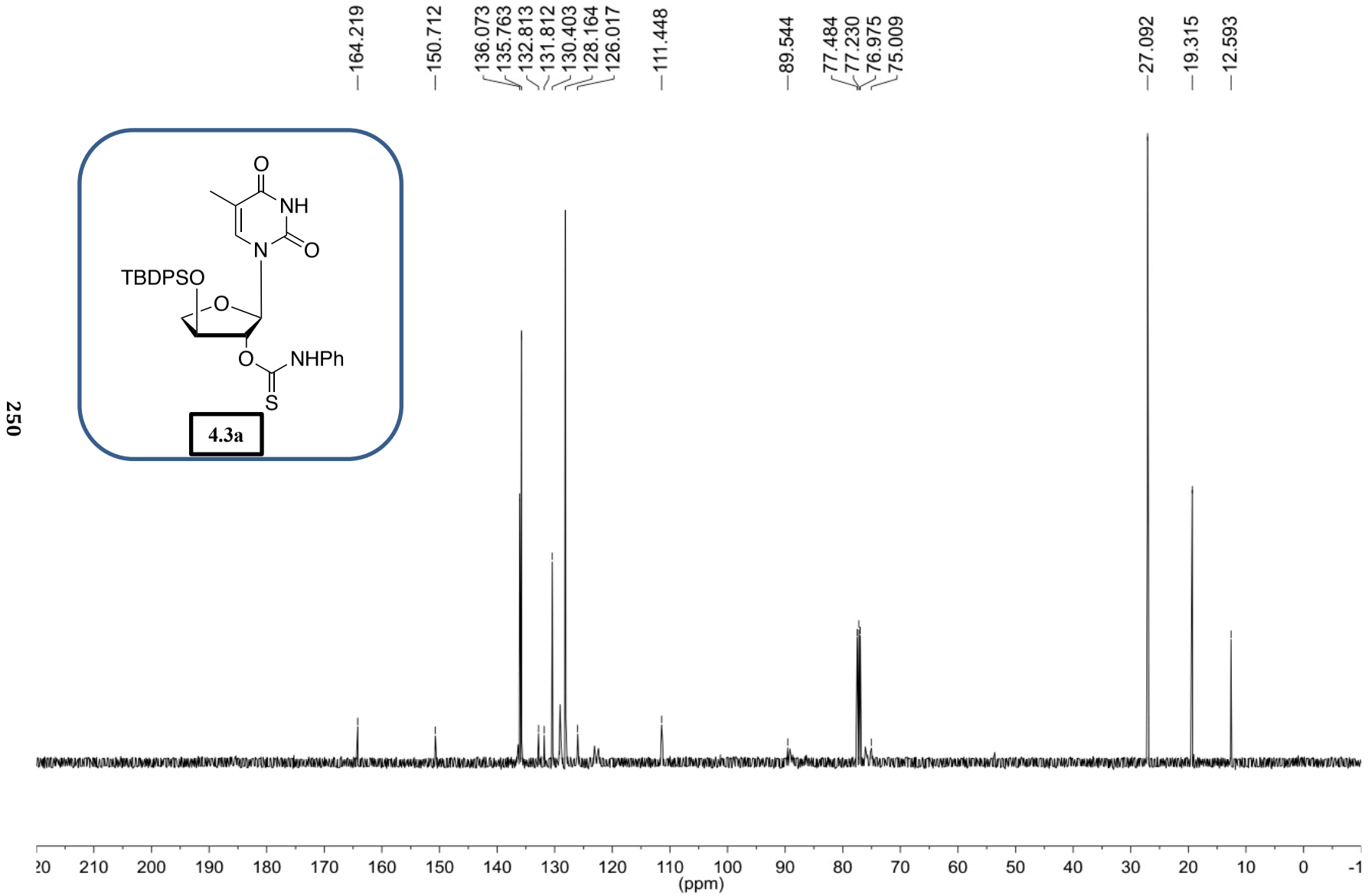
^{1}H NMR spectrum of compound **4.2d'** (500 MHz,CD_3OD)



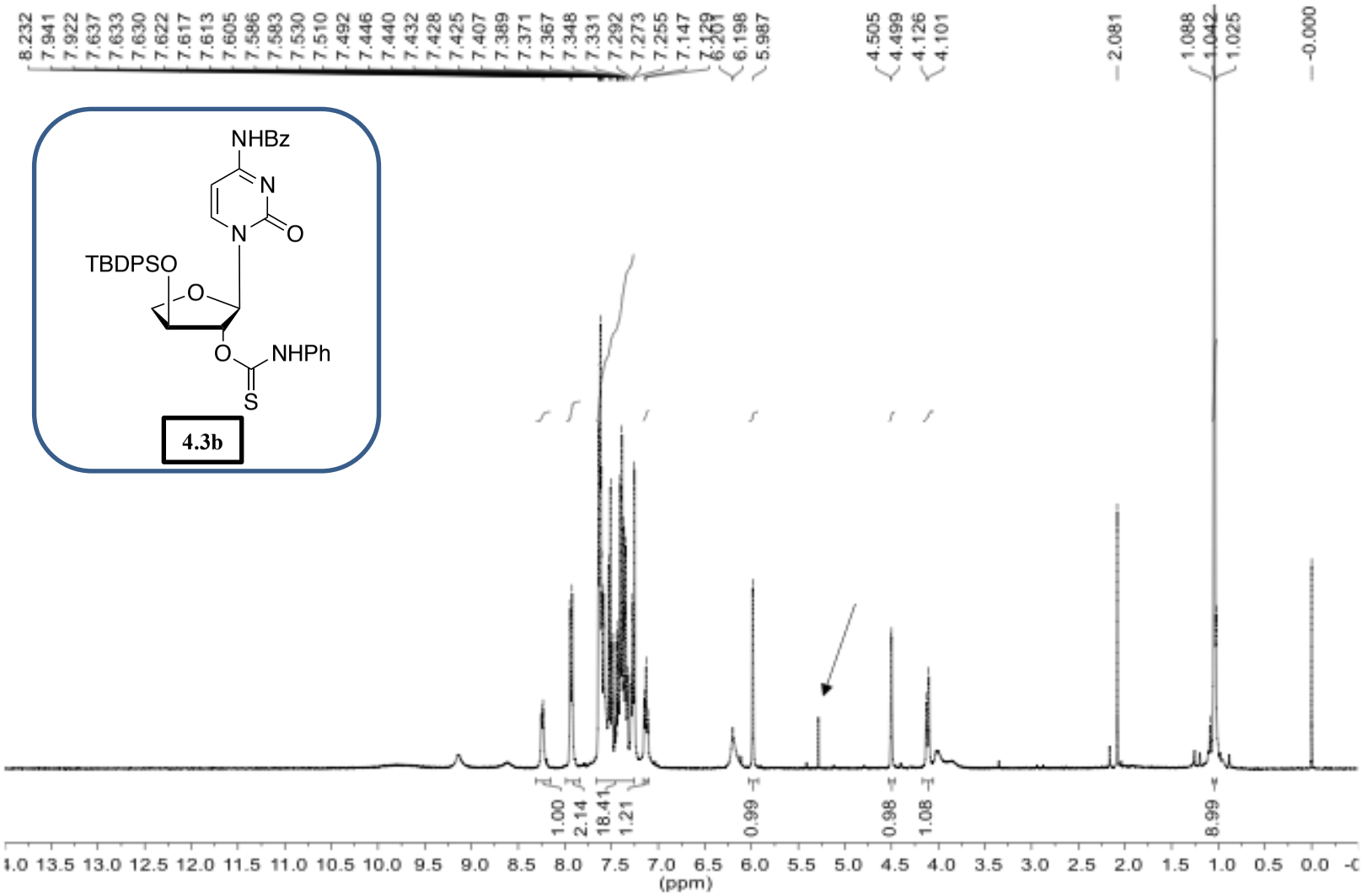
^{13}C NMR spectrum of compound **4.2d'** (500 MHz,$CDCl_3$)



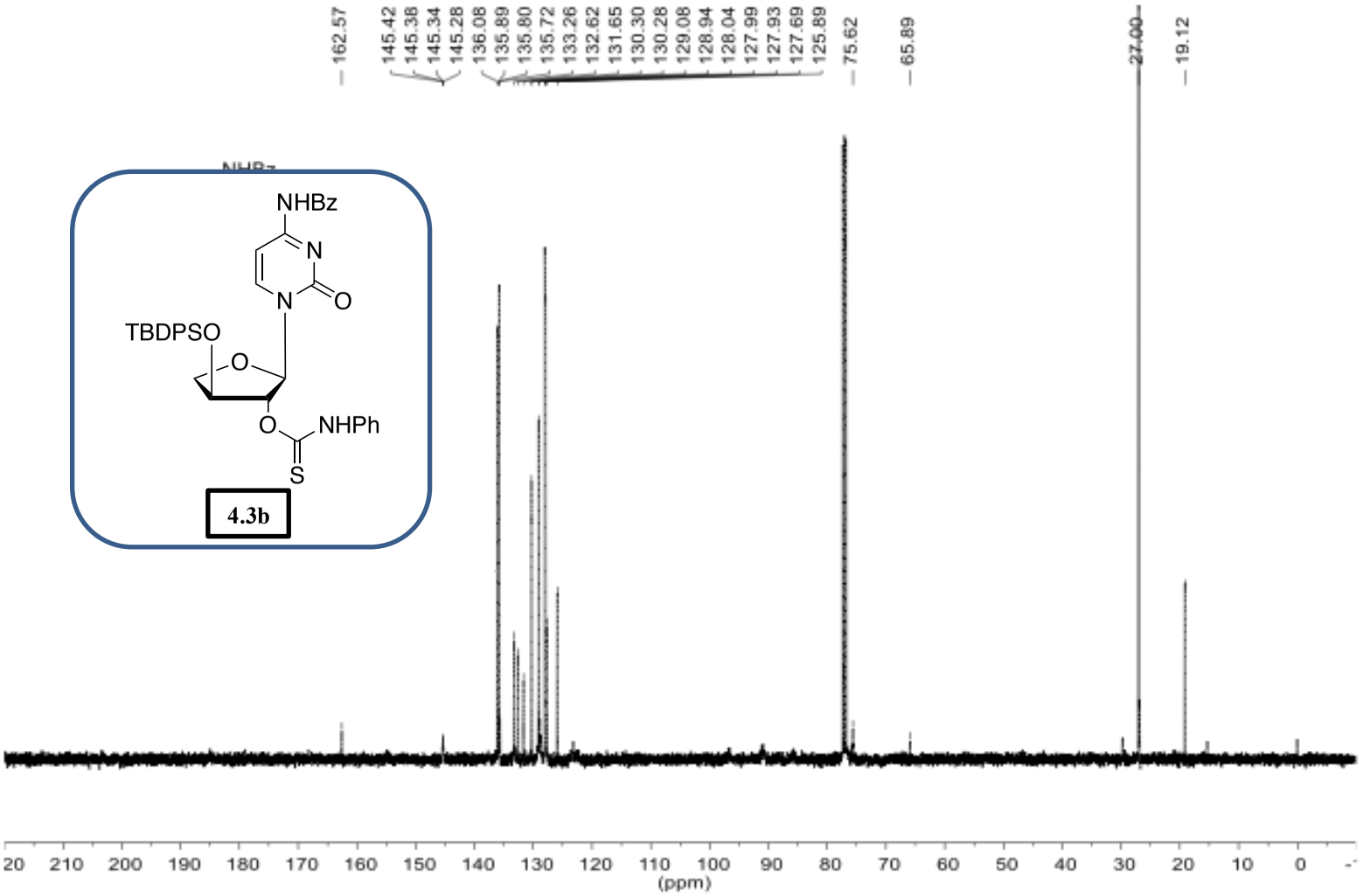
^{1}H NMR spectrum of compound **4.3a** (500 MHz,CD_3OD)



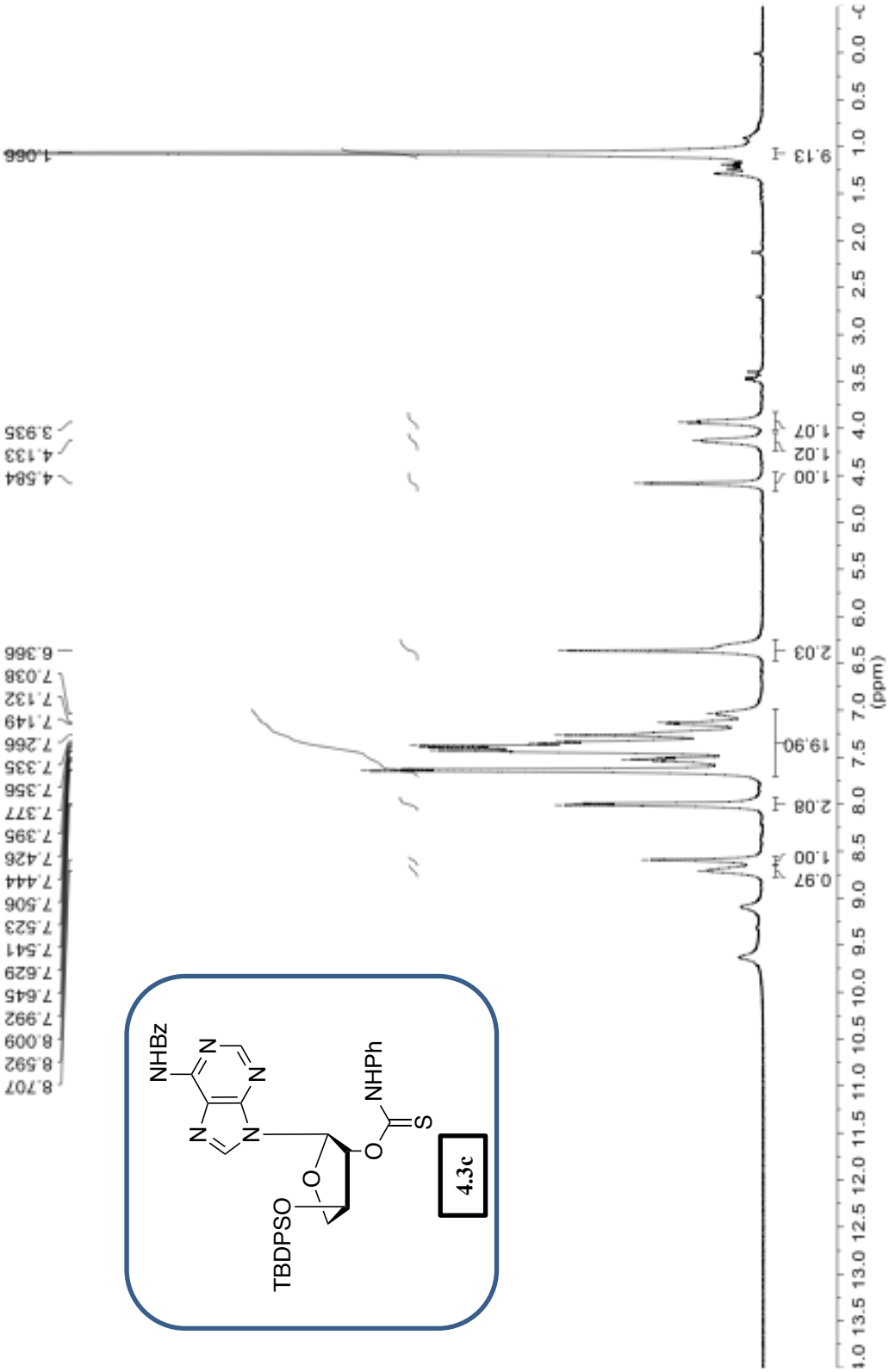
^{13}C NMR spectrum of compound **4.3a** (500 MHz,CD_3OD)



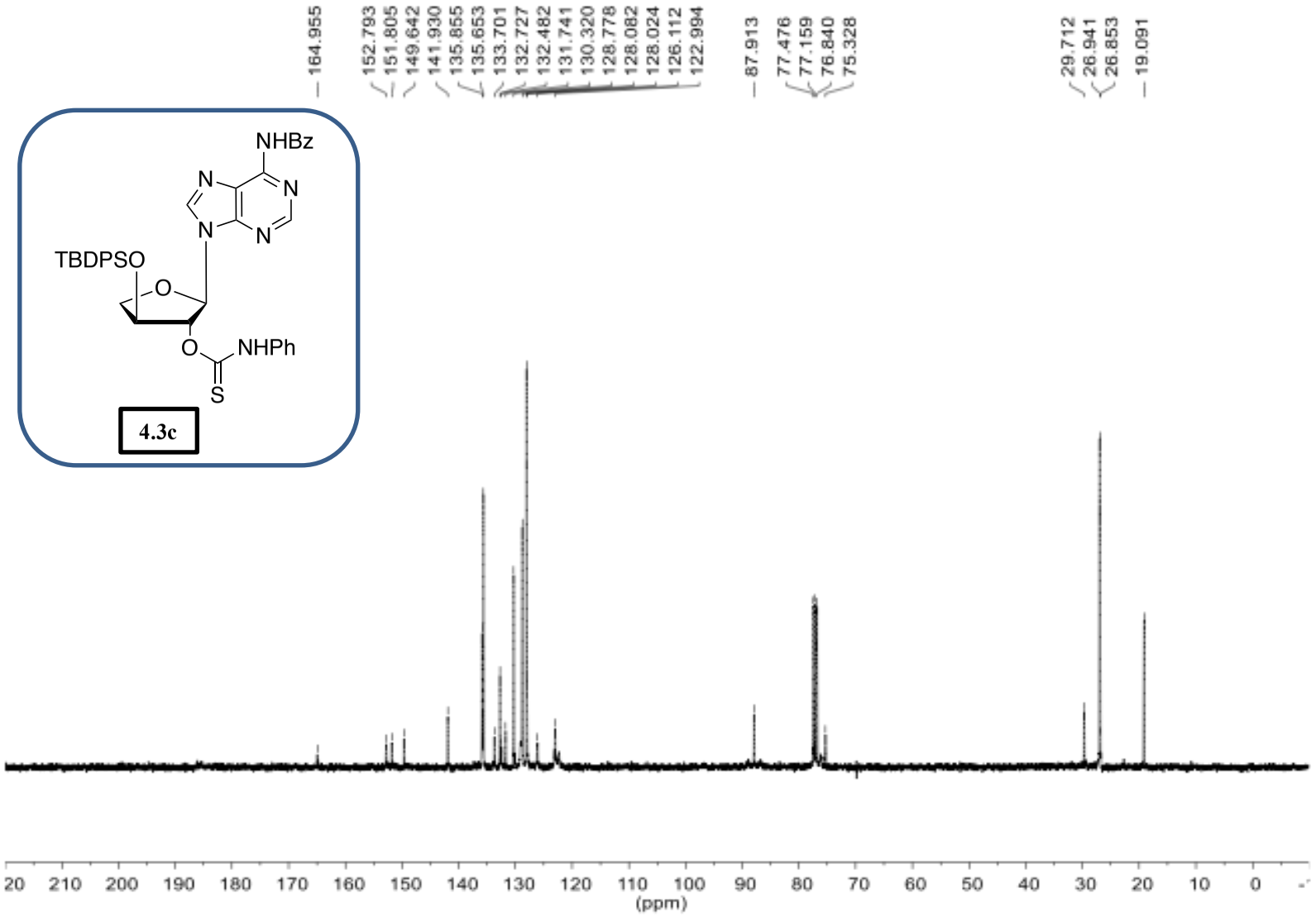
^{1}H NMR spectrum of compound **4.3b** (500 MHz,CD_3OD)



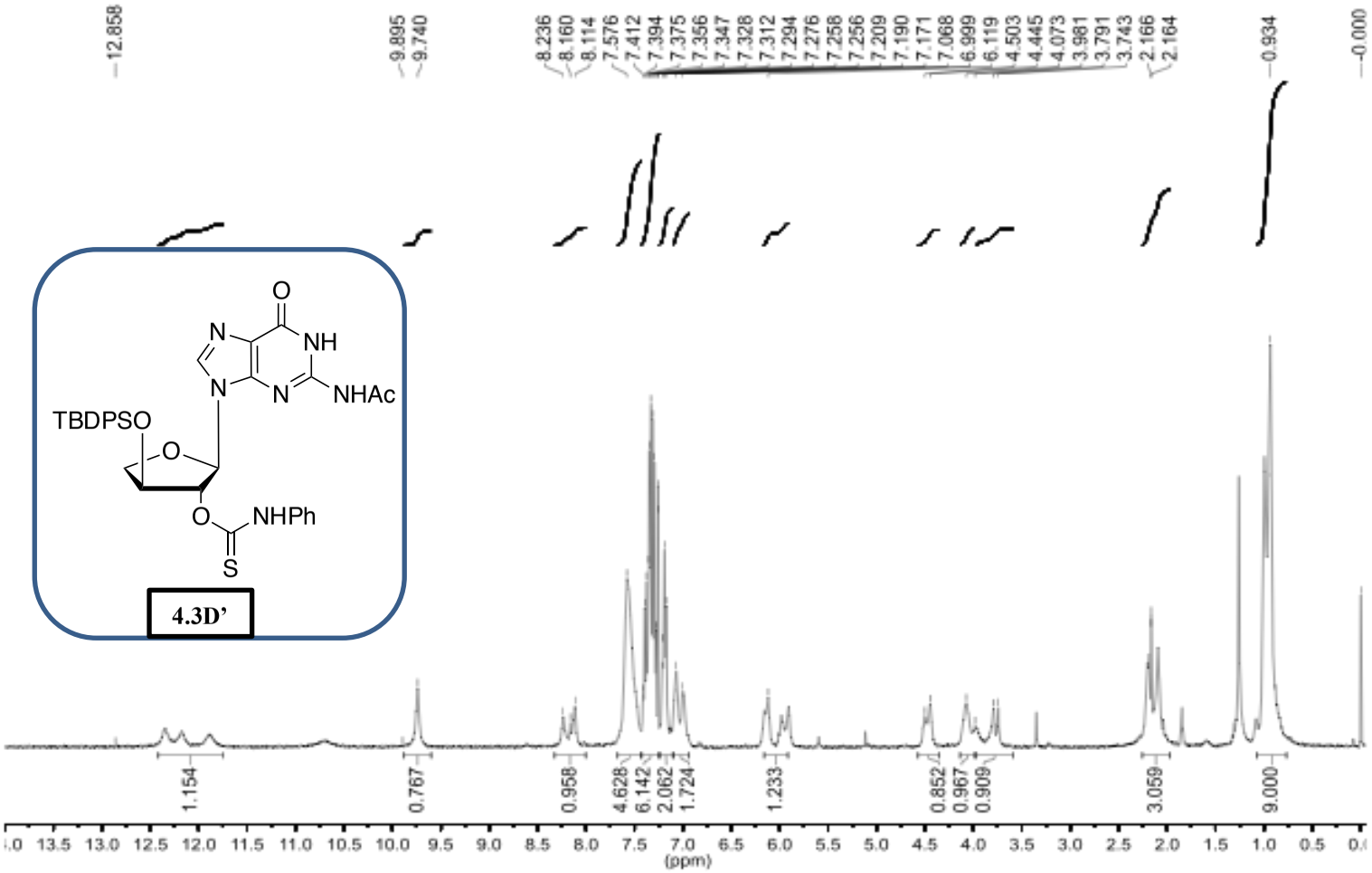
^{13}C NMR spectrum of compound **4.3b** (500 MHz,CD_3OD)



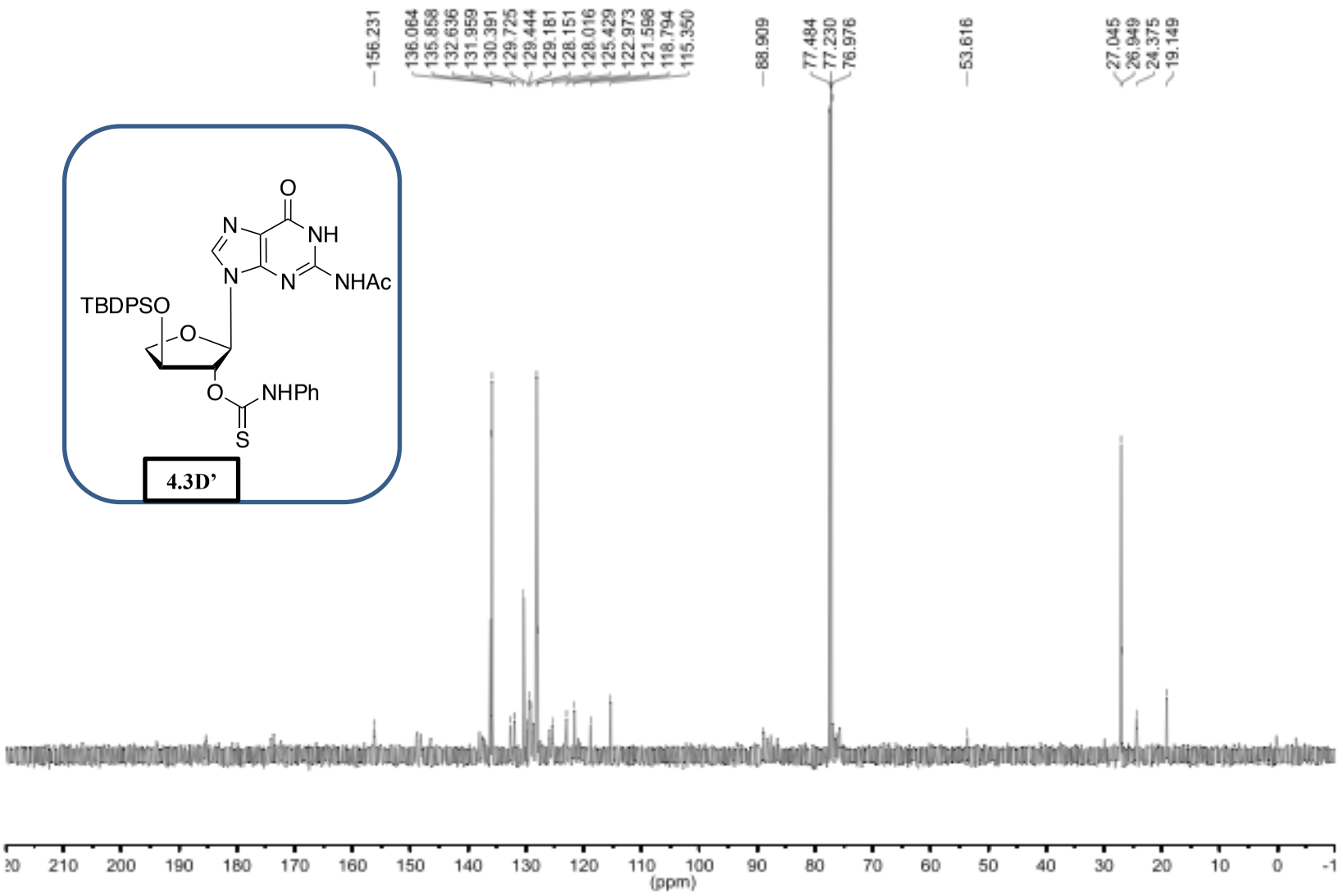
^{1}H NMR spectrum of compound **4.3c** (500 MHz,CD_3OD)



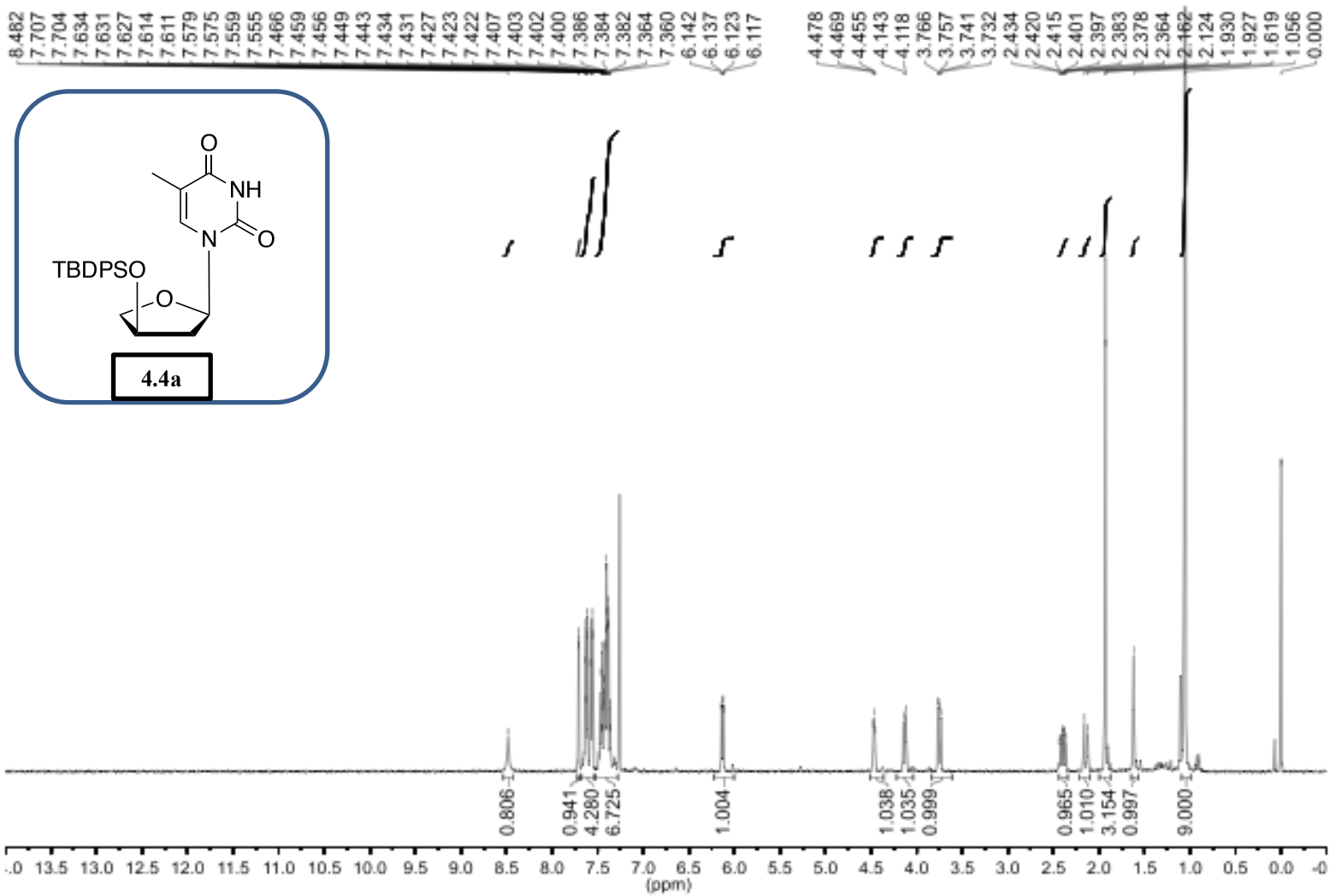
^{13}C NMR spectrum of compound **4.3c** (500 MHz,CD_3OD)



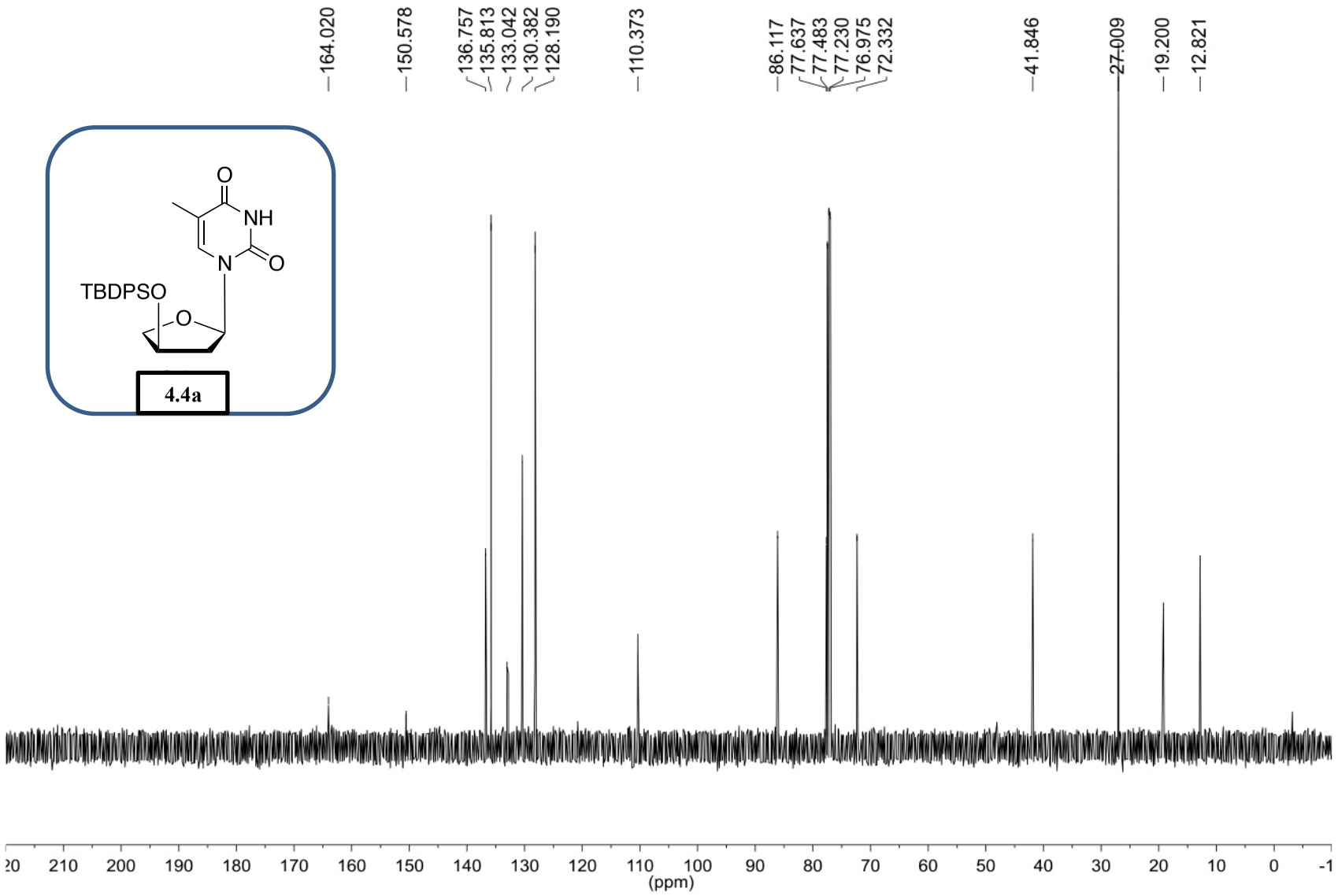
^{1}H NMR spectrum of compound **4.3d`** (500 MHz,CD_3OD)



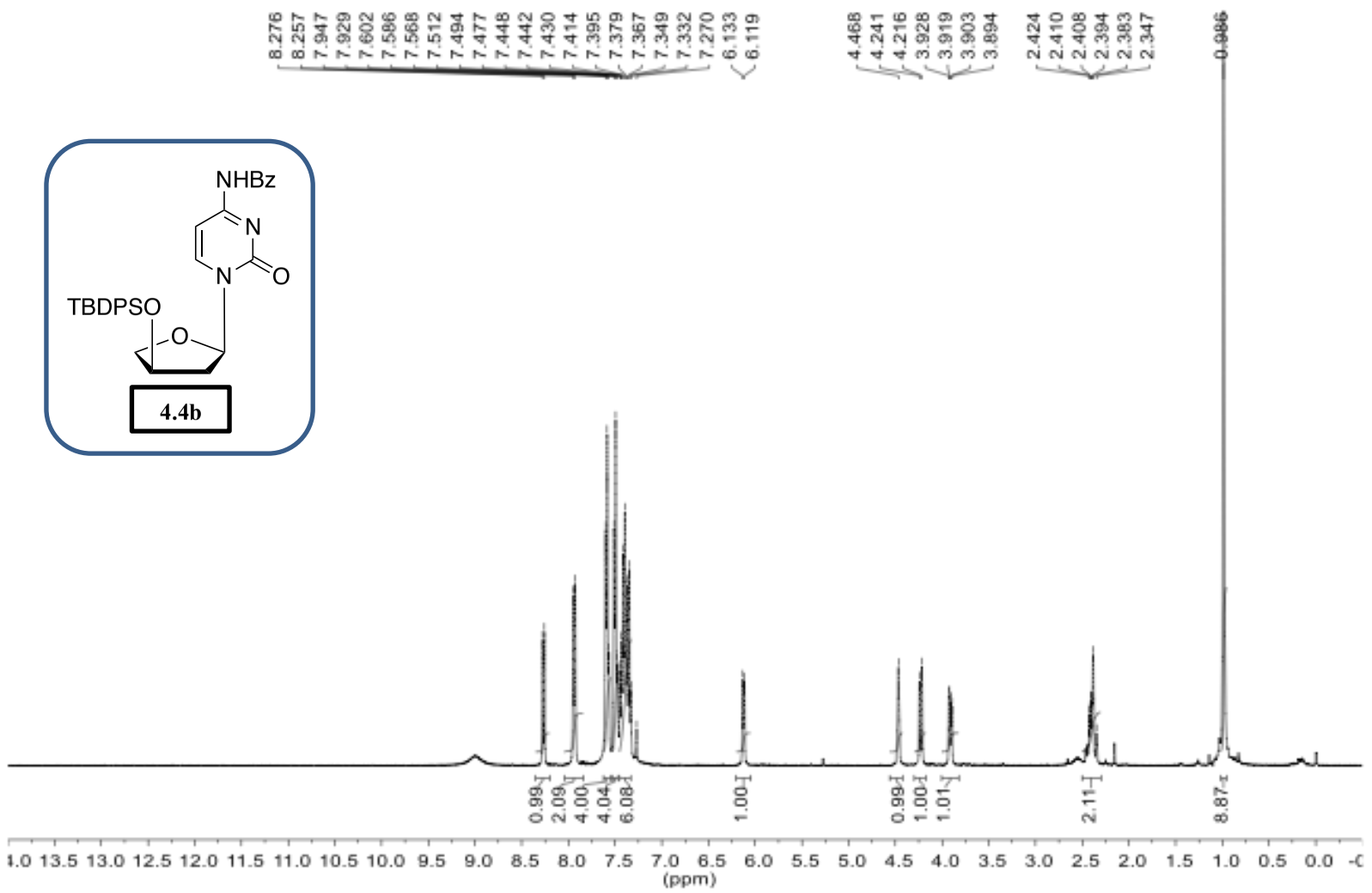
^{13}C NMR spectrum of compound **3.3d'** (500 MHz,CD_3OD)



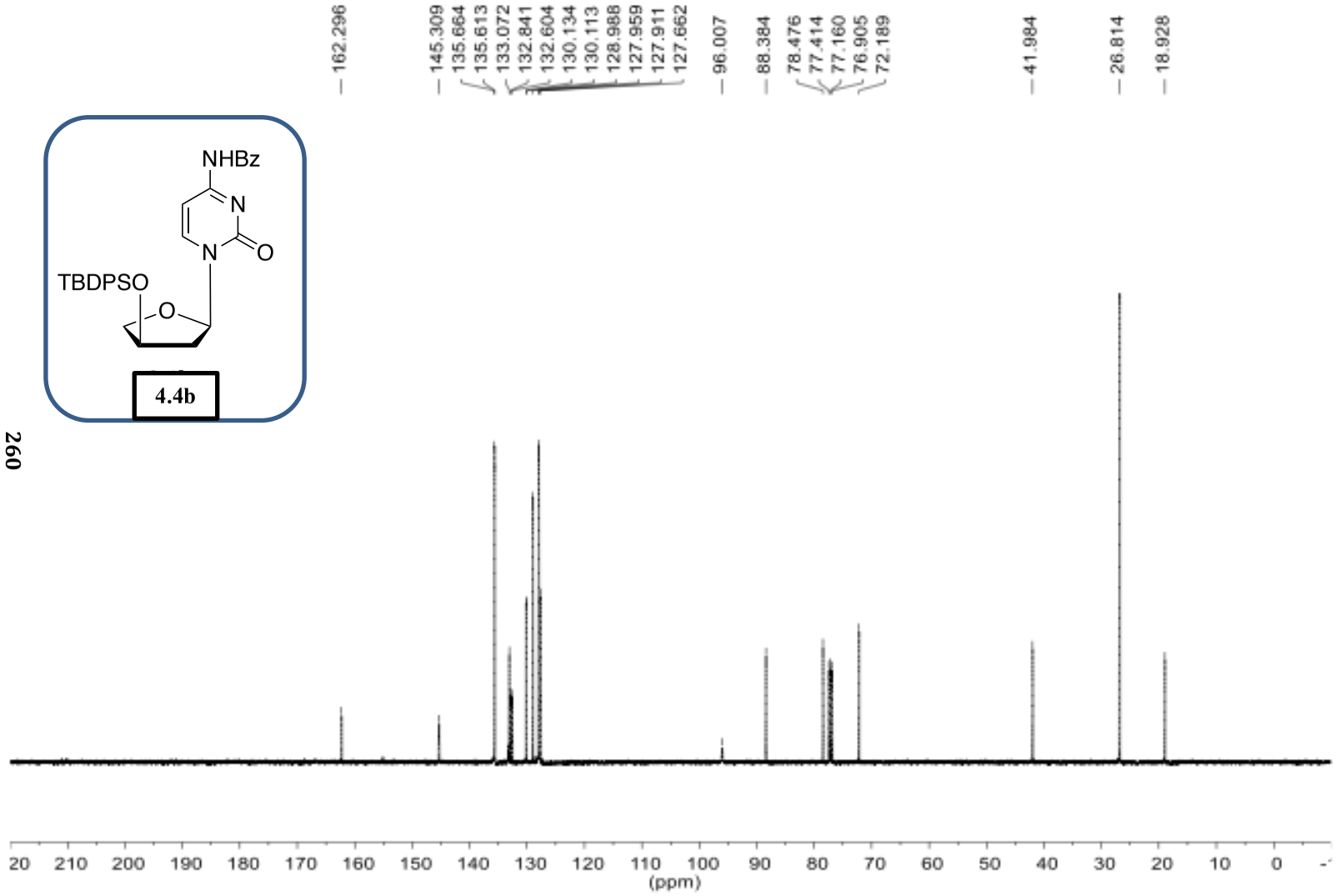
^{1}H NMR spectrum of compound **4.4a** (500 MHz,CD_3OD)



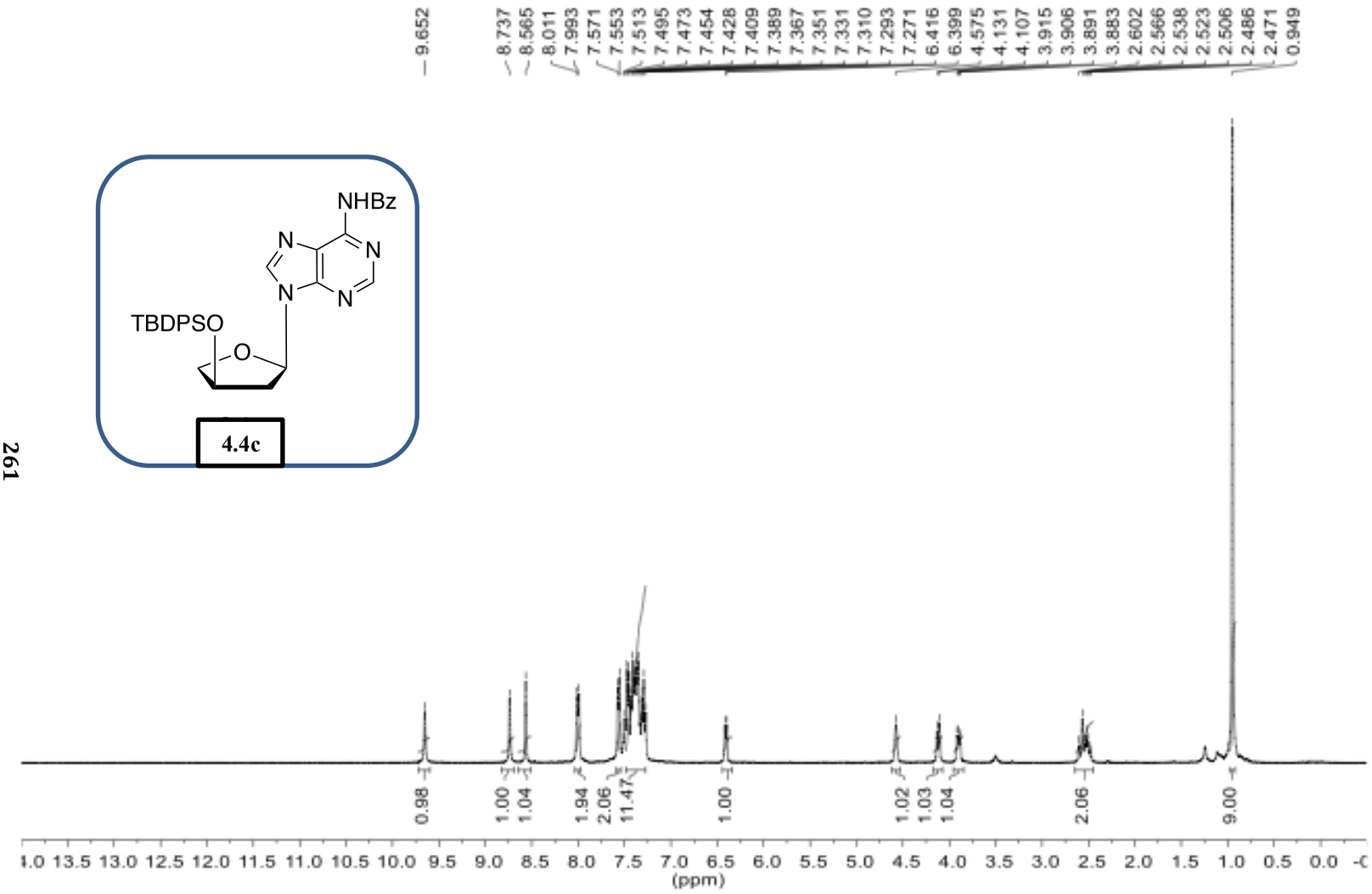
^{13}C NMR spectrum of compound **4.4a** (500 MHz,CD_3OD)



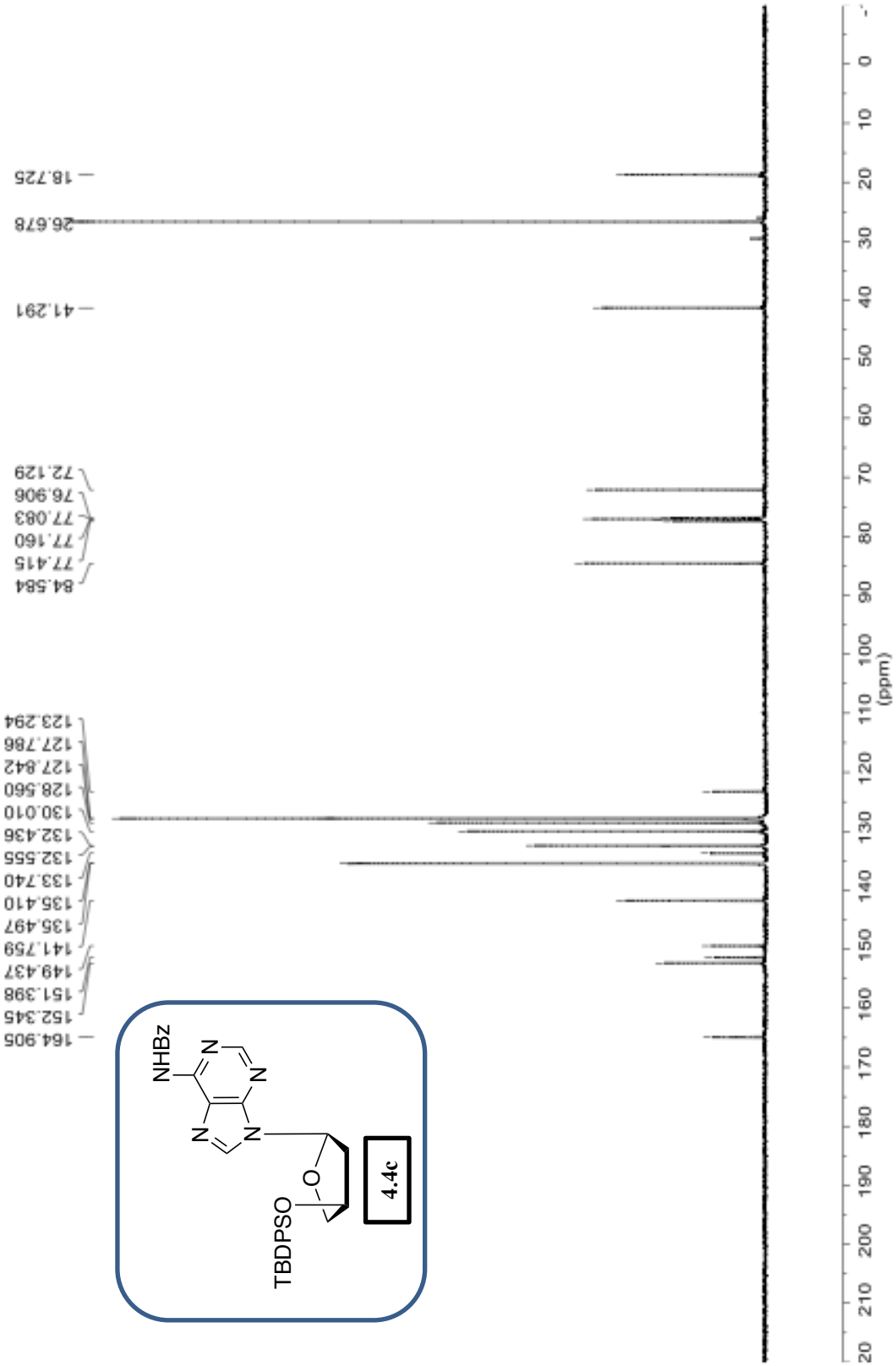
^{1}H NMR spectrum of compound **4.4b** (500 MHz,CD_3OD)



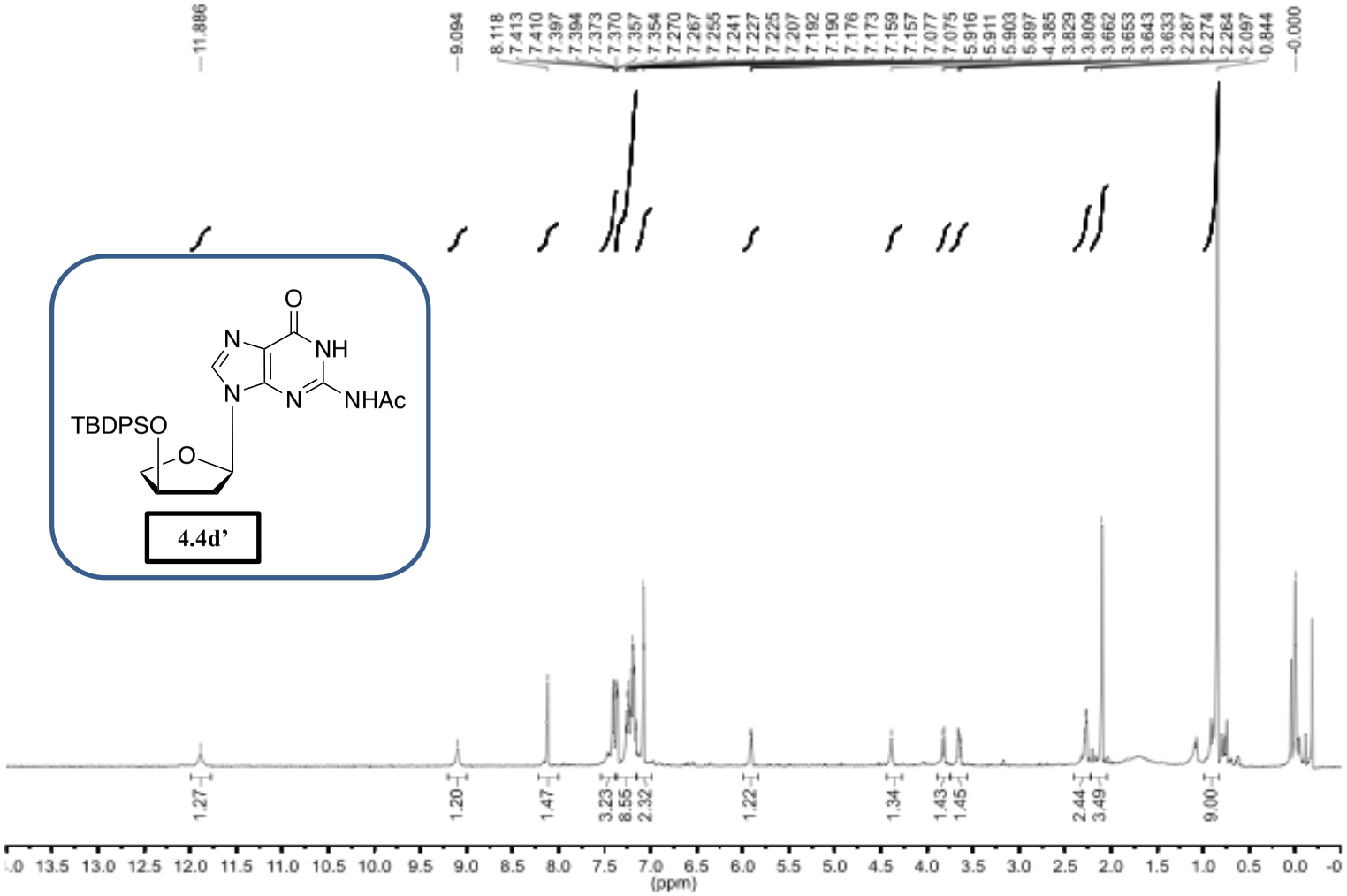
^{13}C NMR spectrum of compound **4.4b** (500 MHz,CD_3OD)



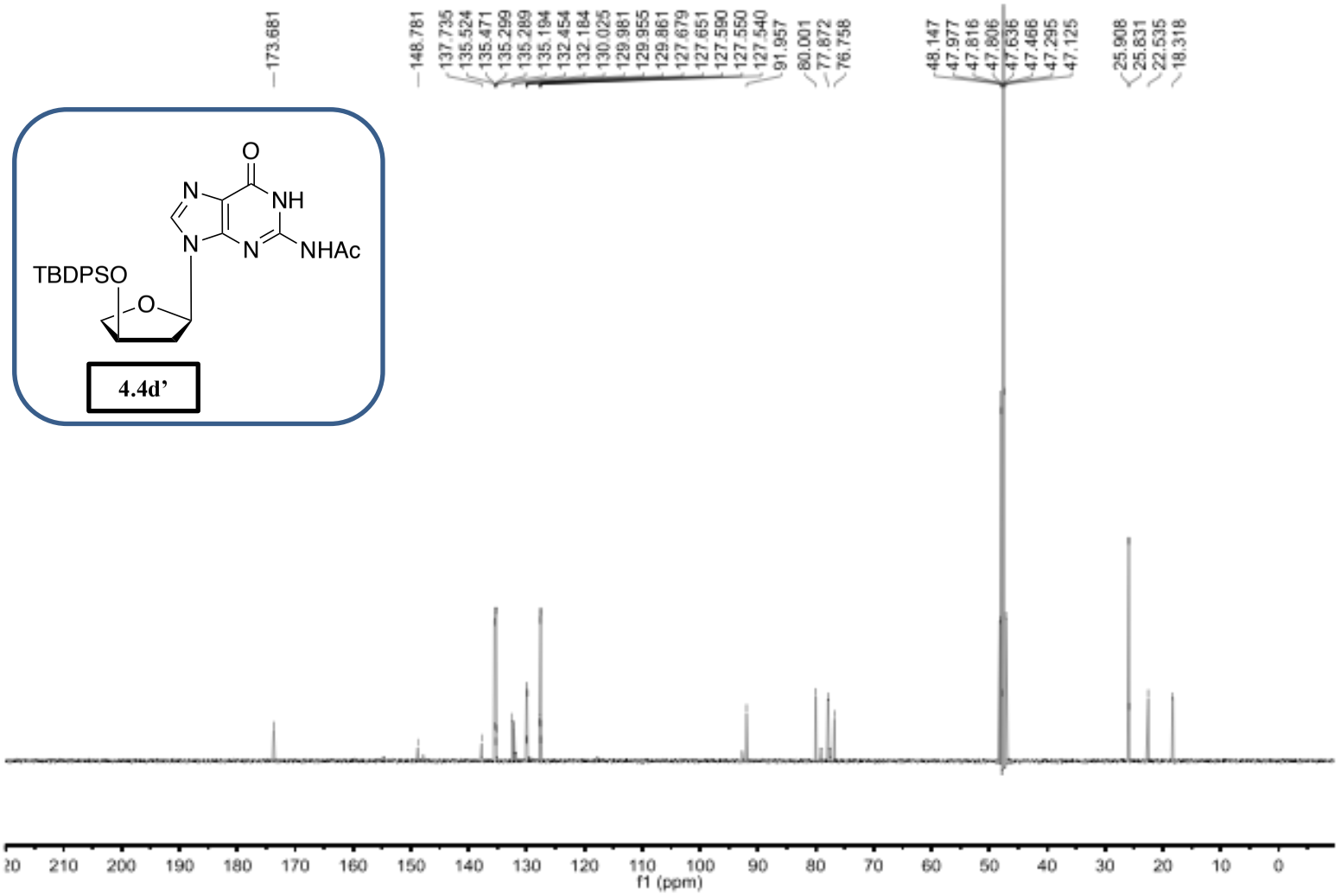
^{1}H NMR spectrum of compound **4.4c** (500 MHz,CD_3OD)



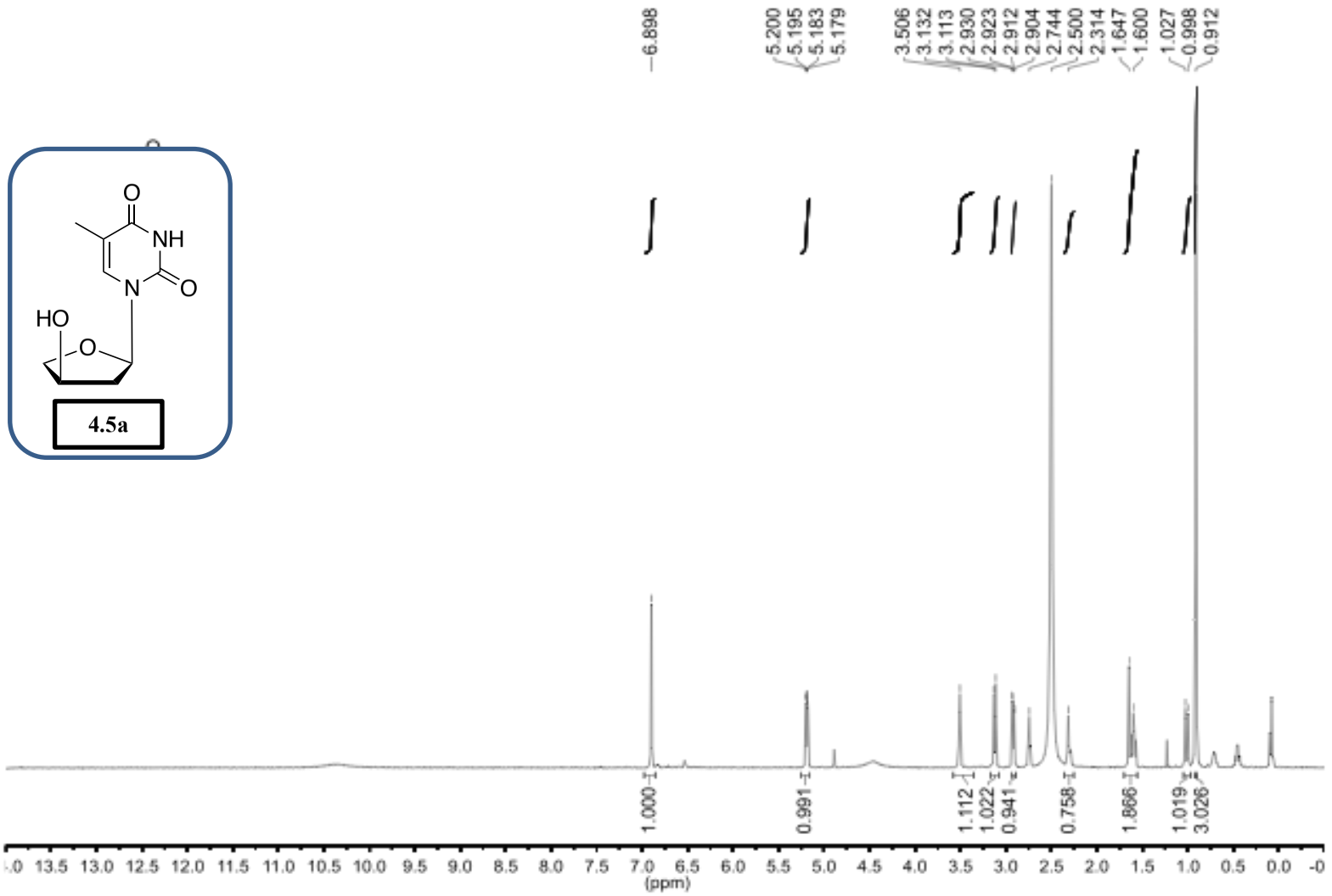
^{13}C NMR spectrum of compound **4.4c** (500 MHz,CD_3OD)



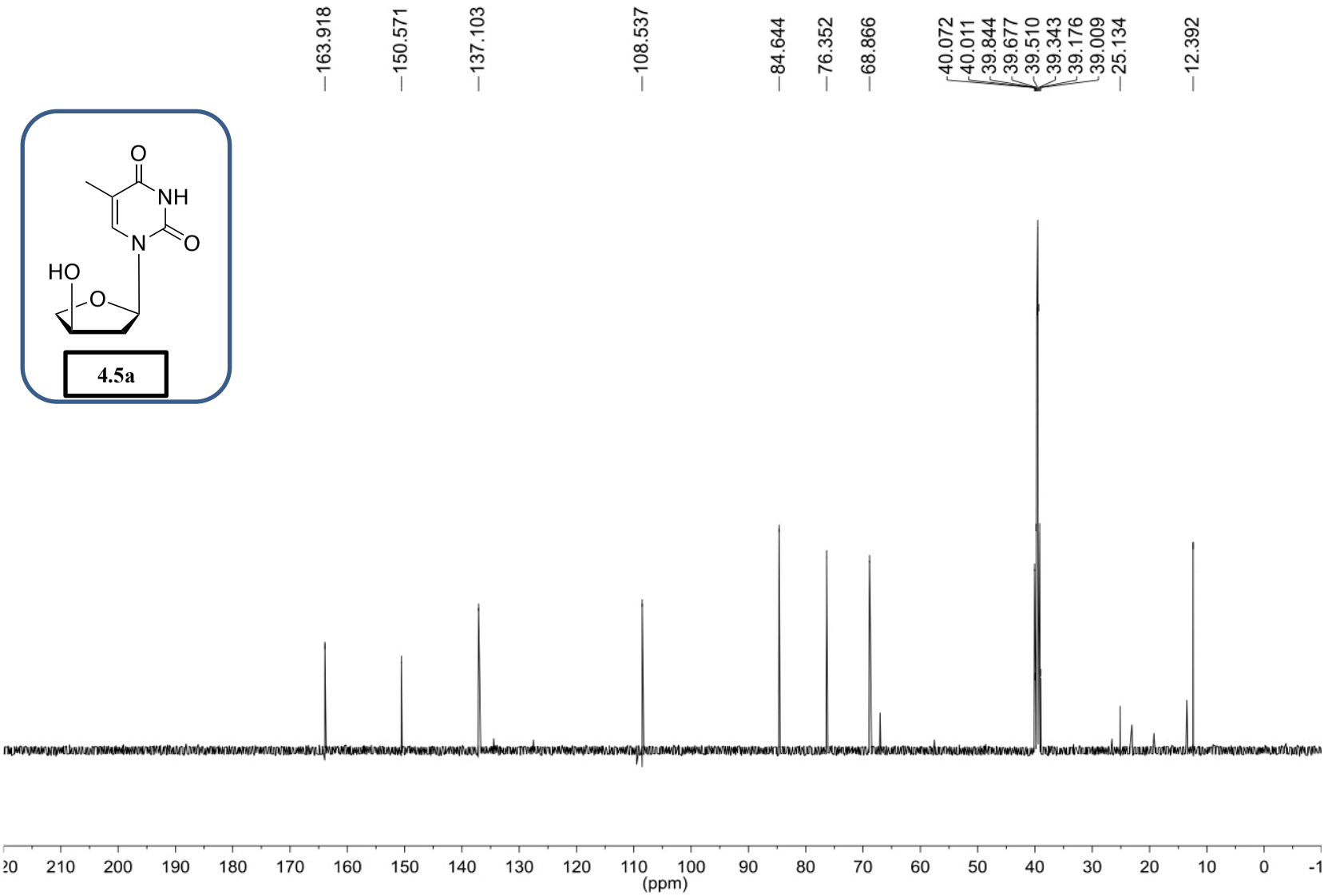
^{1}H NMR spectrum of compound **4.4d'** (500 MHz,CD_3OD)



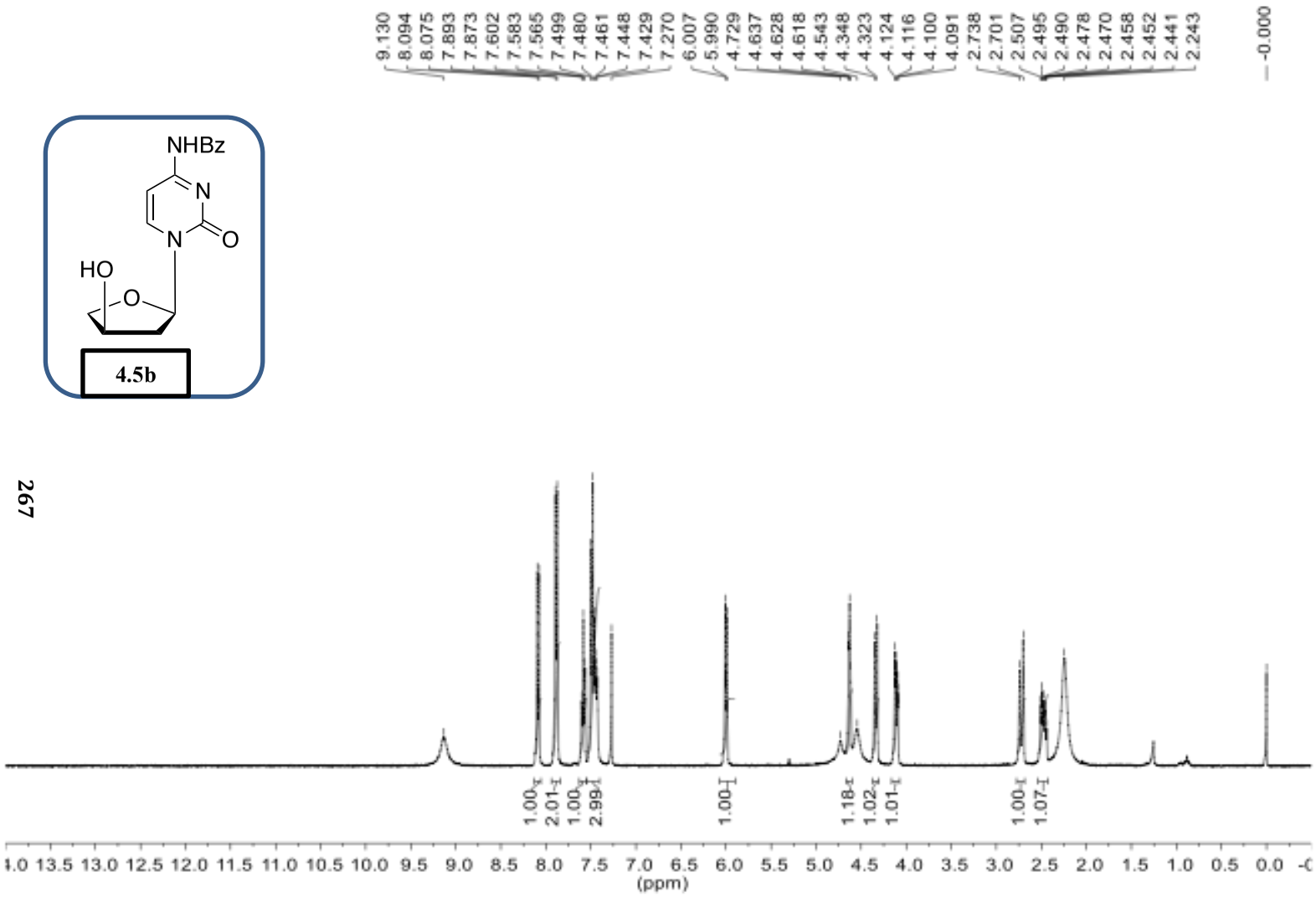
^{13}C NMR spectrum of compound **3.4d`** (500 MHz,CD_3OD)



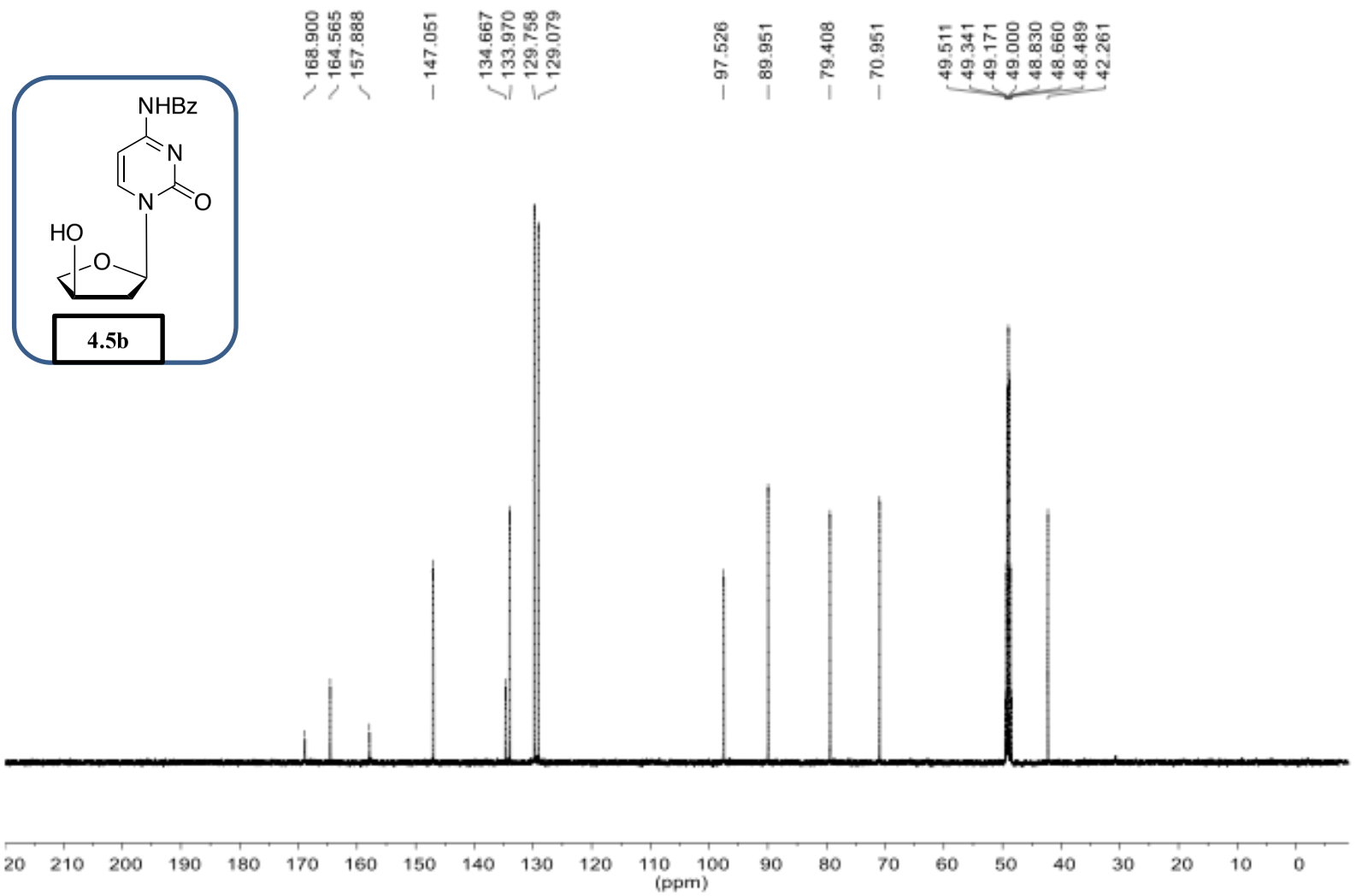
^{1}H NMR spectrum of compound **4.5a** (500 MHz,CD_3OD)



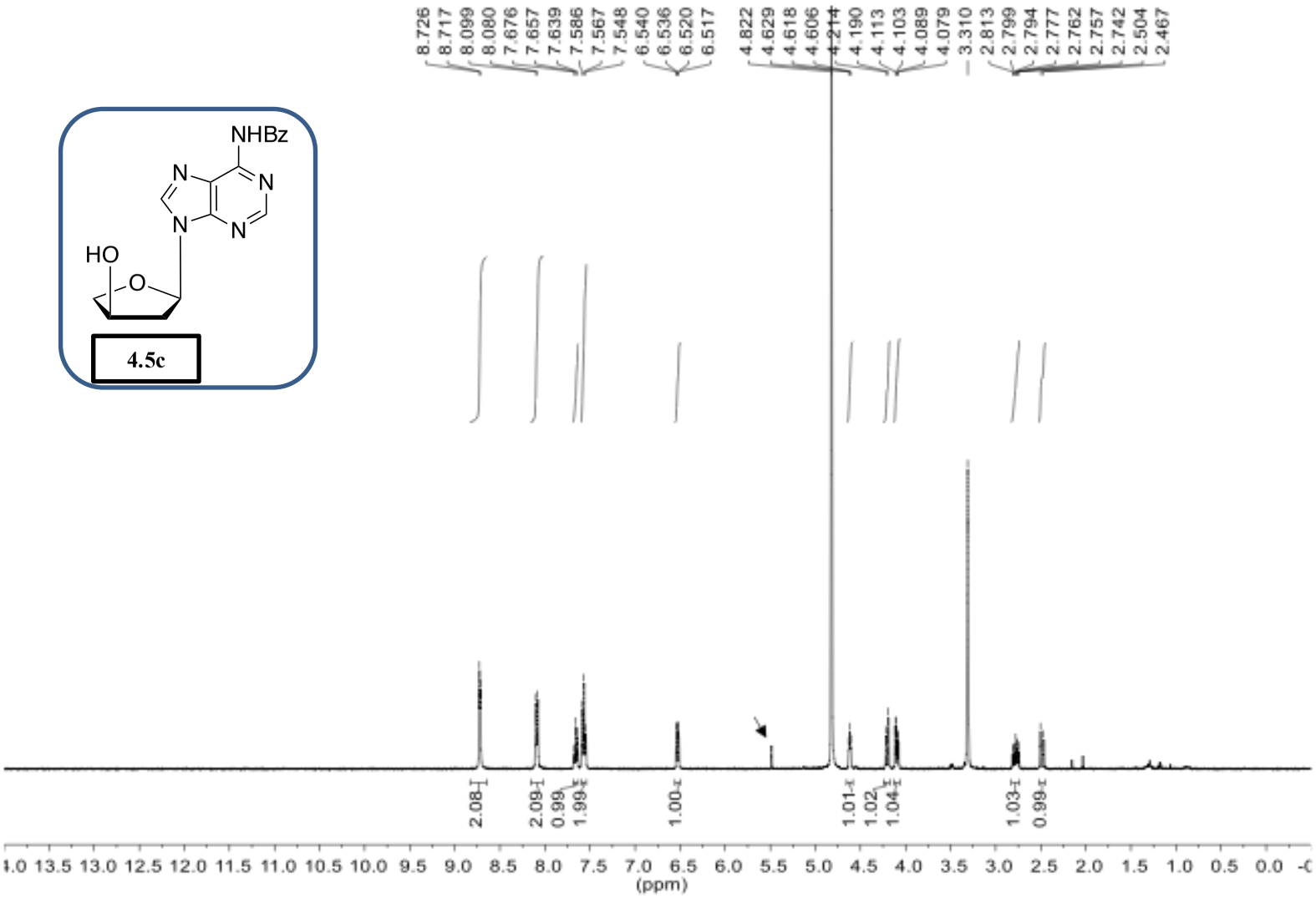
^{13}C NMR spectrum of compound **4.5a** (500 MHz,CD_3OD)



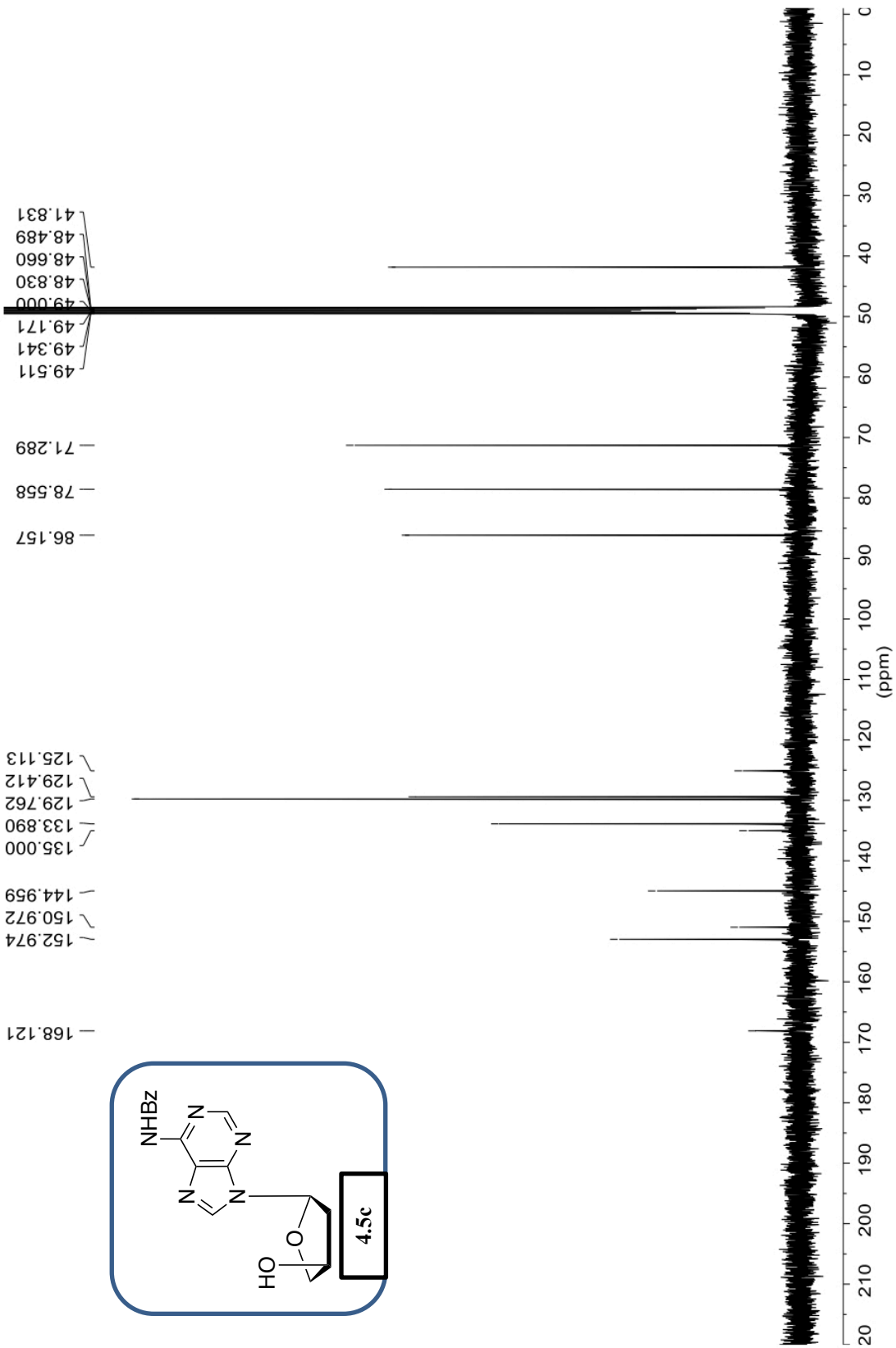
^{1}H NMR spectrum of compound **4.5b** (500 MHz,CD_3OD)



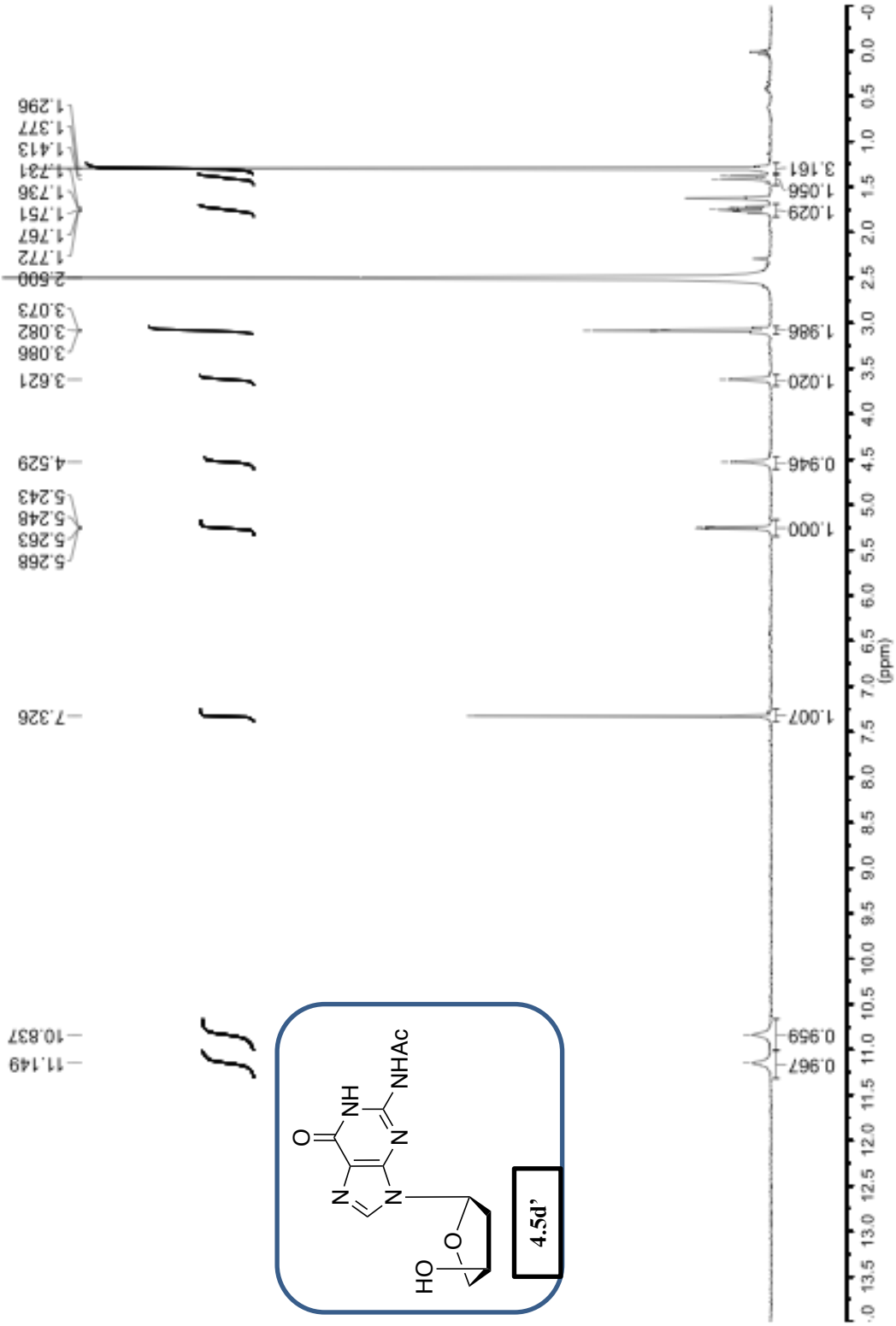
^{13}C NMR spectrum of compound **4.5b** (500 MHz,CD_3OD)



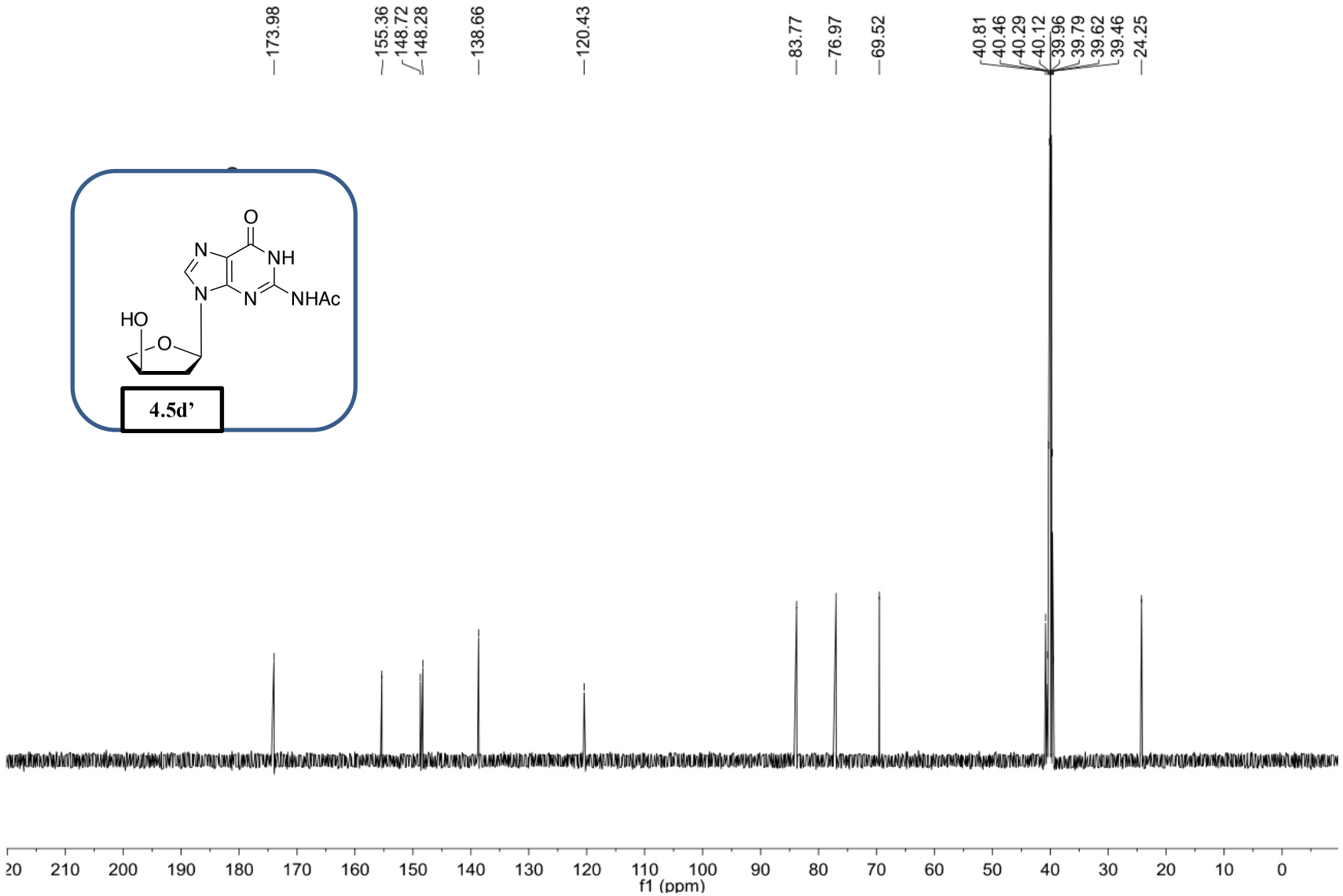
^{1}H NMR spectrum of compound **4.5c** (500 MHz,CD_3OD)



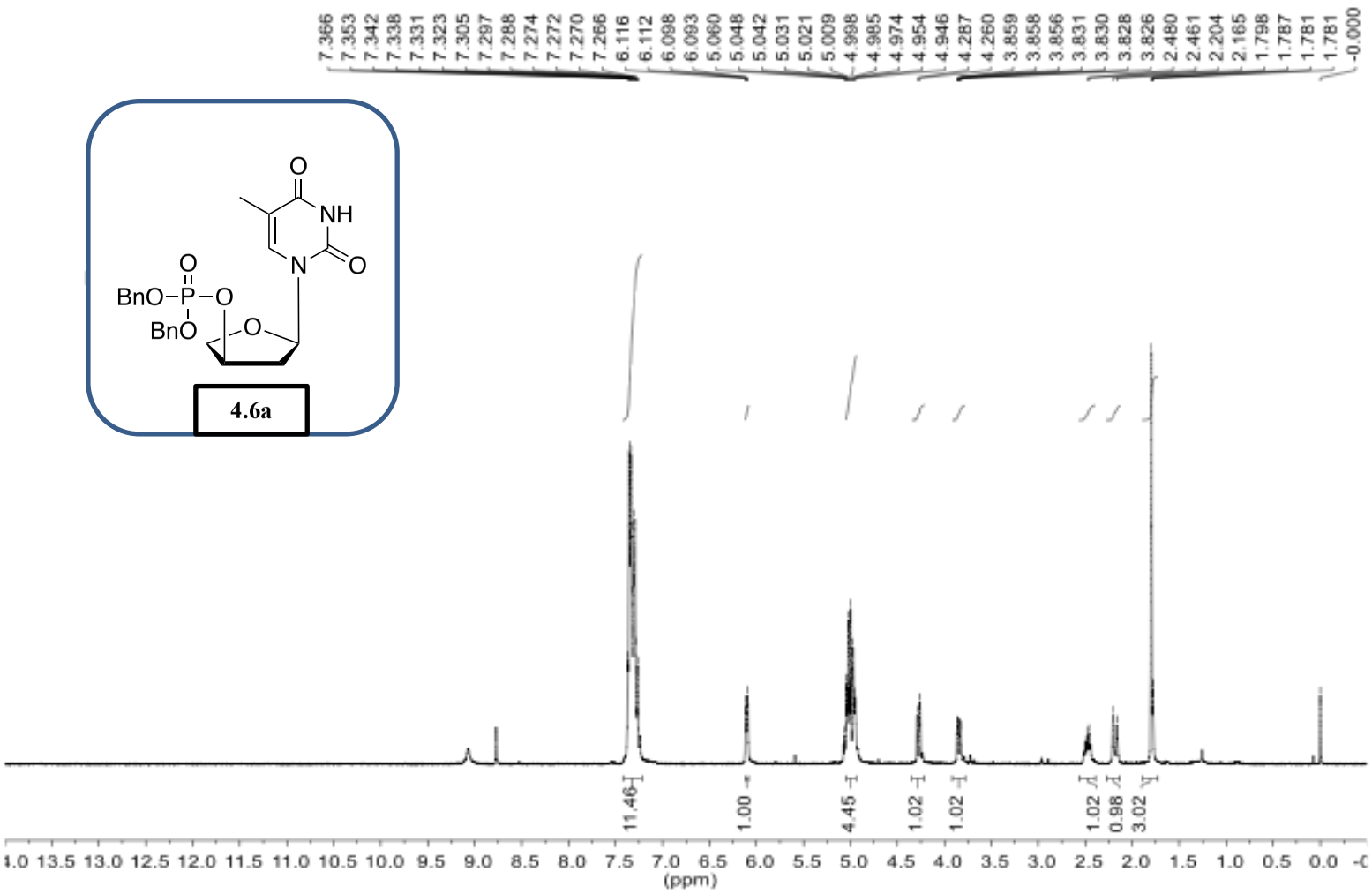
^{13}C NMR spectrum of compound **4.5c** (500 MHz,CD_3OD)



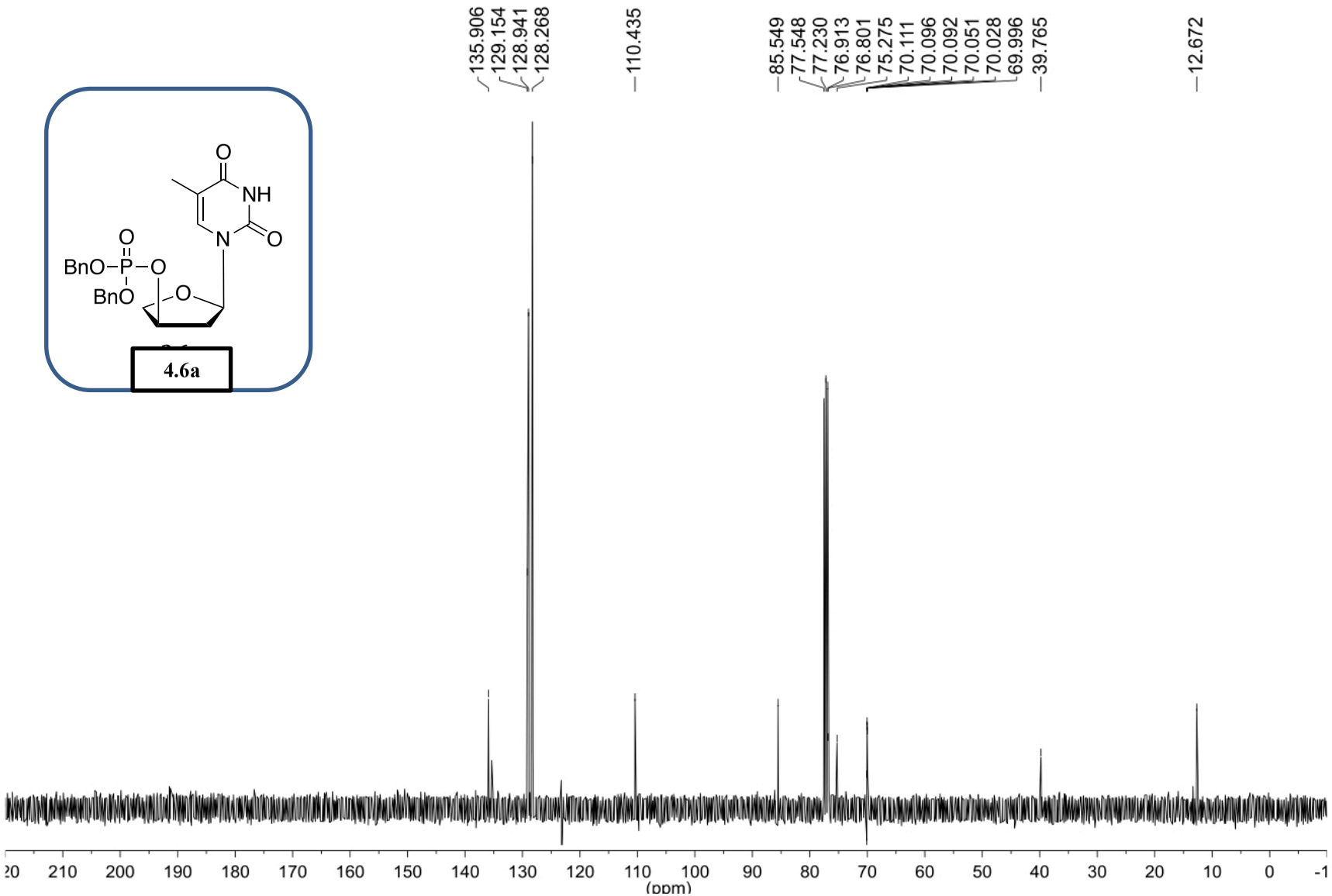
^{1}H NMR spectrum of compound **4.5d-** (500 MHz,CD_3OD)



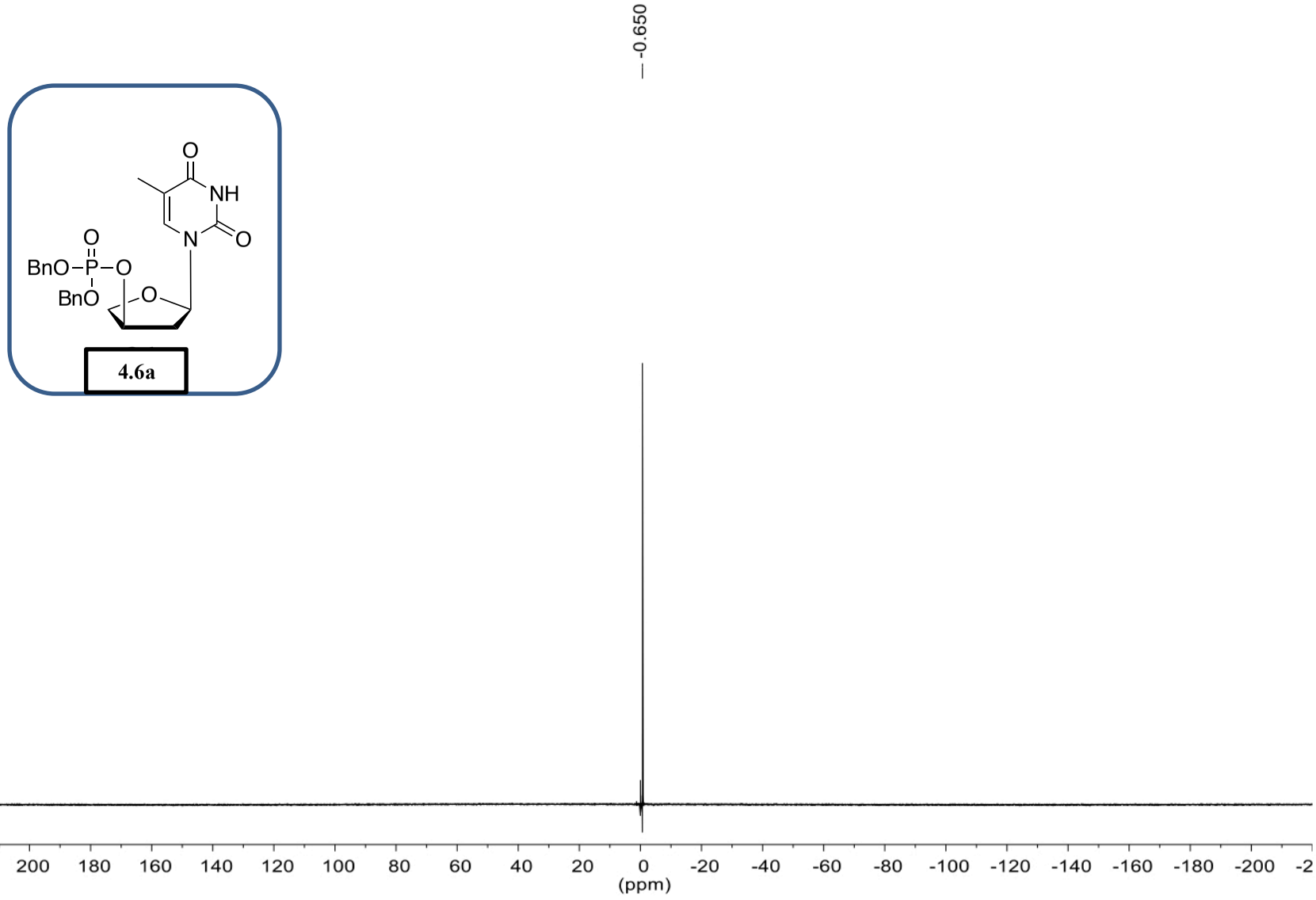
^{13}C NMR spectrum of compound **4.5d'** (500 MHz,CD_3OD)



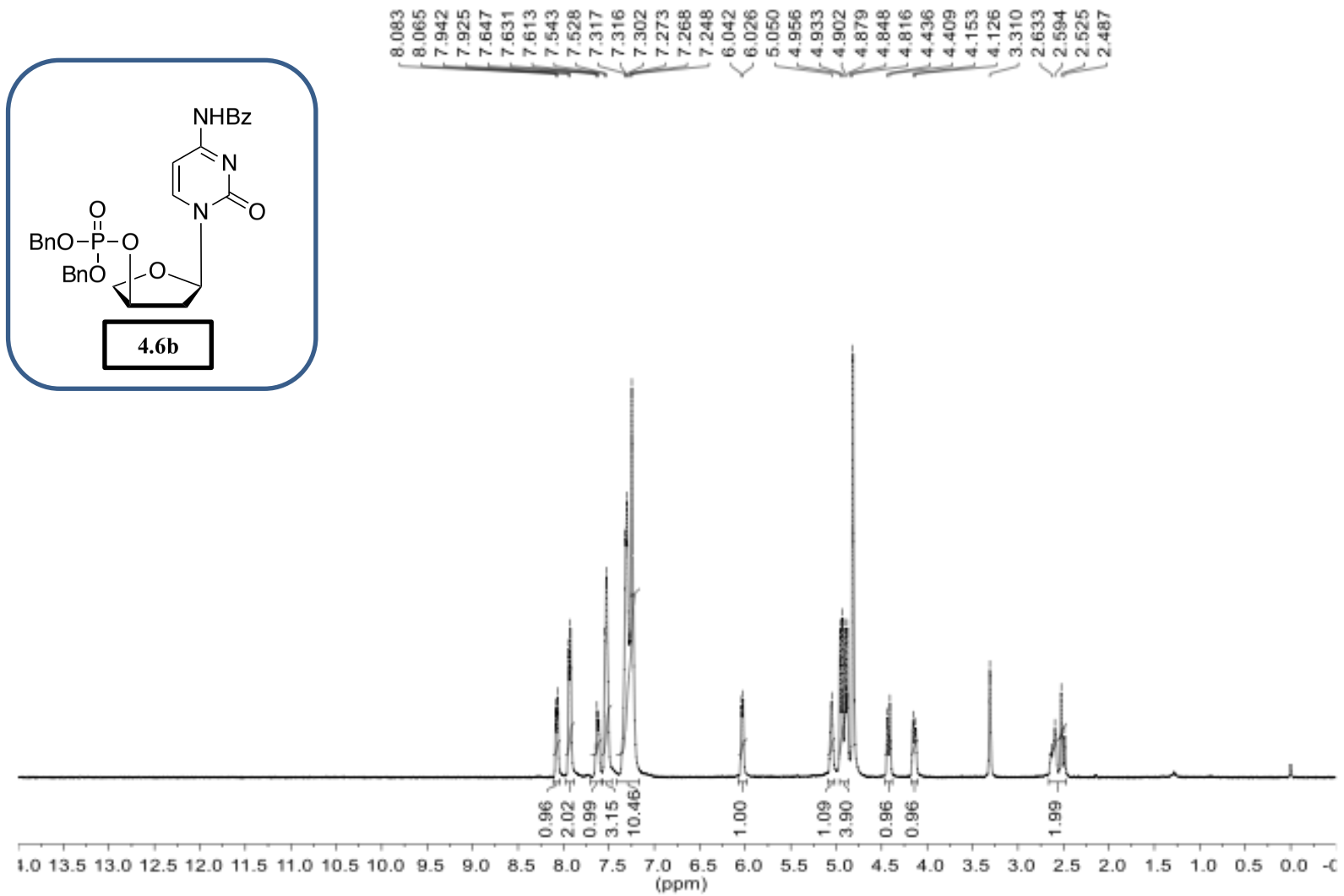
^{1}H NMR spectrum of compound **4.6a** (500 MHz,CD_3OD)



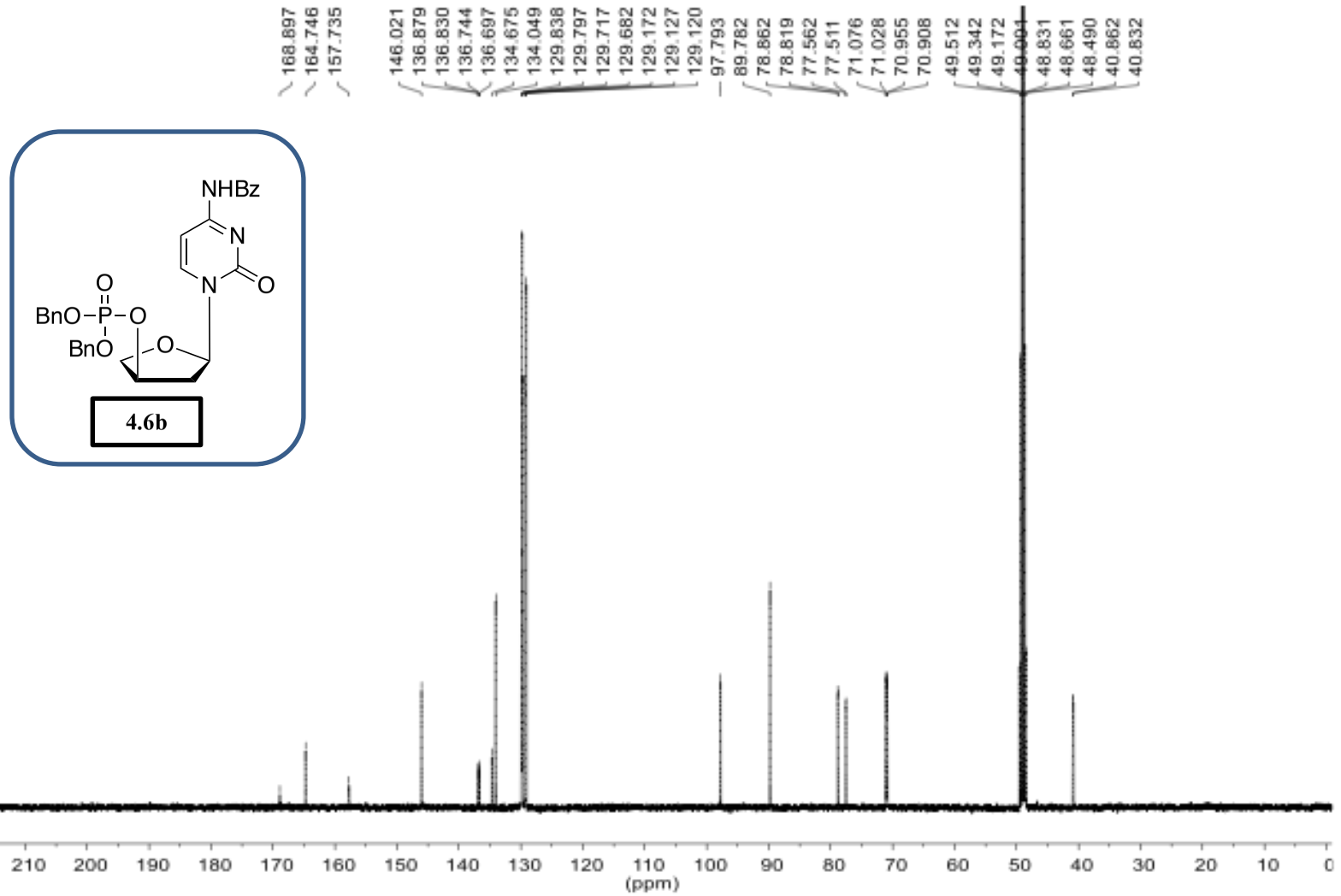
^{13}C NMR spectrum of compound **3.6a** (500 MHz,CD_3OD)



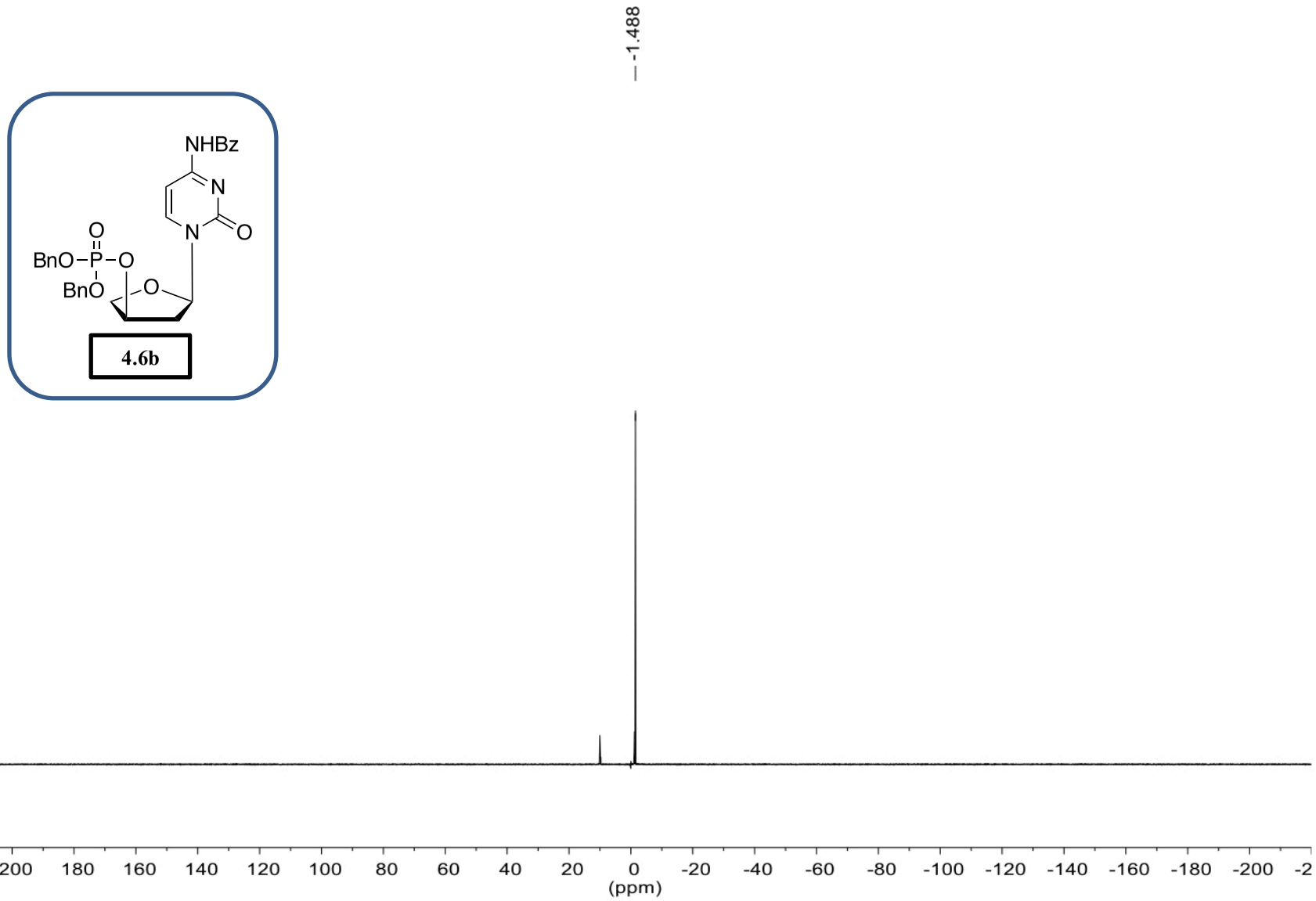
^{11}P NMR spectrum of compound **4.6a** (500 MHz,CD_3OD)



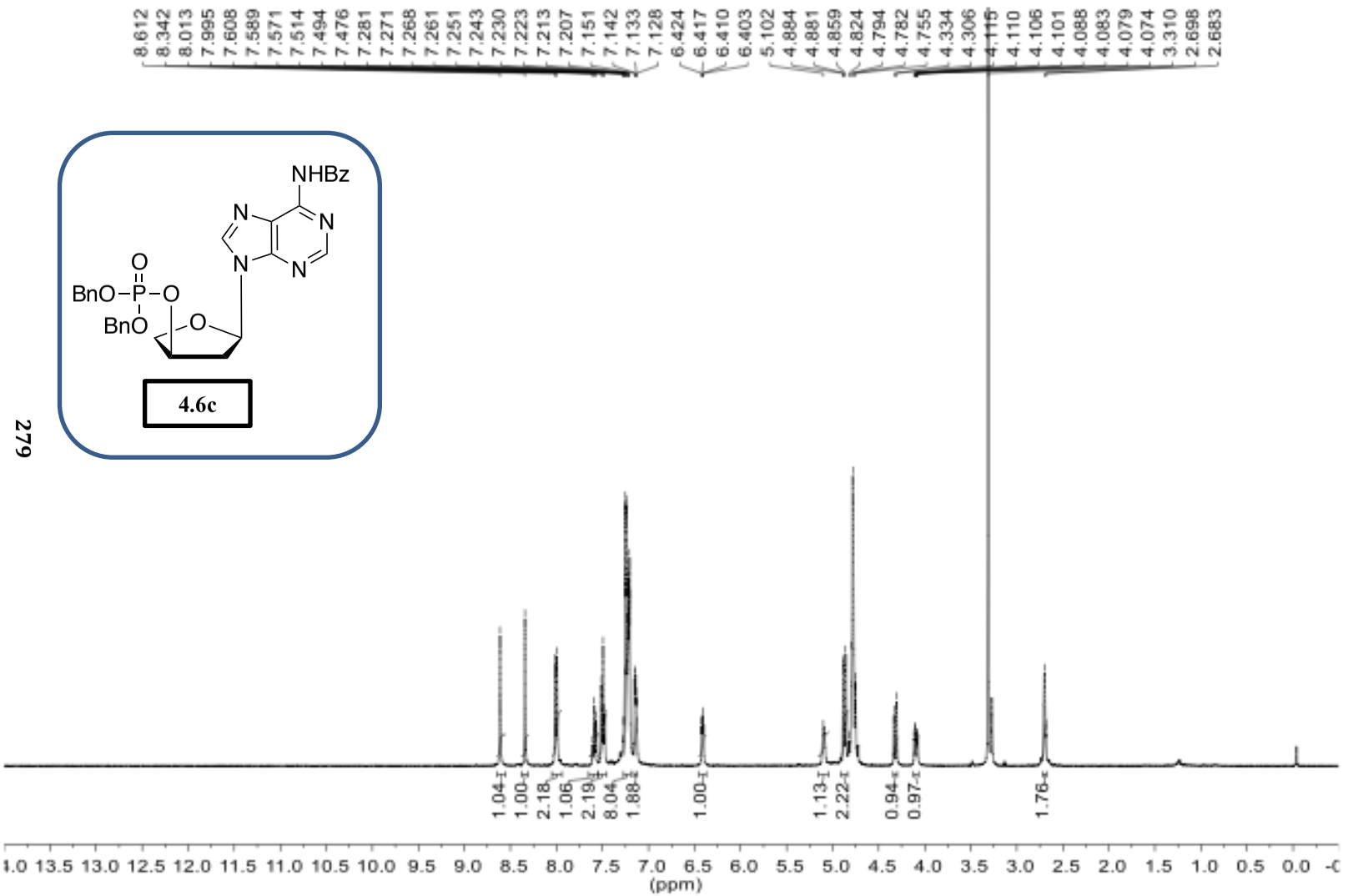
^{1}H NMR spectrum of compound **4.6b** (500 MHz,CD_3OD)



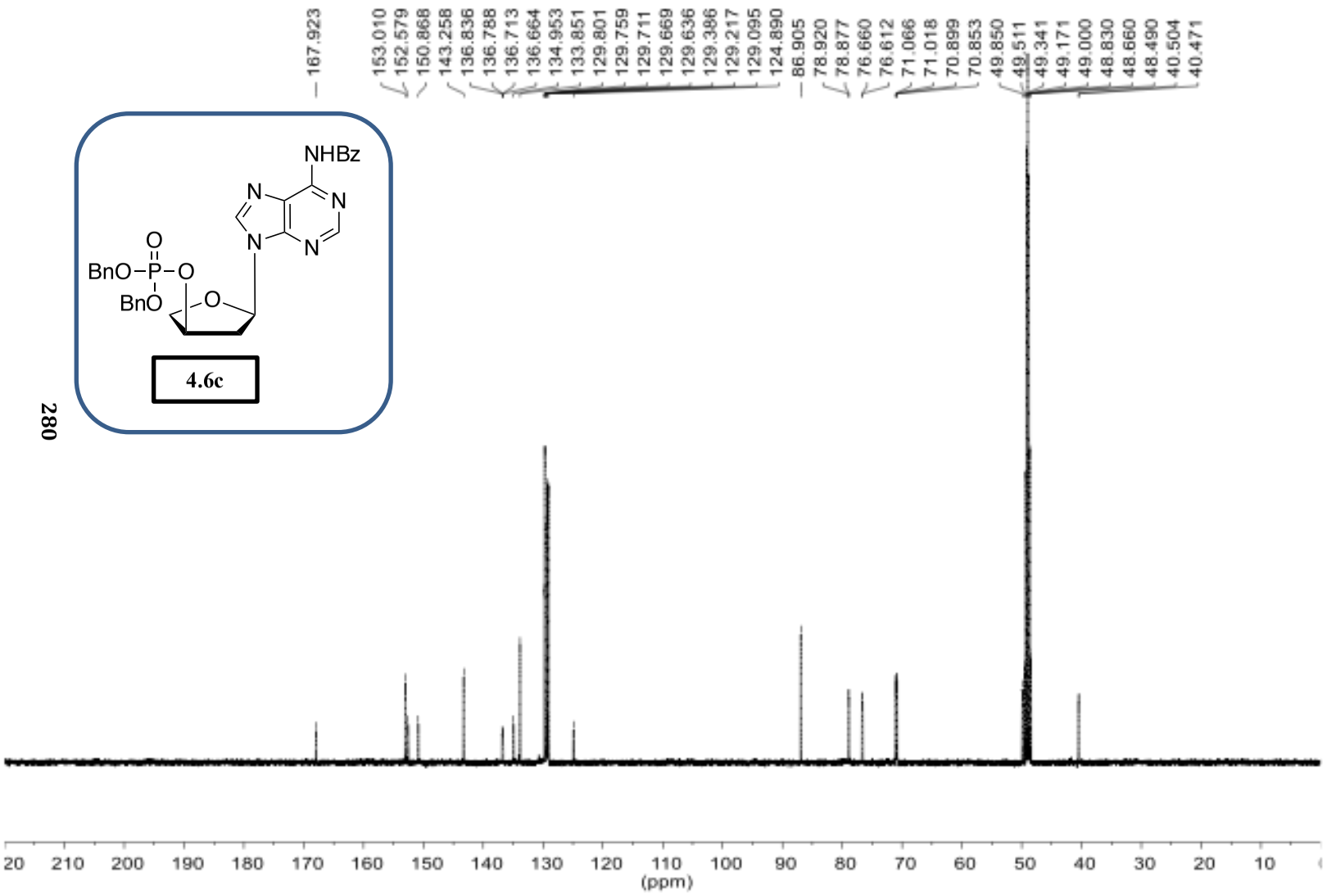
^{13}C NMR spectrum of compound **4.6b** (500 MHz,CD_3OD)



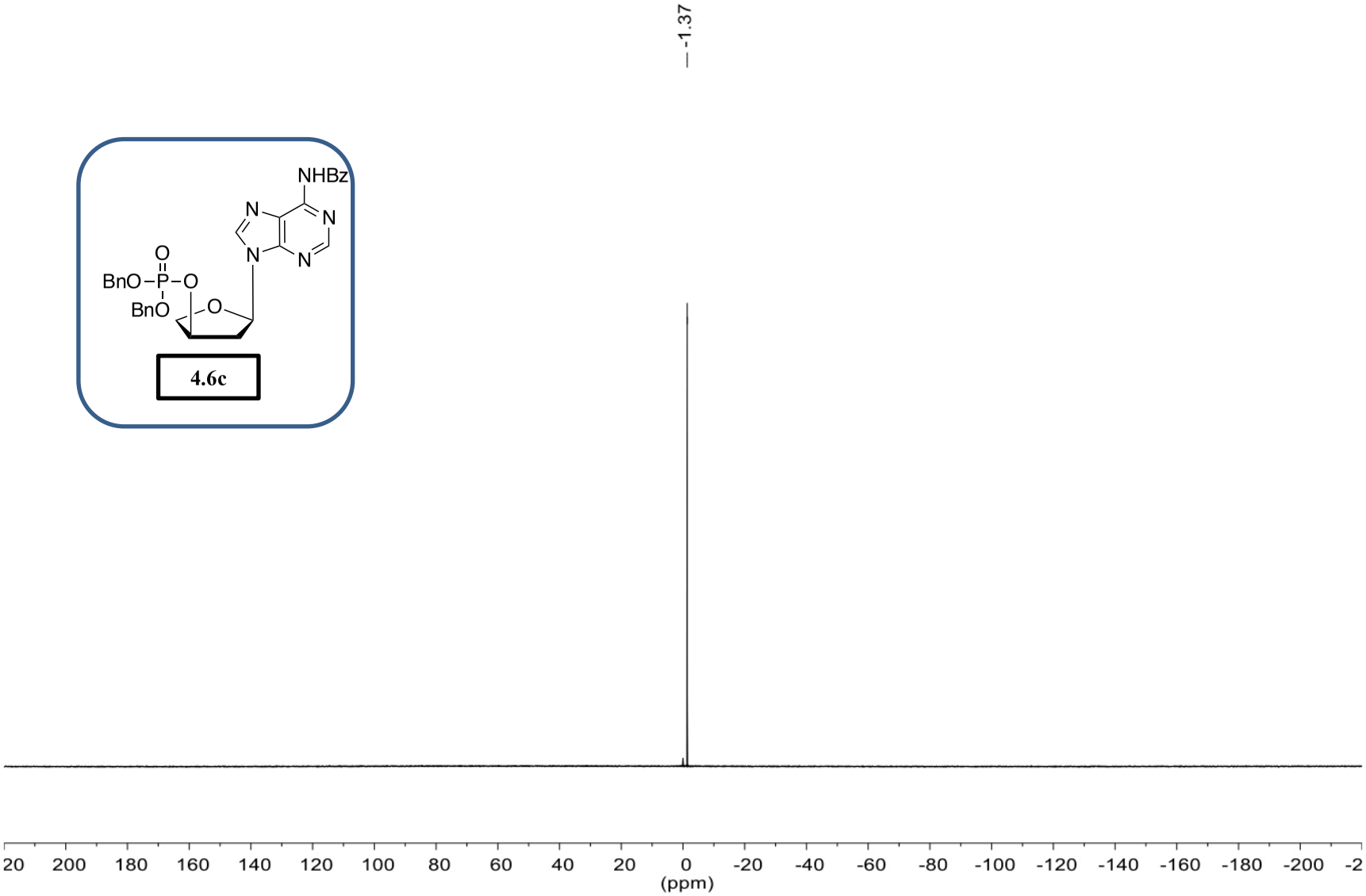
^{31}P NMR spectrum of compound **4.6b** (500 MHz,CD_3OD)



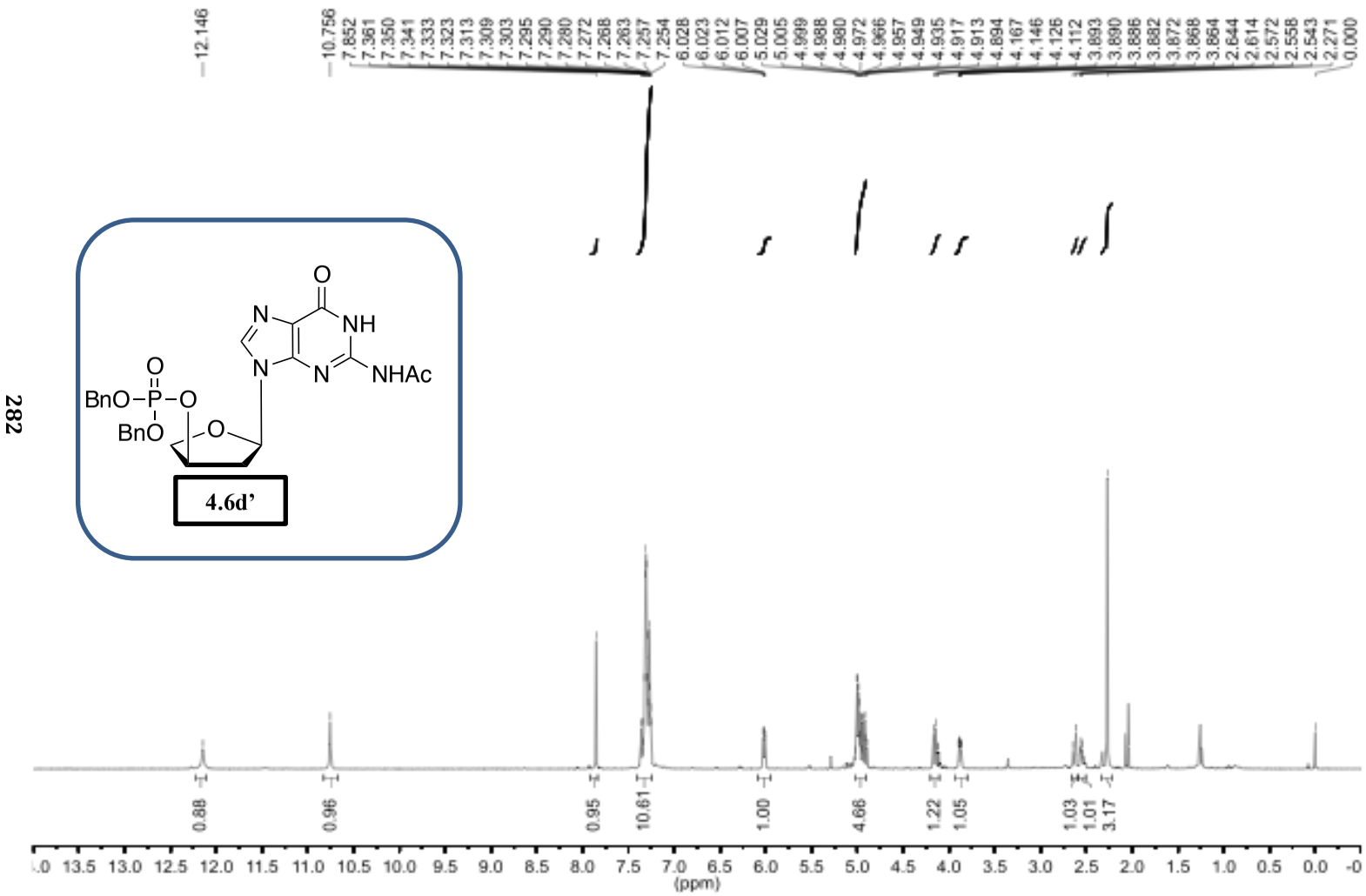
^{1}H NMR spectrum of compound **4.6c** (500 MHz,CD_3OD)



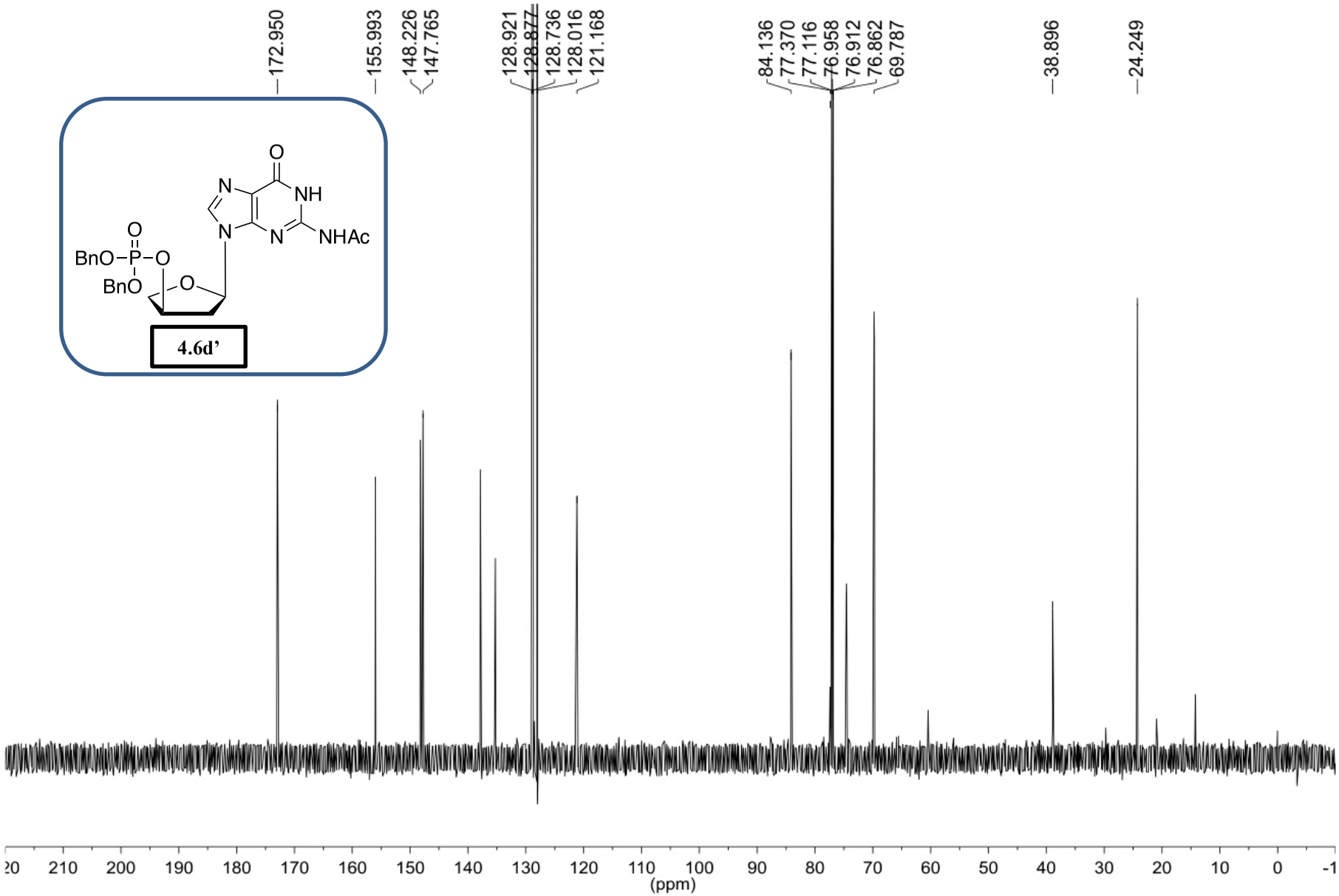
^{13} CNMR spectrum of compound **4.6c** (500 MHz,CD_3OD)



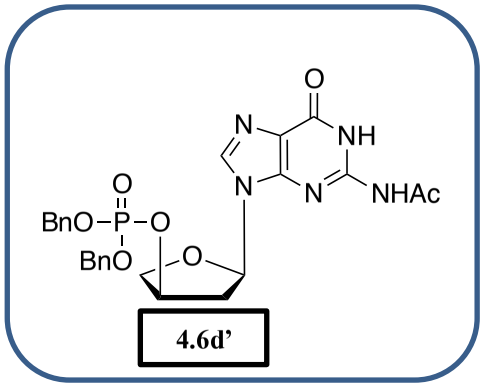
[31] PNMR spectrum of compound **4.6c** (500 MHz,CD_3OD)



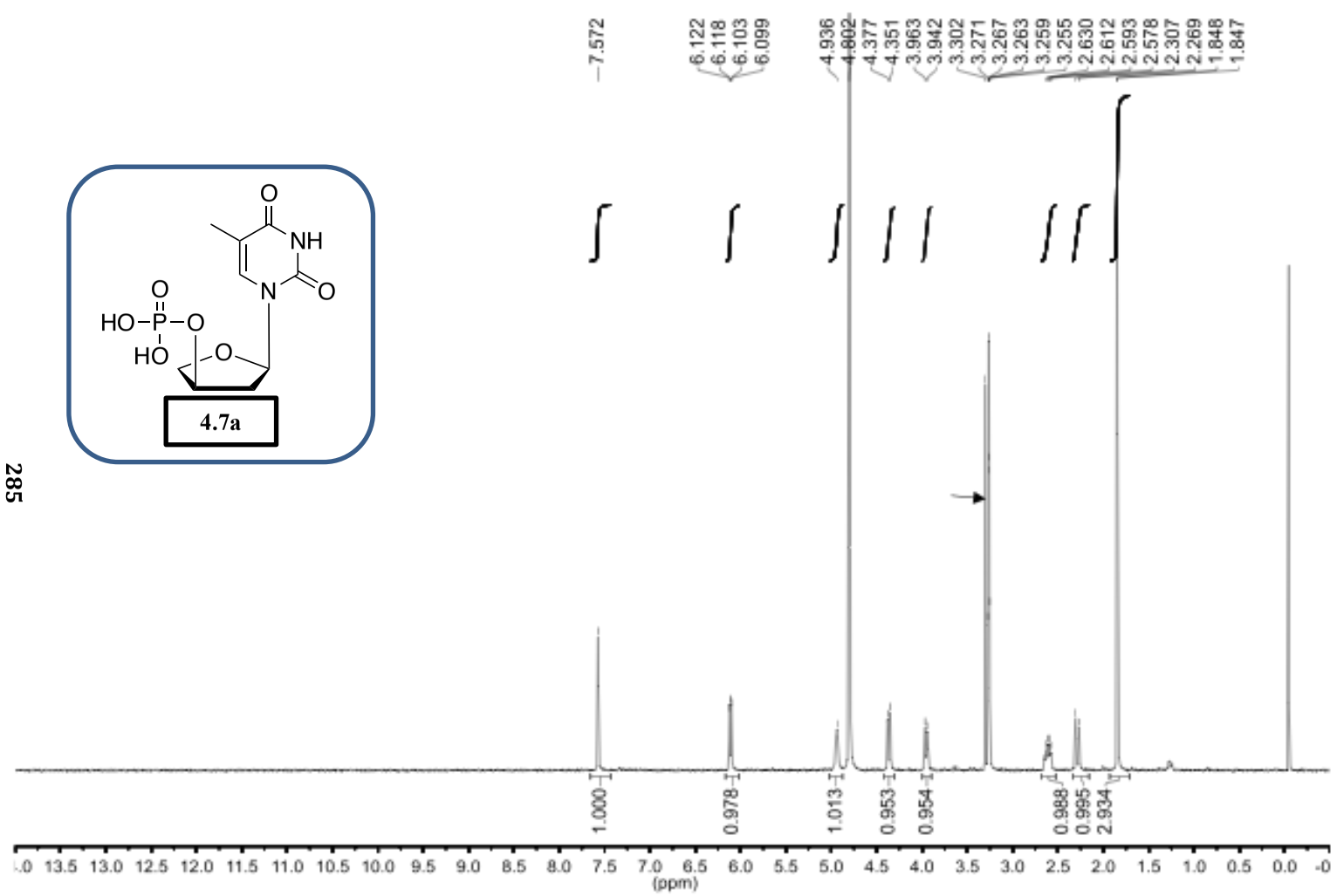
[1] HNMR spectrum of compound **4.6d'** (500 MHz,CD_3OD)



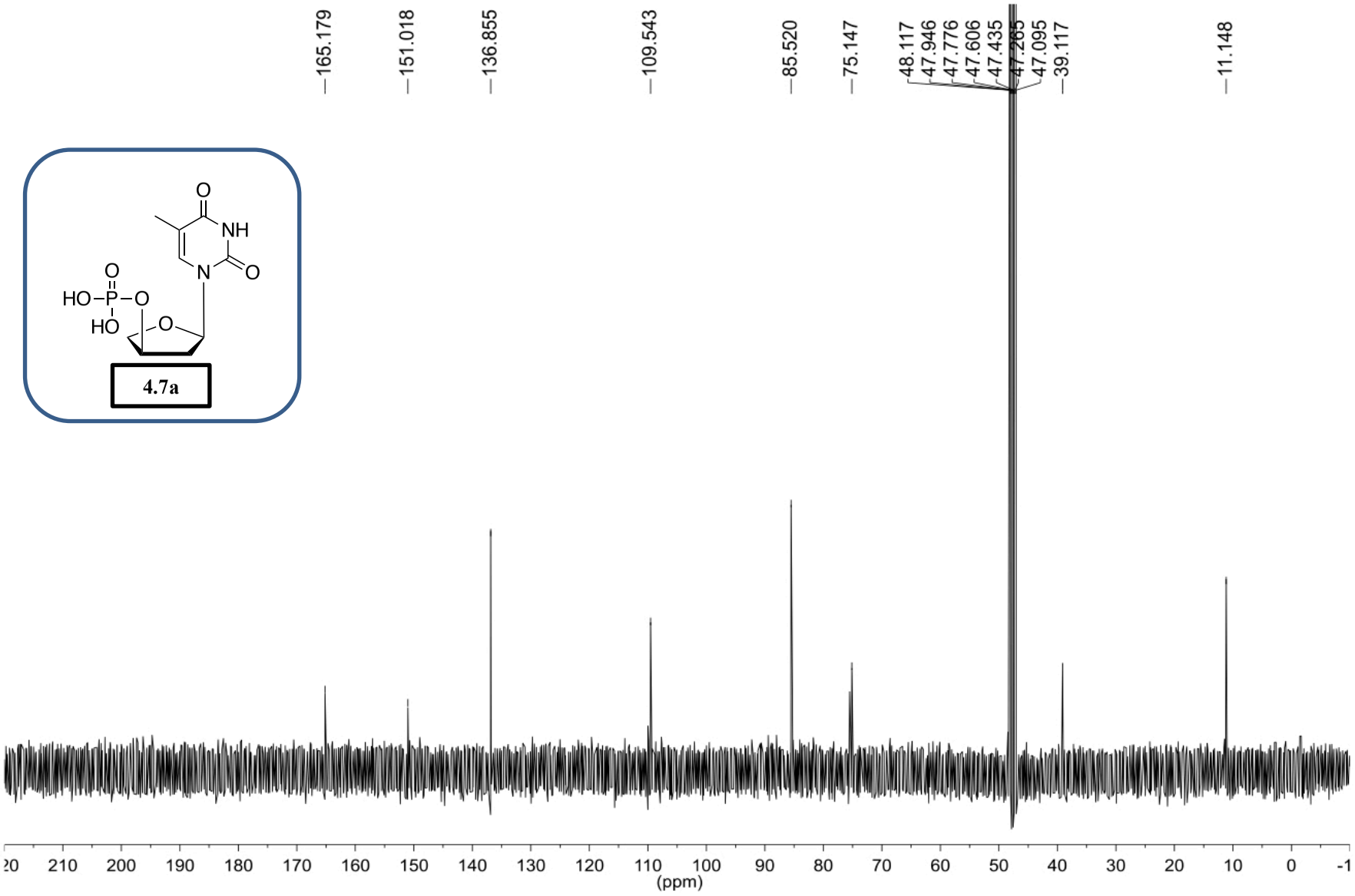
^{13}CNMR spectrum of compound **4.6d'** (500 MHz,CD_3OD)



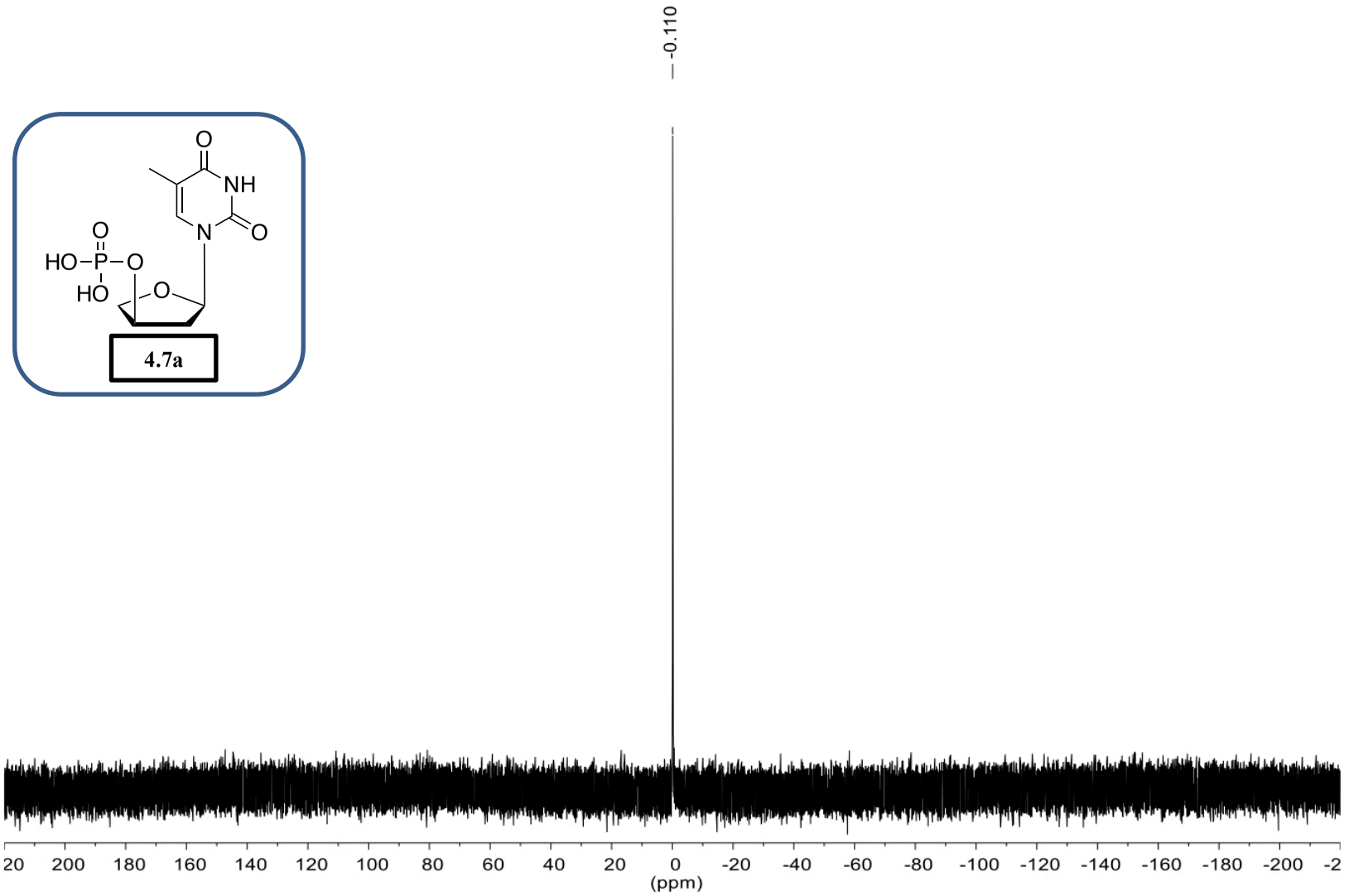
^{31} PNMR spectrum of compound **4.6d'** (500 MHz,CD_3OD)



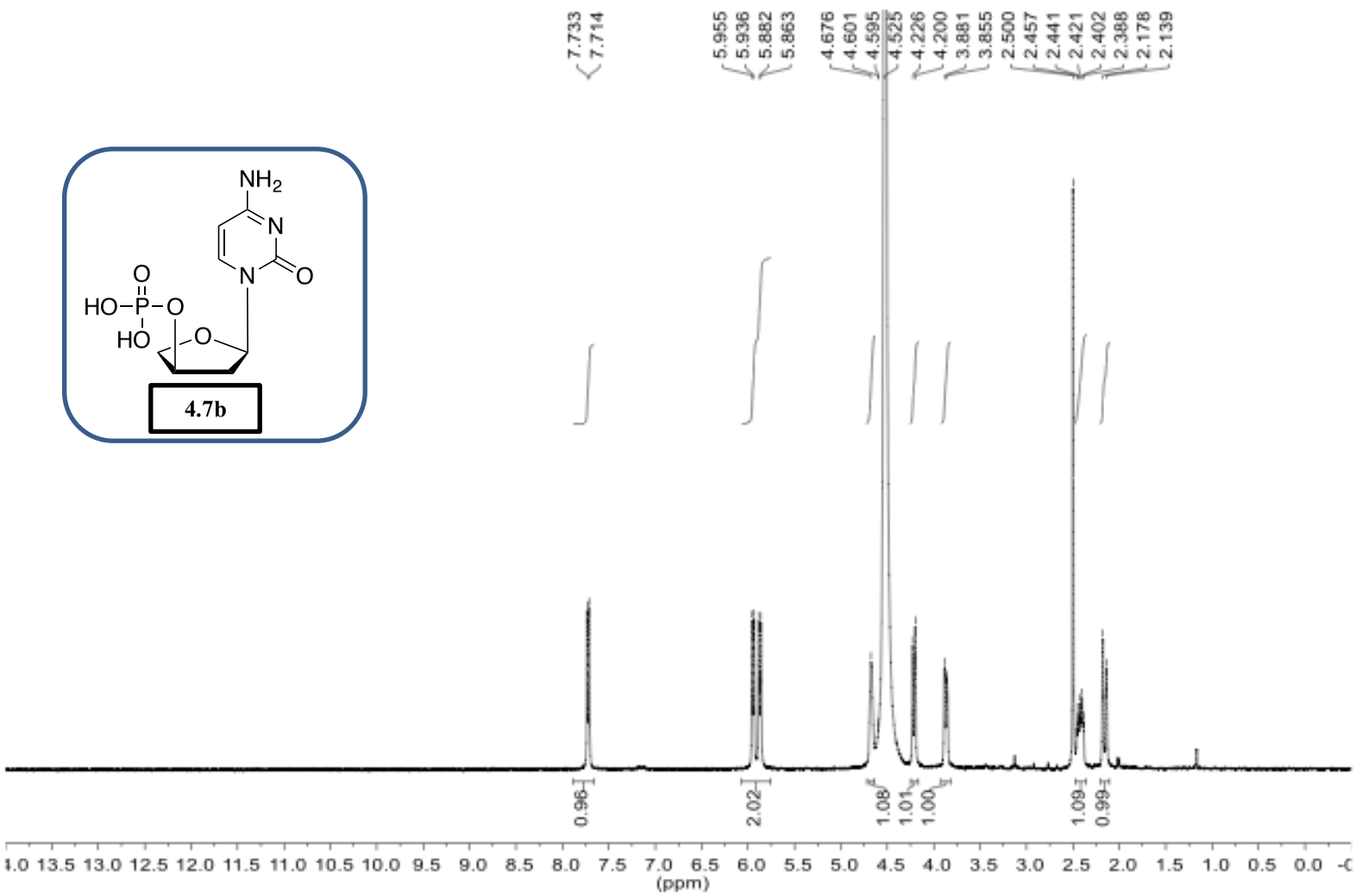
[1] HNMR spectrum of compound **4.7a** (500 MHz,CD_3OD)



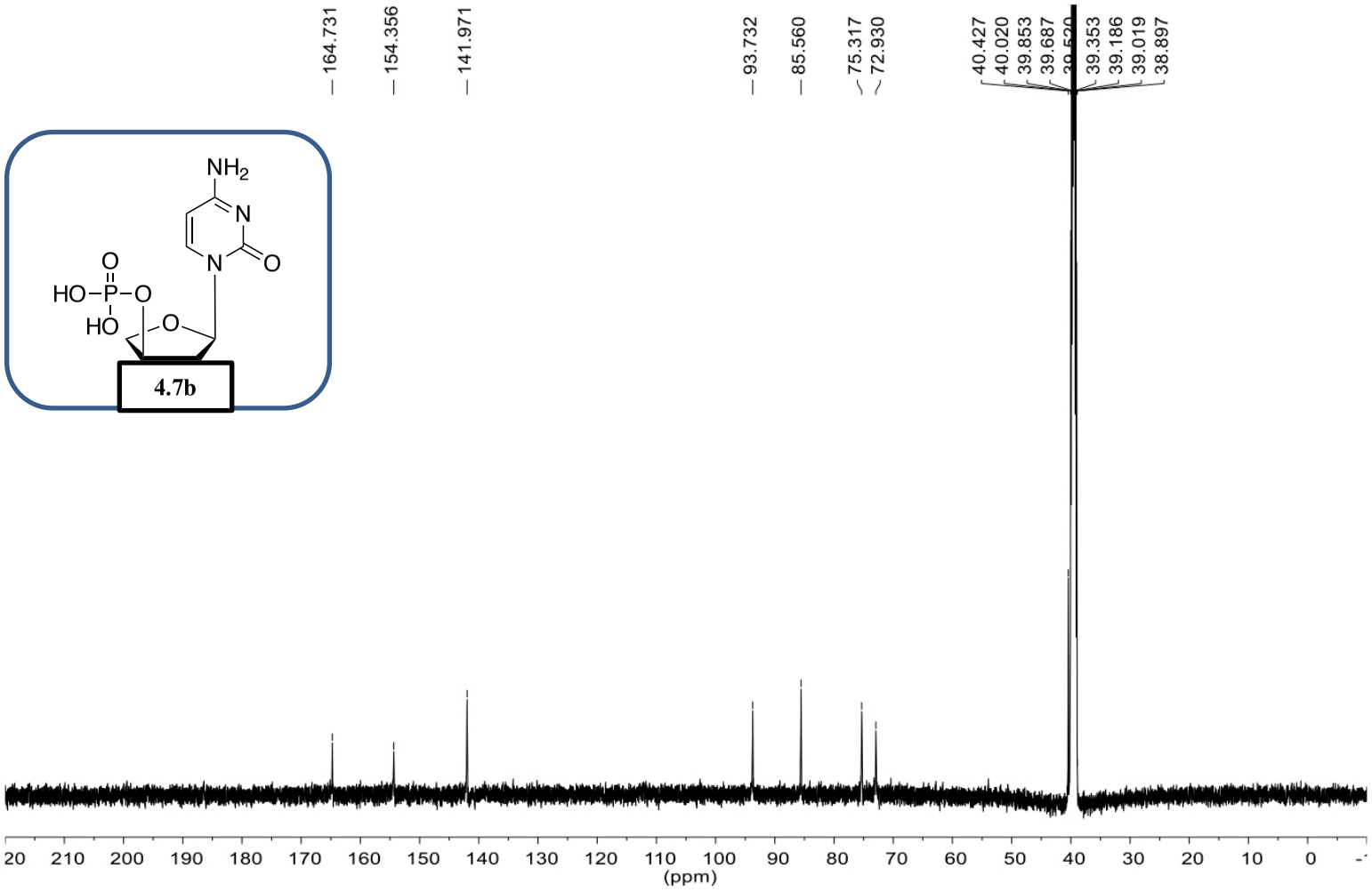
^{13}CNMR spectrum of compound **4.7a** (500 MHz,CD_3OD)



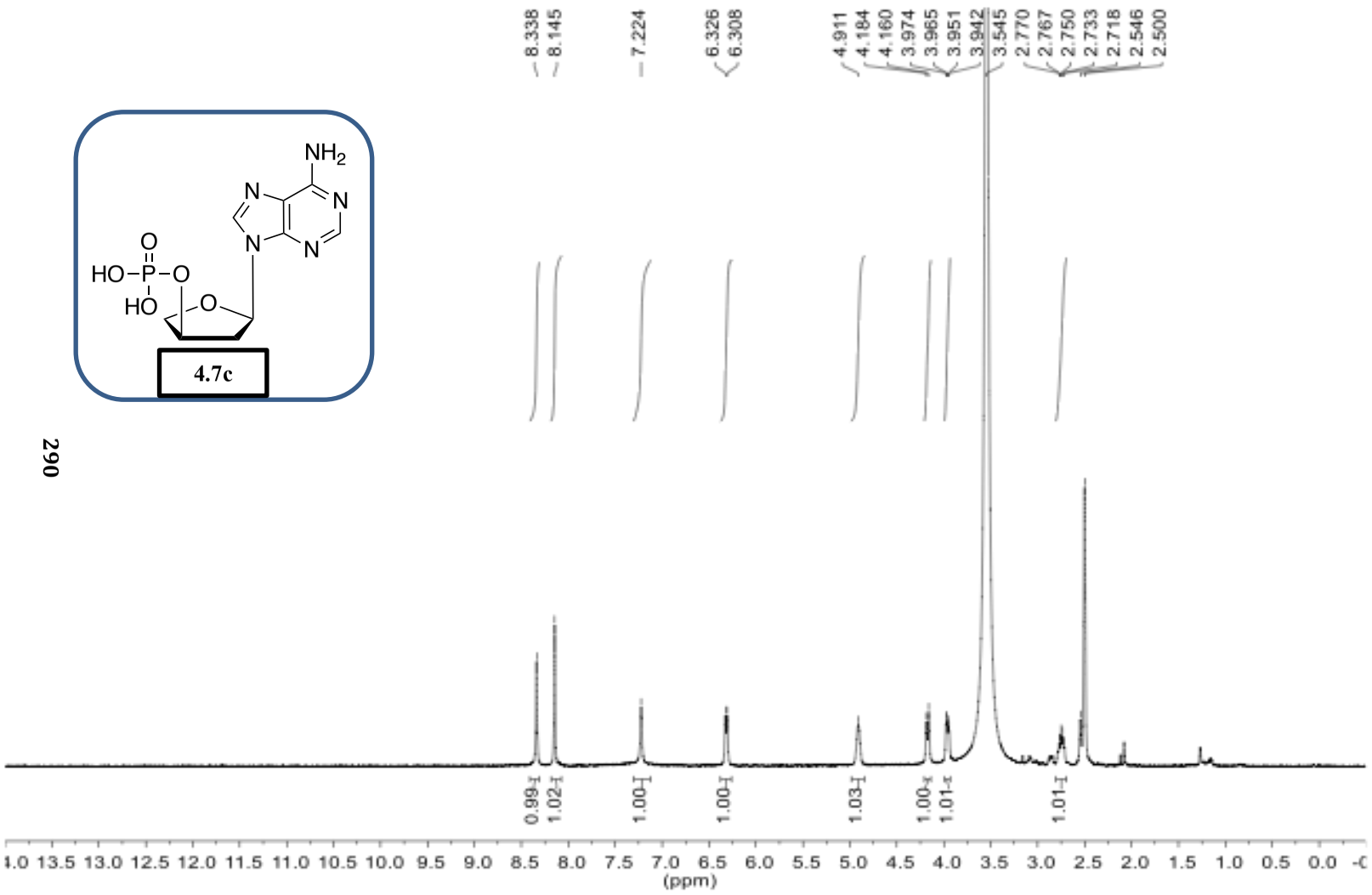
[31] PNMR spectrum of compound **4.7a** (500 MHz,CD_3OD)



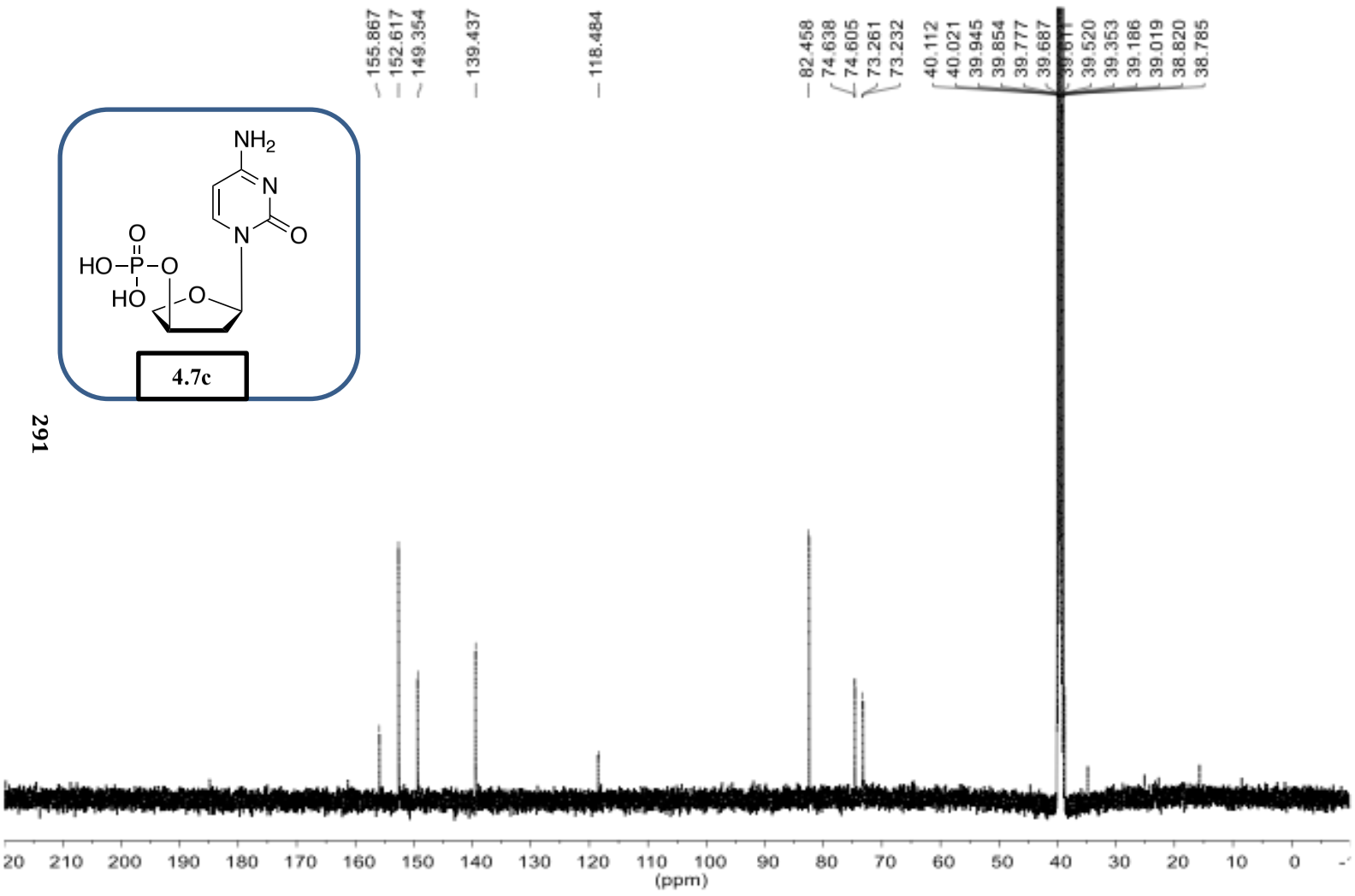
[1] HNMR spectrum of compound **4.7b** (500 MHz,CD_3OD)



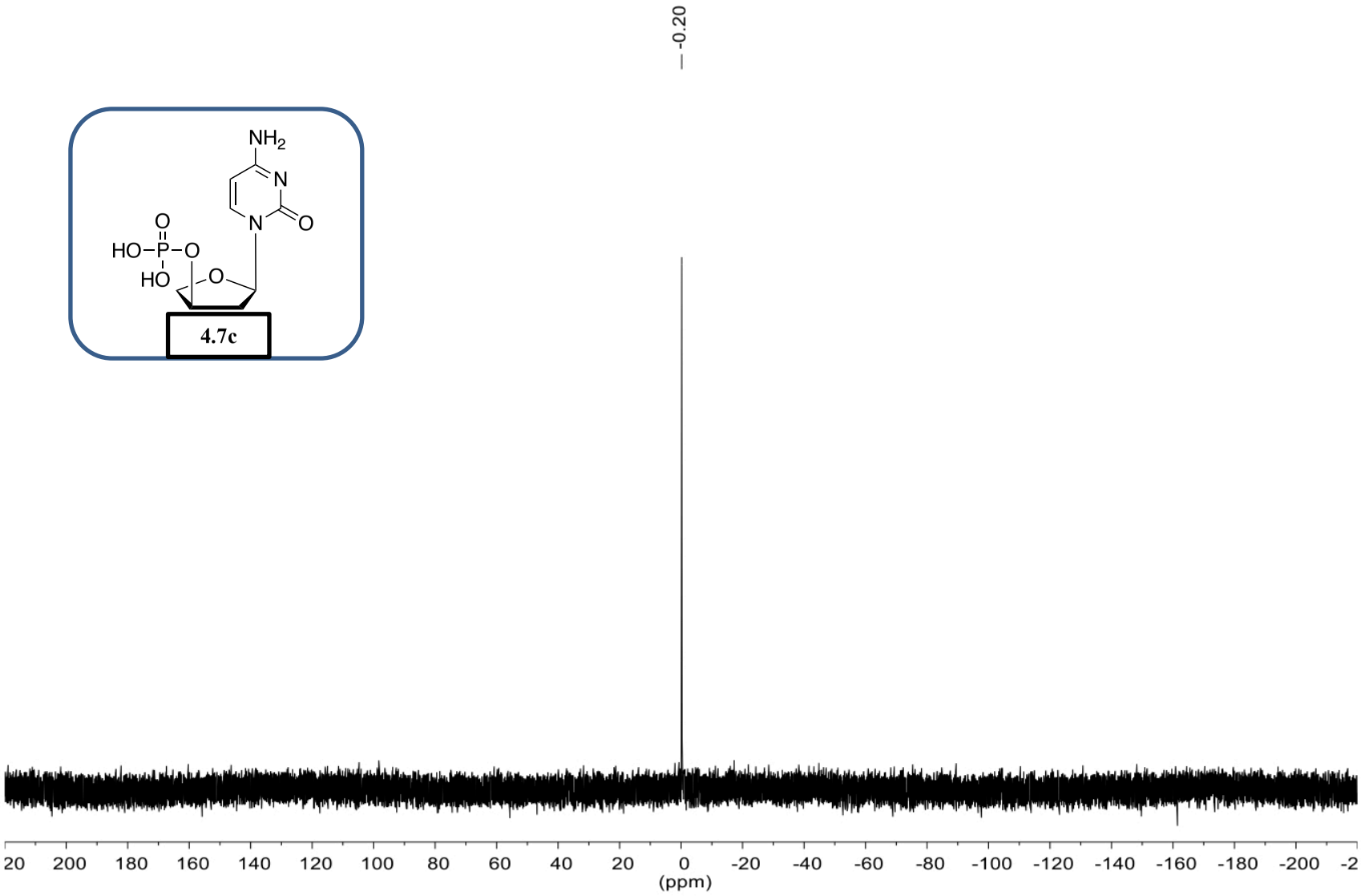
^{13}C NMR spectrum of compound **4.7b** (500 MHz,CD_3OD)



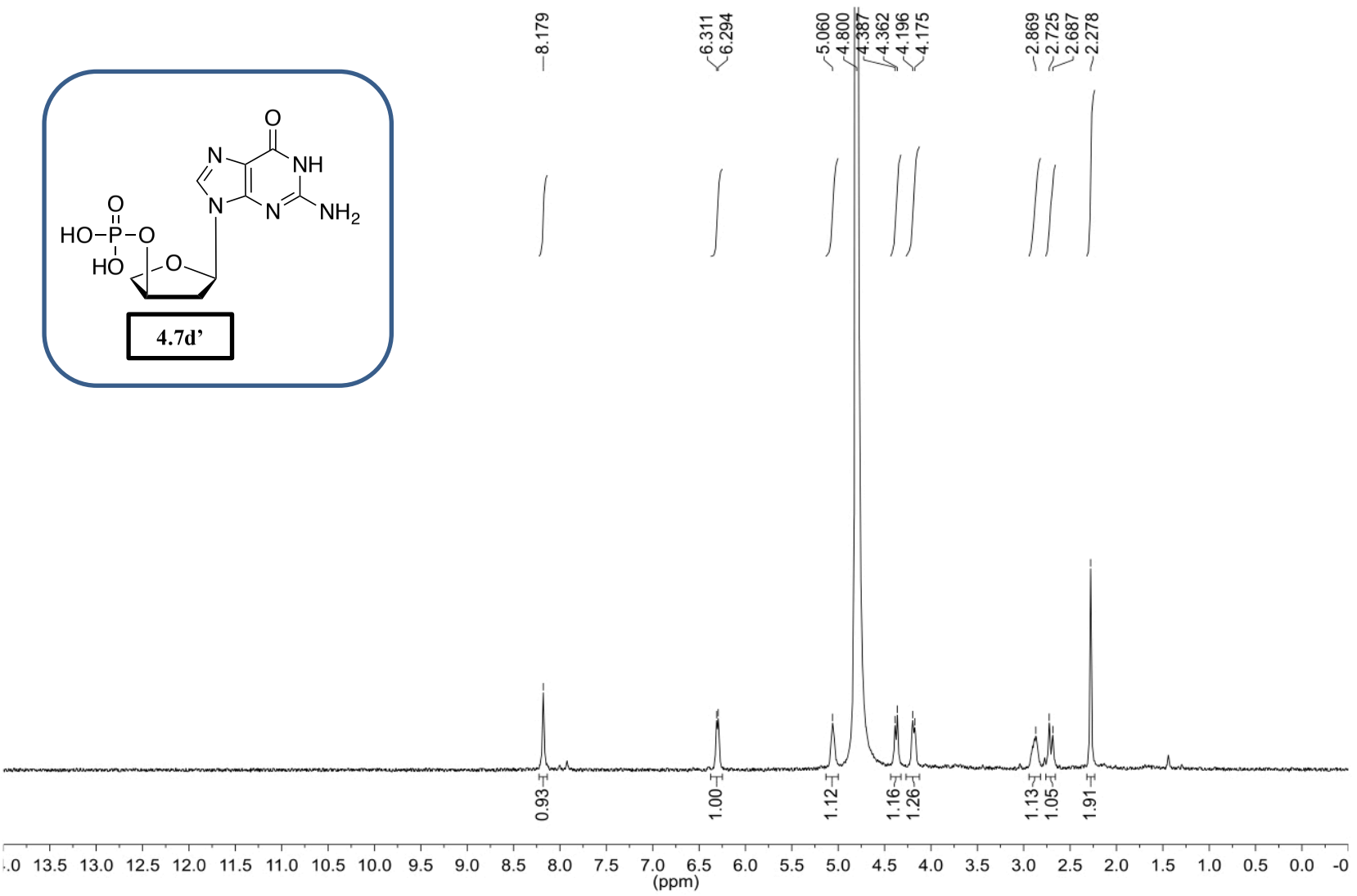
^{1}H NMR spectrum of compound **4.7c** (500 MHz,CD_3OD)



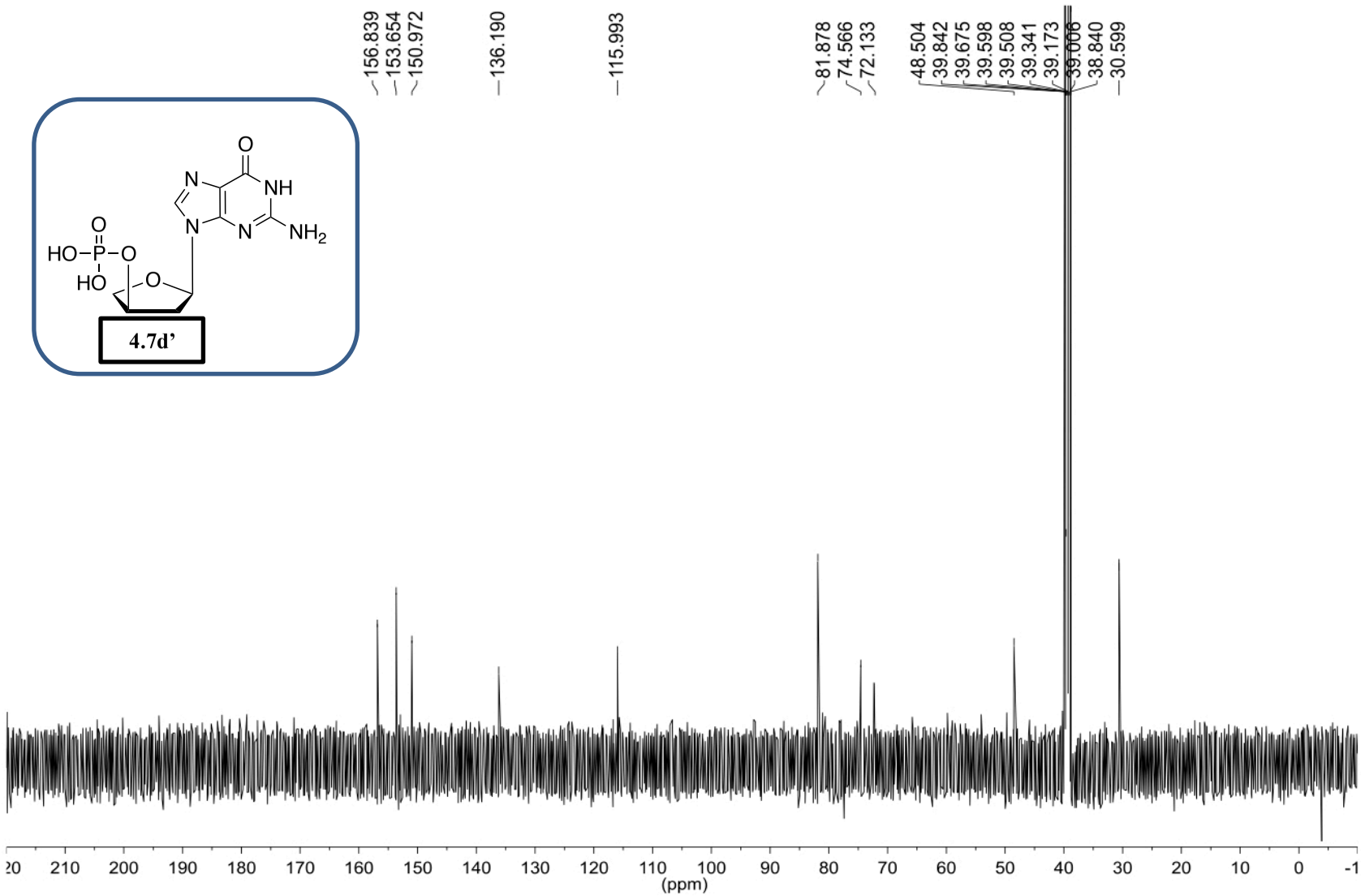
^{13}C NMR spectrum of compound **4.7c** (500 MHz,CD_3OD)



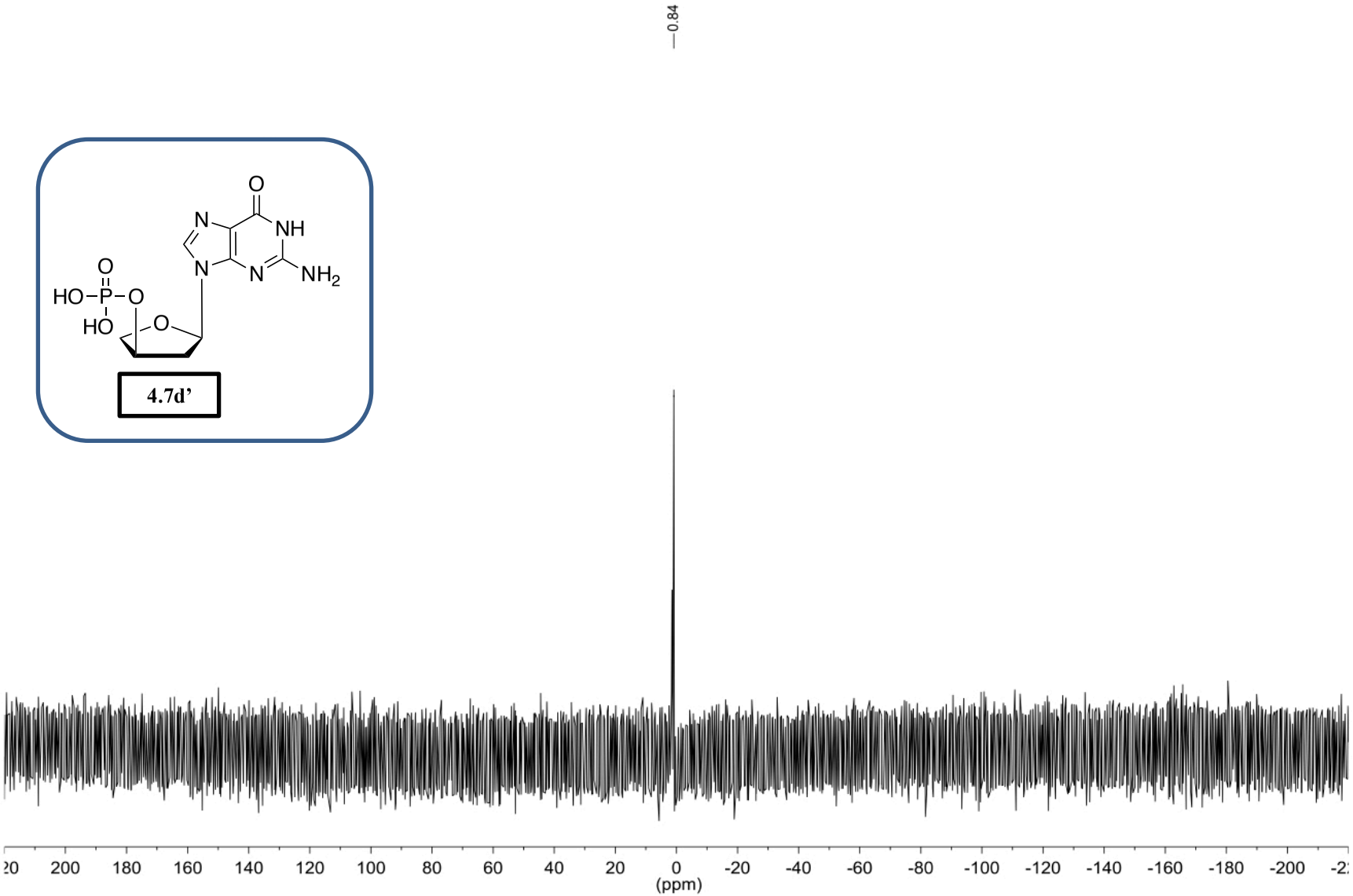
^{31}P NMR spectrum of compound **4.7c** (500 MHz,CD_3OD)



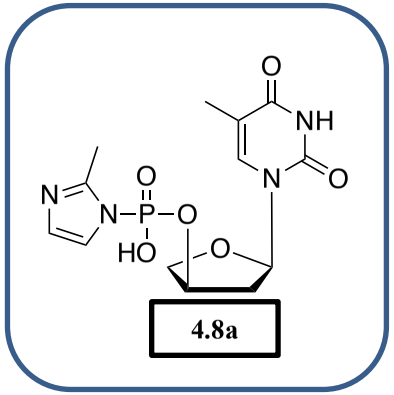
^{1}H NMR spectrum of compound **4.7d'** (500 MHz,CD_3OD)



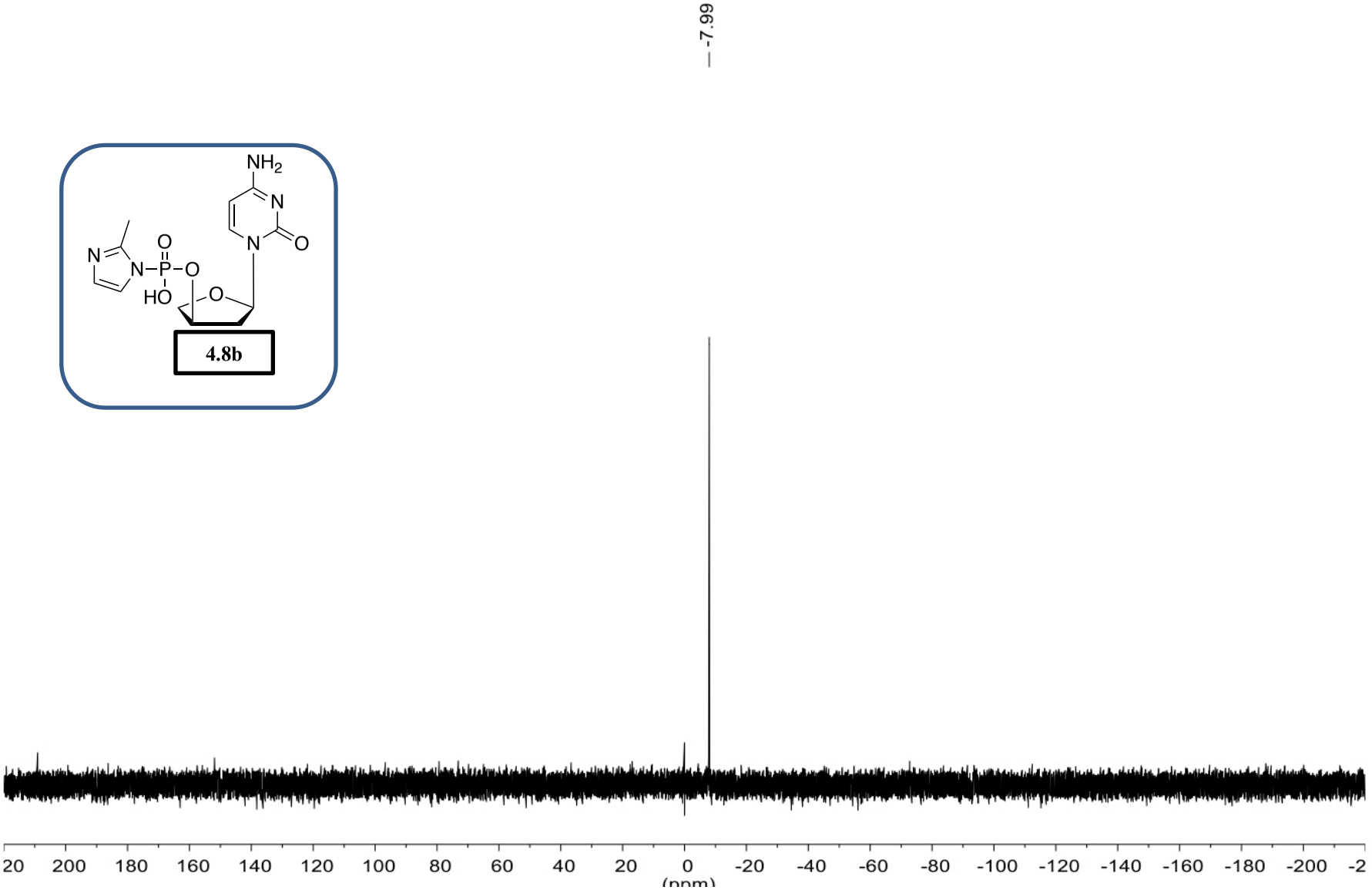
^{13}C NMR spectrum of compound **4.7d'** (500 MHz,CD_3OD)



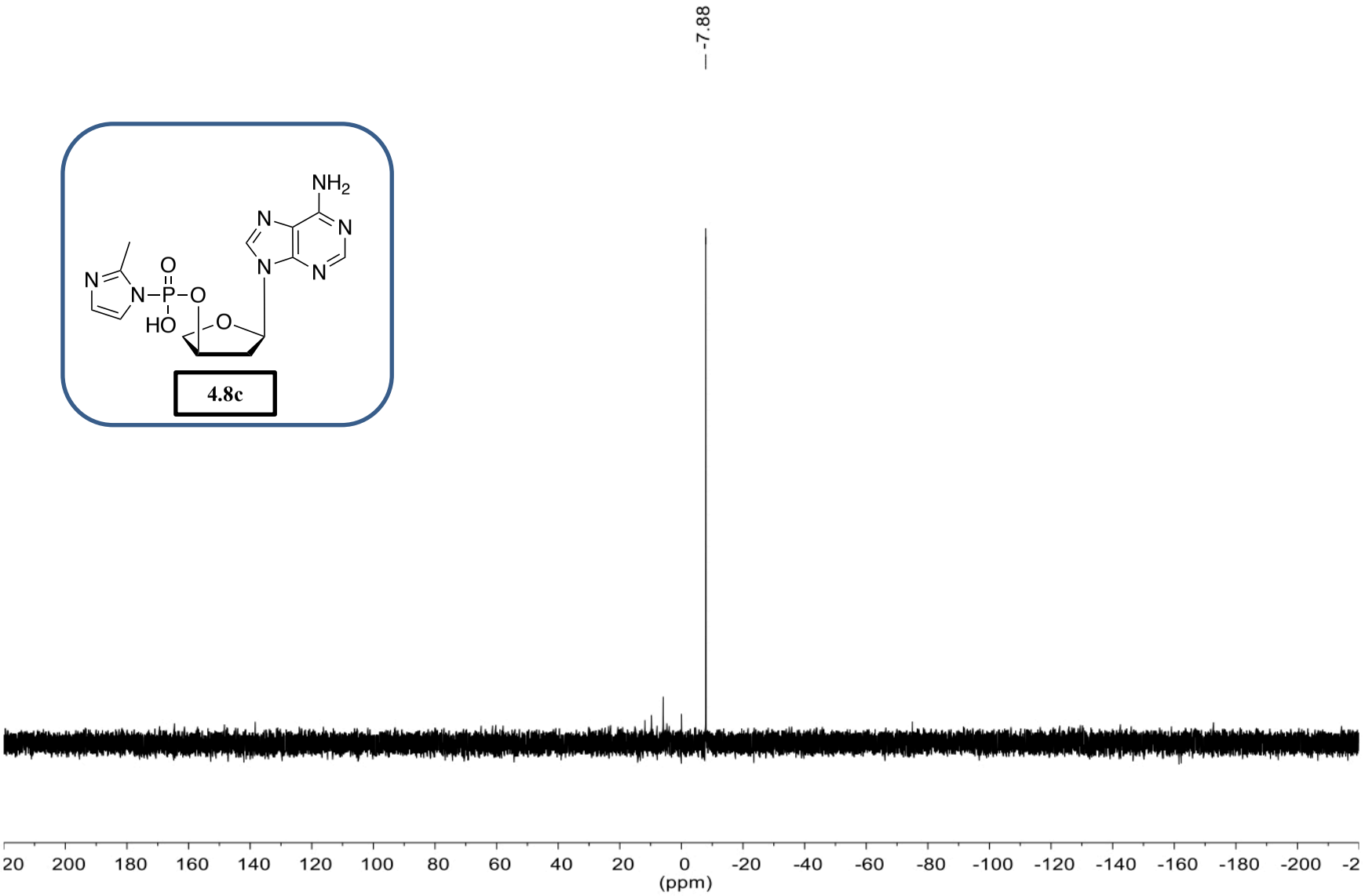
^{31}P NMR spectrum of compound **4.7d'** (500 MHz,CD_3OD)



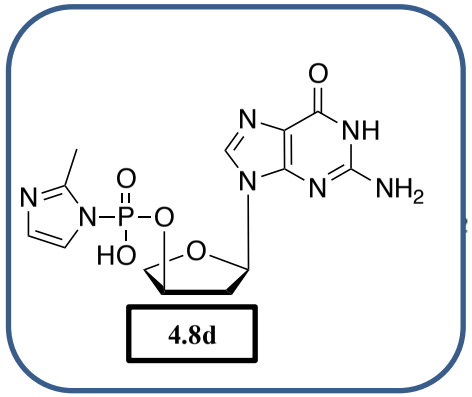
^{31}P NMR spectrum of compound **4.8a** (500 MHz,D_2O)



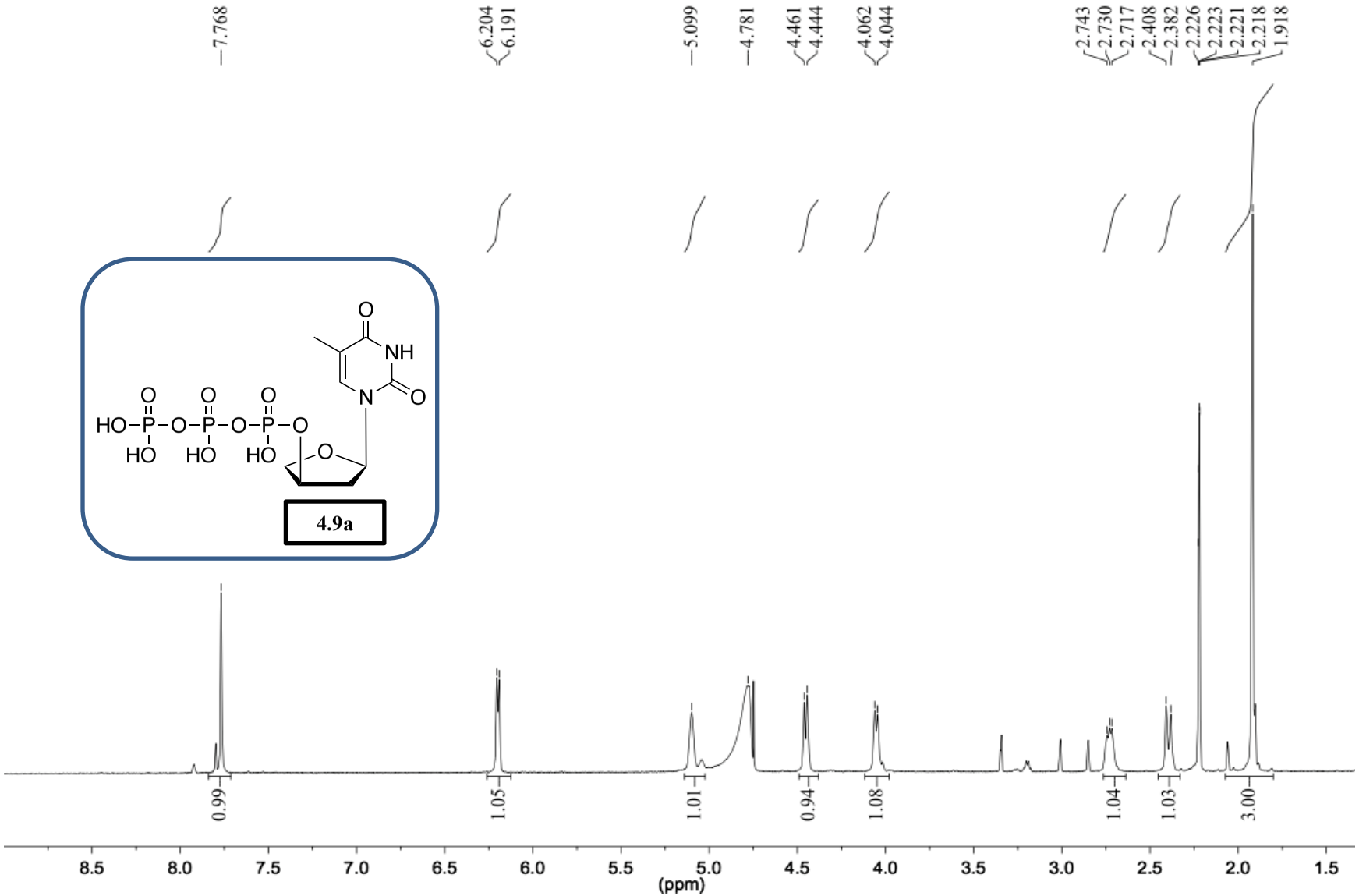
^{31}P NMR spectrum of compound **4.8b** (500 MHz,D_2O)



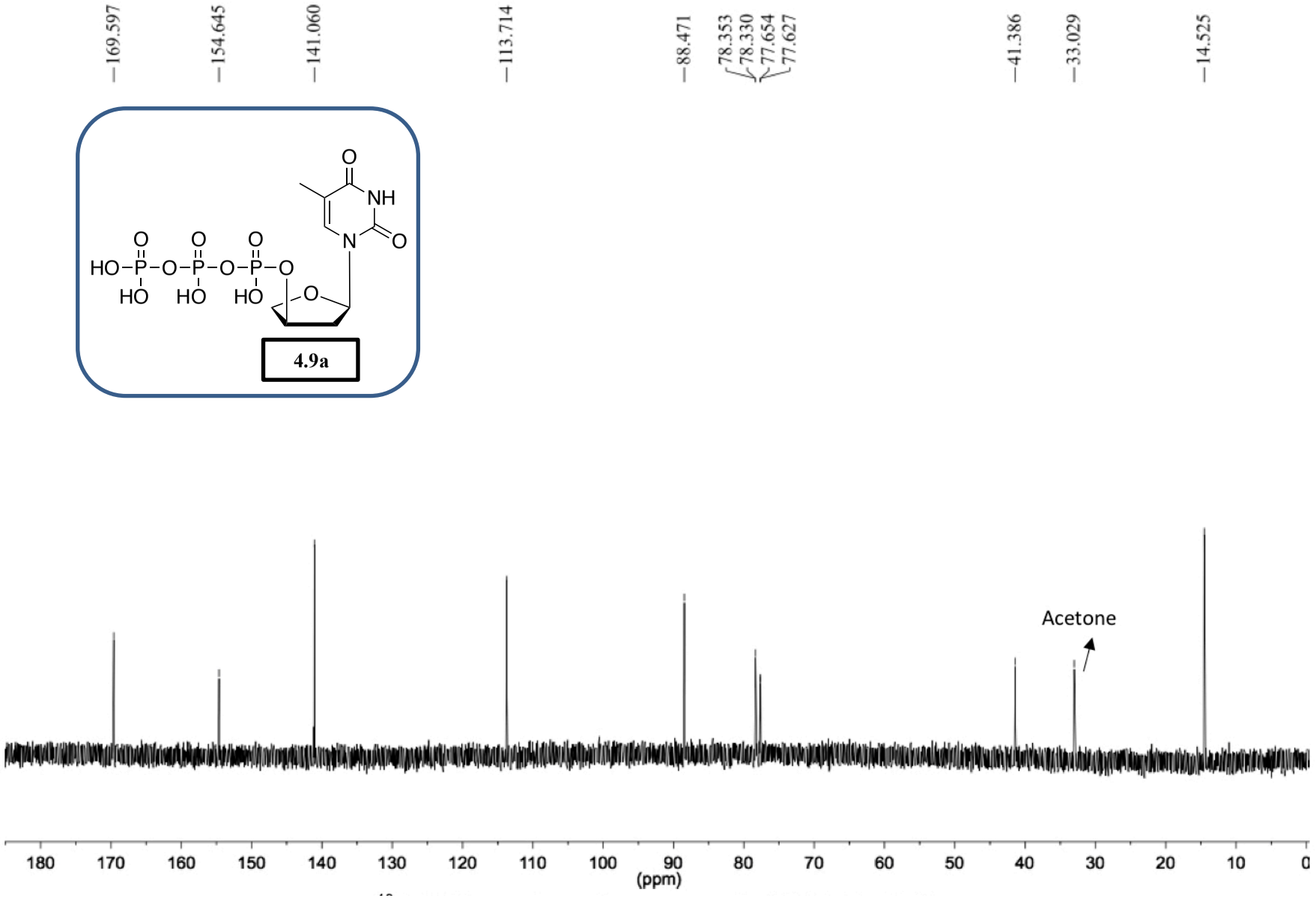
^{31}P NMR spectrum of compound **4.8c** (500 MHz,D_2O)



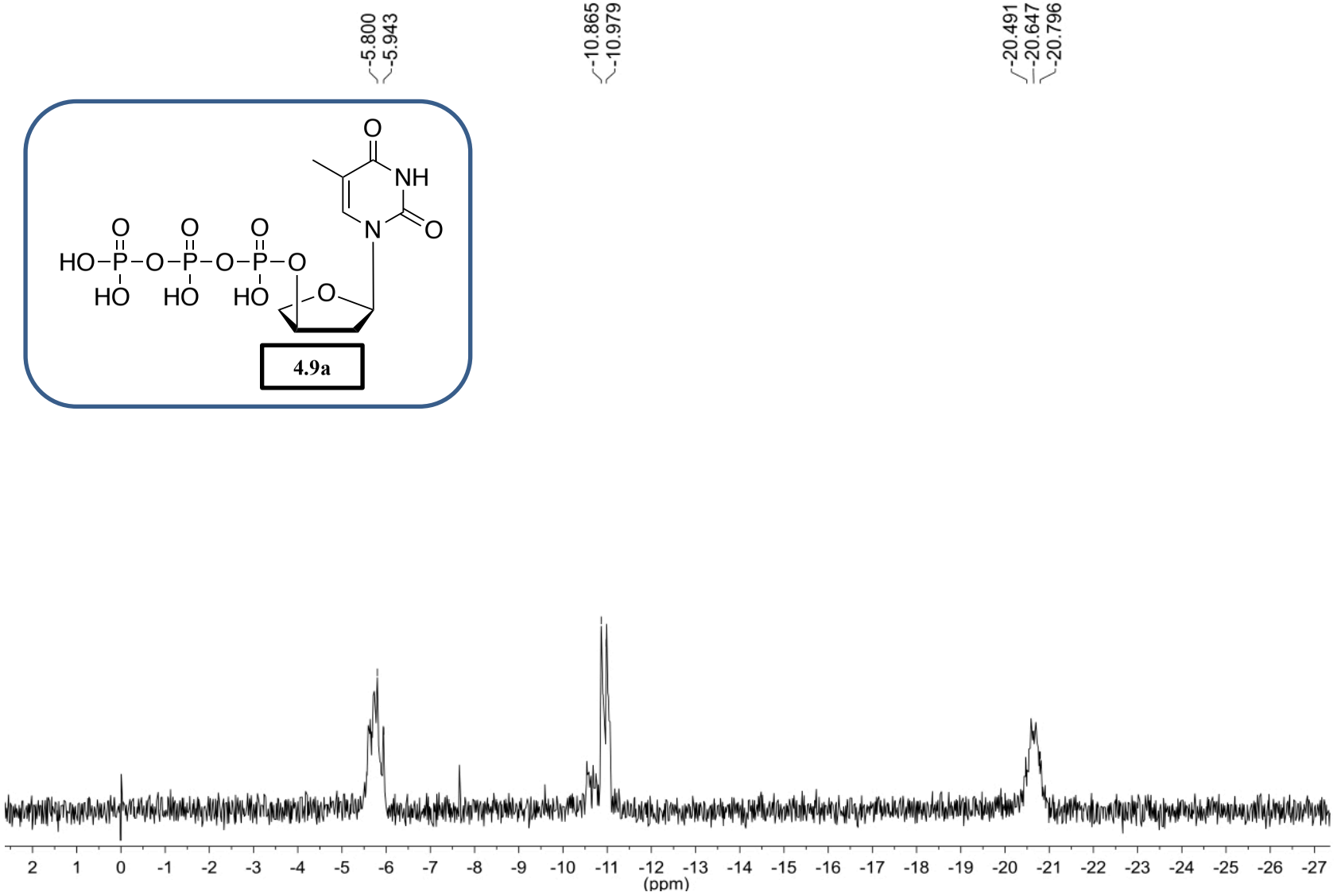
^{31}P NMR spectrum of compound **4.8d** (500 MHz,D_2O)



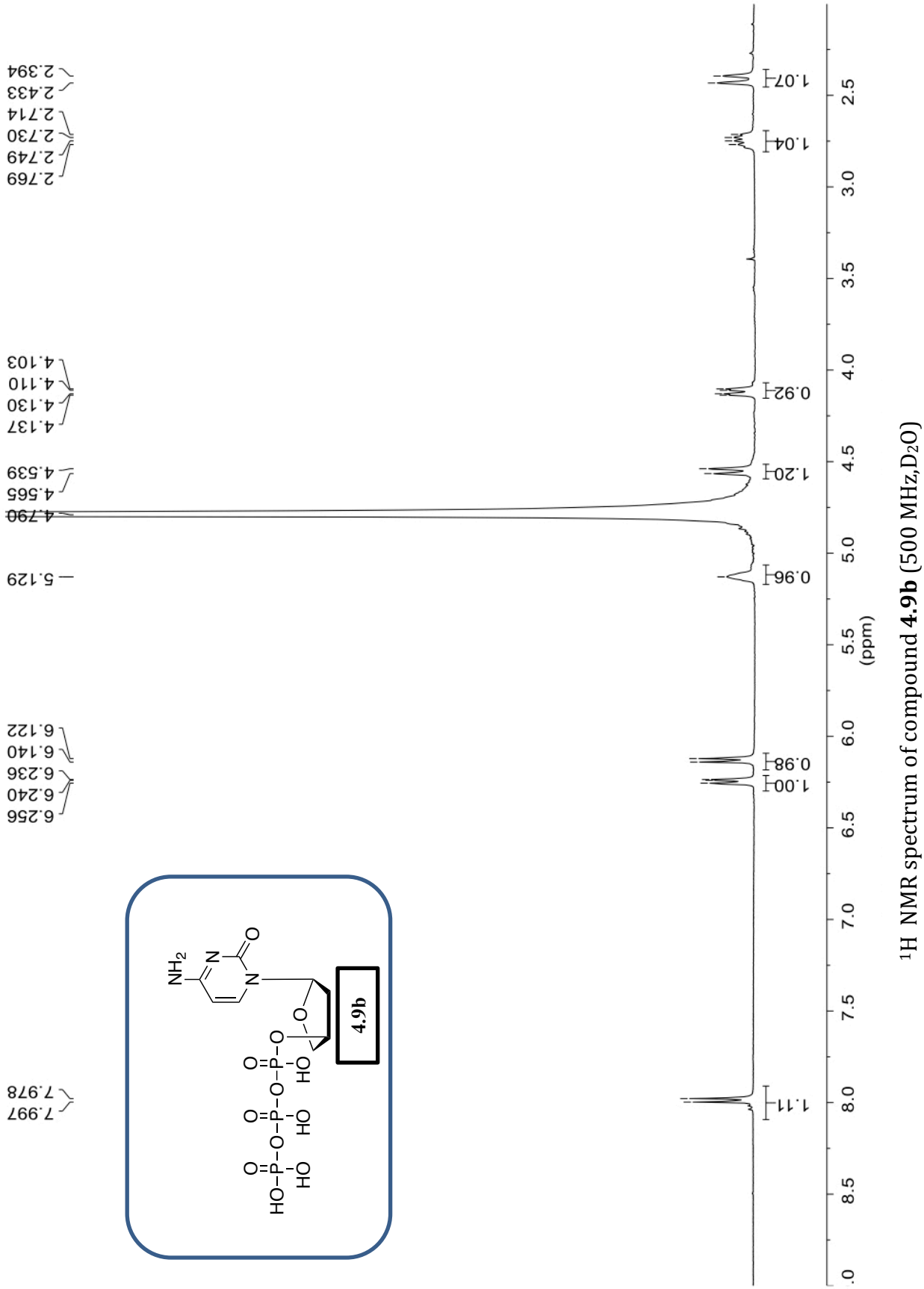
^{1}H NMR spectrum of compound **4.9a** (500 MHz,D_2O)



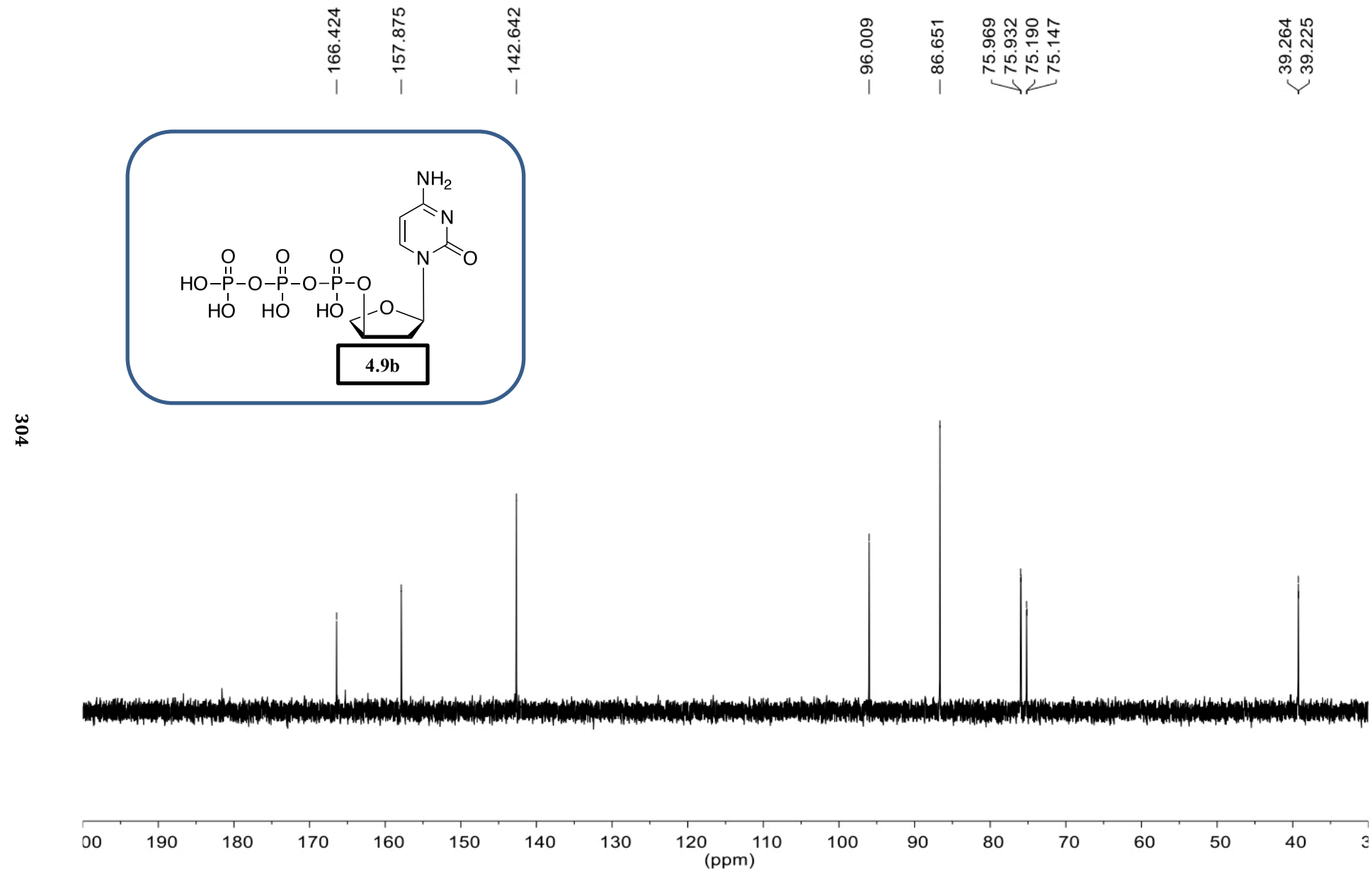
^{13}C NMR spectrum of compound **4.9a** (500 MHz,D_2O)



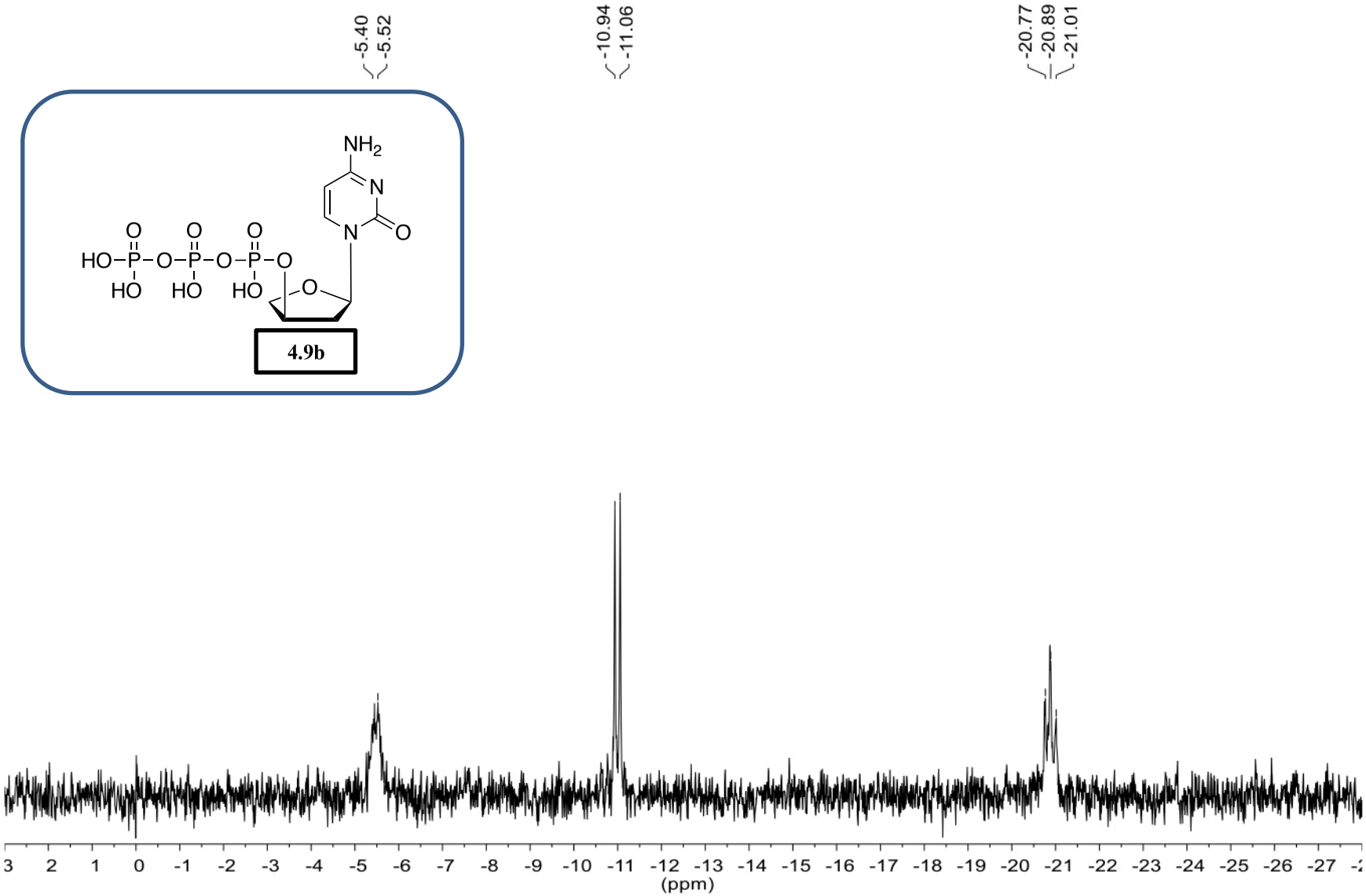
^{31}P NMR spectrum of compound **4.9a** (500 MHz,D_2O)



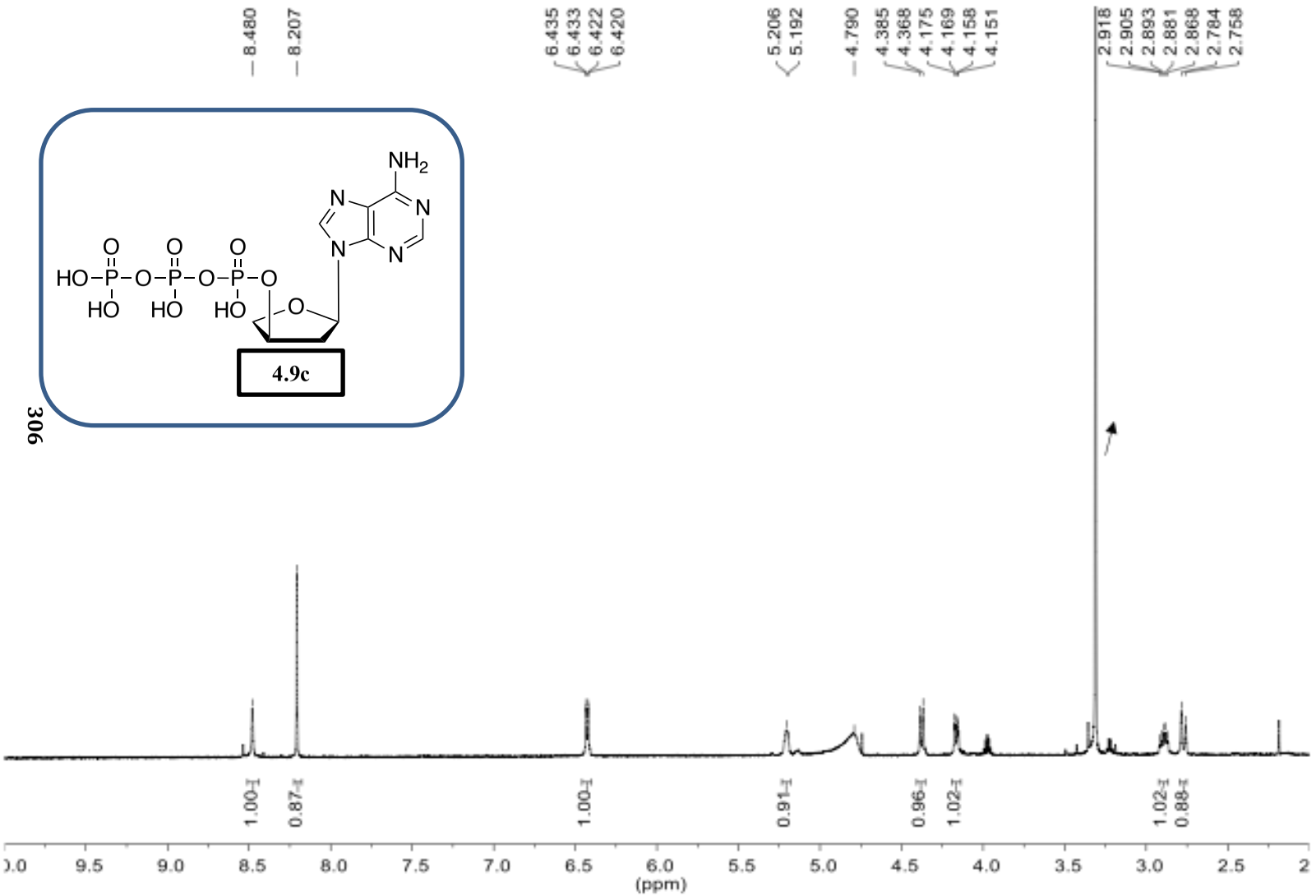
^{1}H NMR spectrum of compound **4.9b** (500 MHz,D_2O)



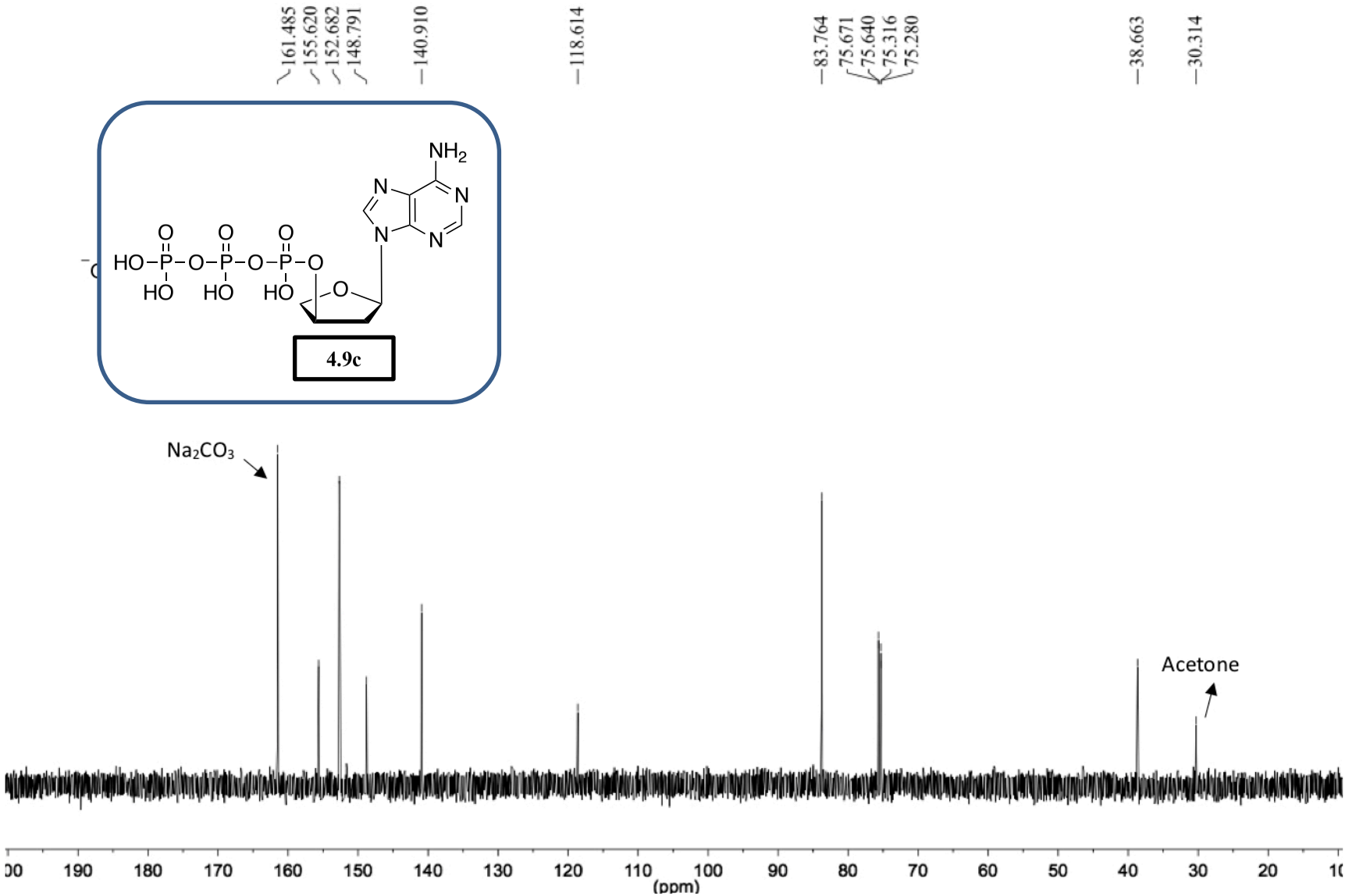
^{13}C NMR spectrum of compound **4.9b** (500 MHz,D_2O)



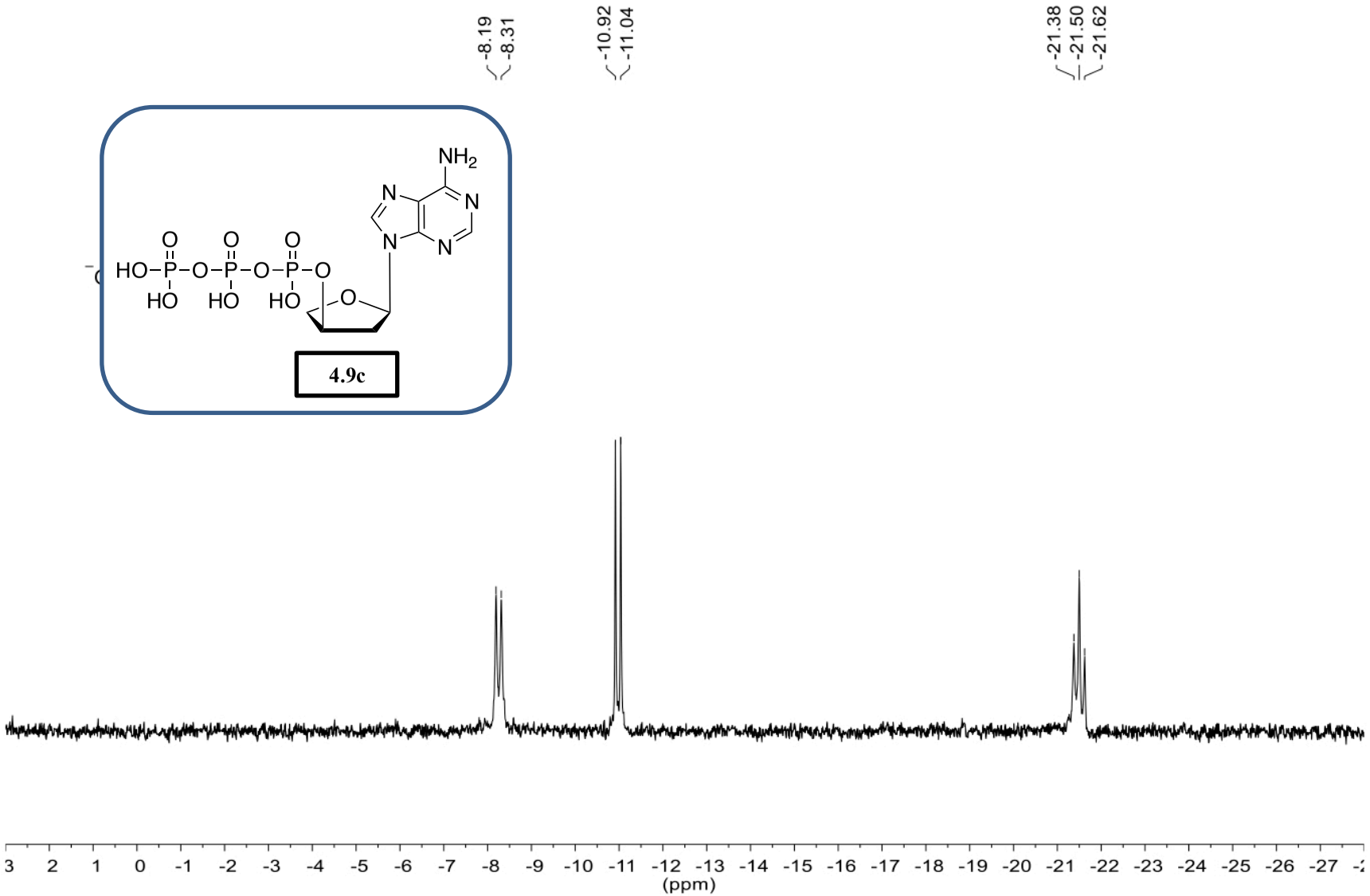
^{31}P NMR spectrum of compound **4.9b** (500 MHz,D_2O)



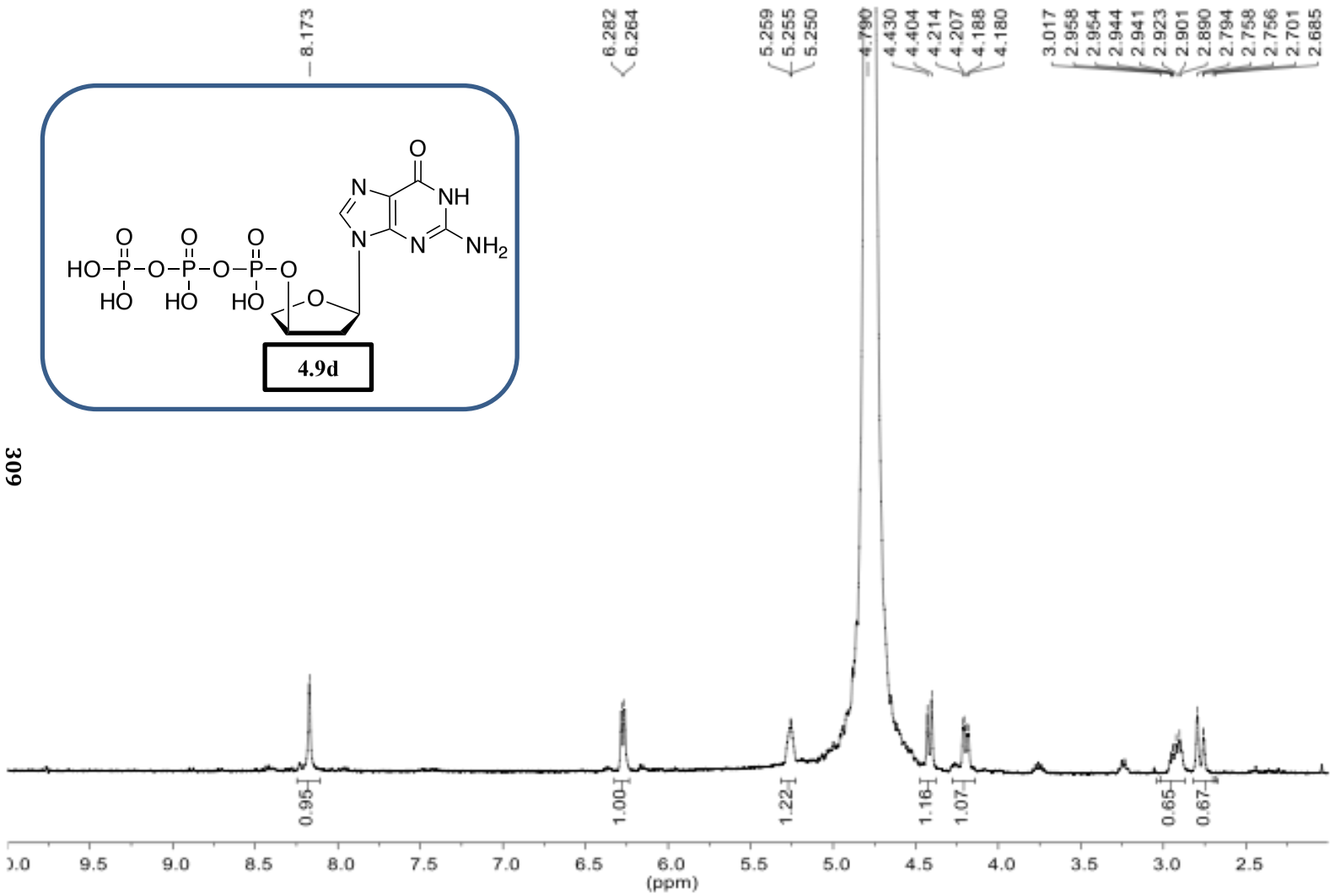
^{1}H NMR spectrum of compound **4.9c** (500 MHz,D_2O)



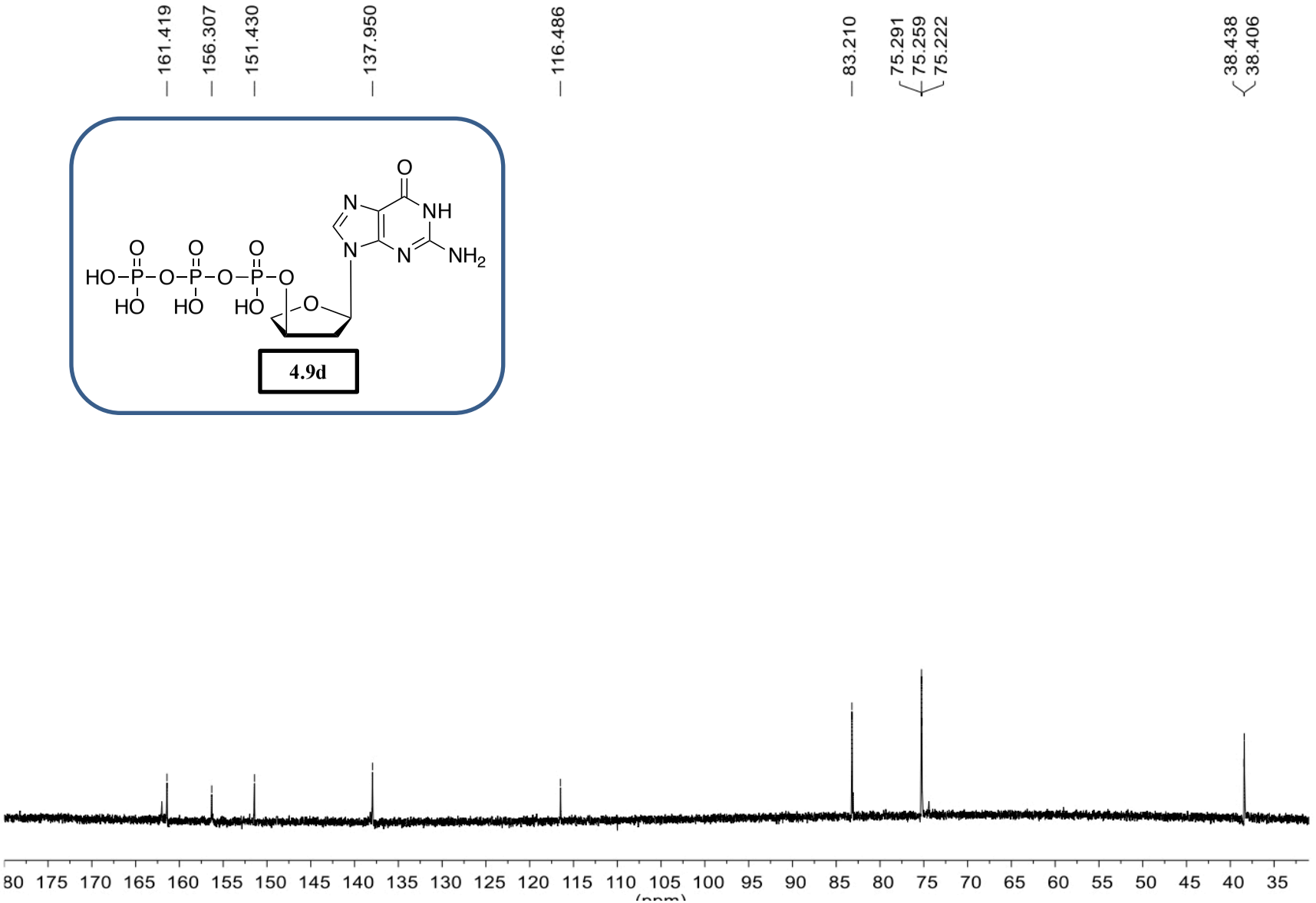
^{13}C NMR spectrum of compound **4.9c** (500 MHz,D_2O)



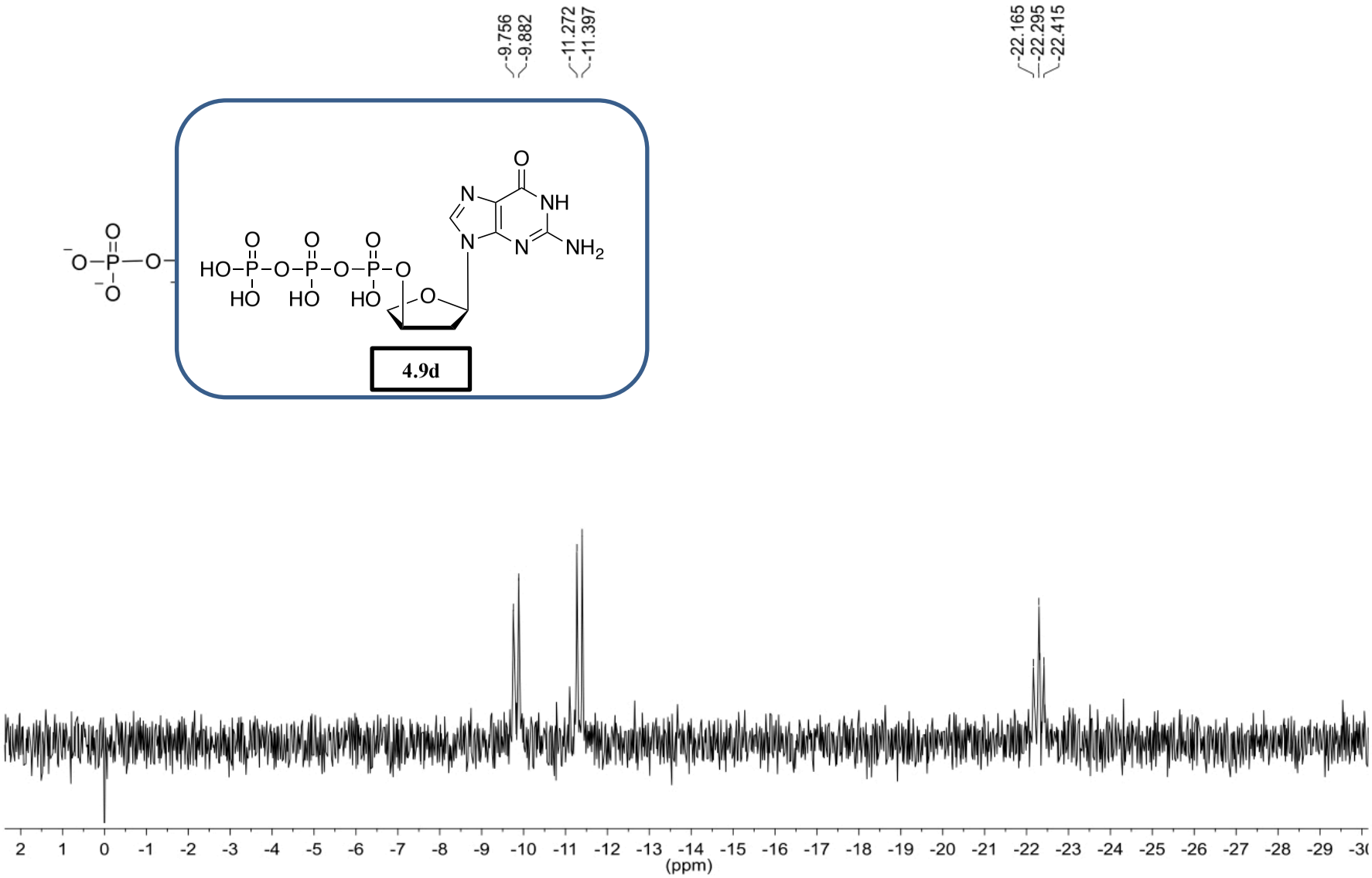
^{31}P NMR spectrum of compound **4.9c** (500 MHz,D_2O)



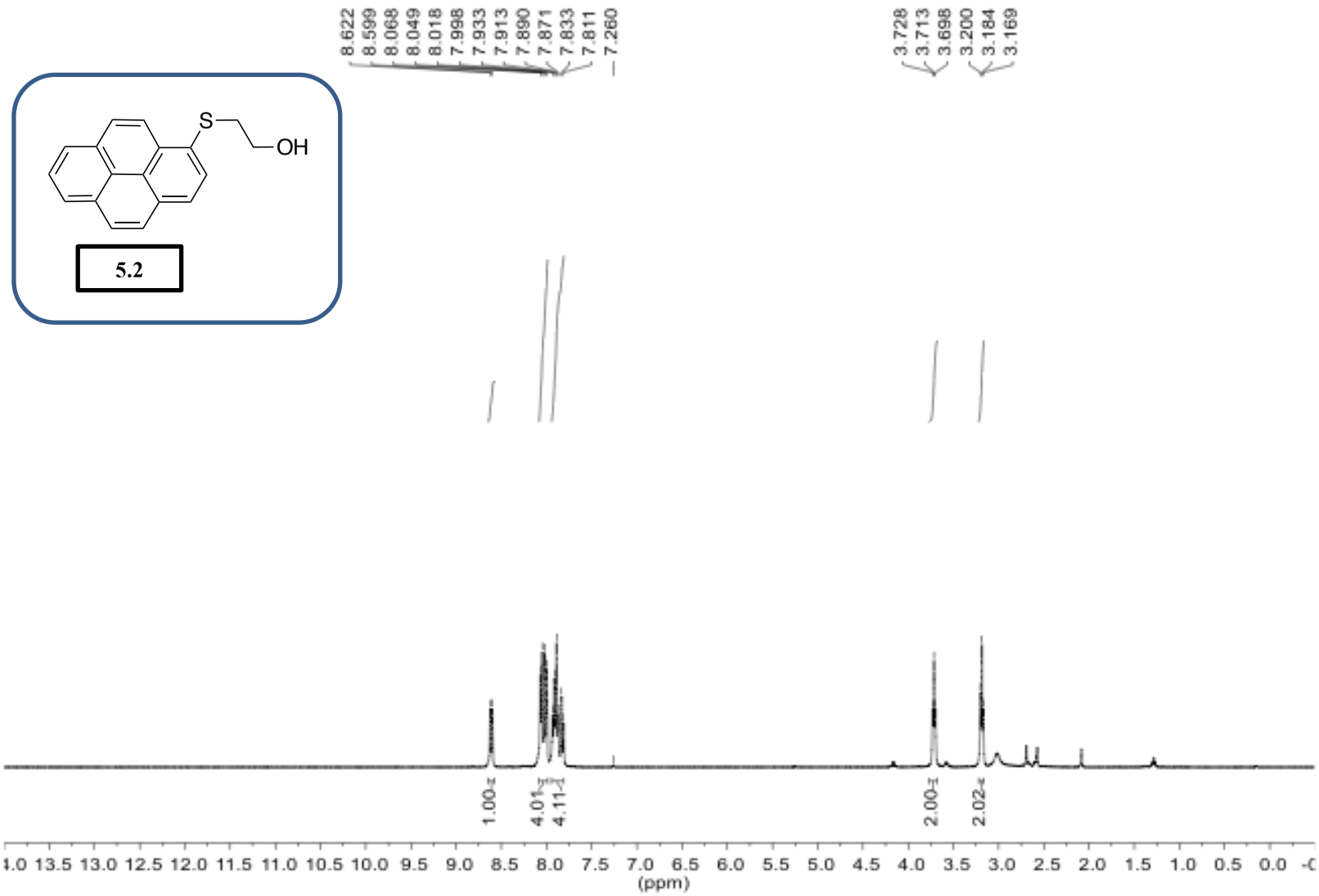
^{1}H NMR spectrum of compound **4.9d** (500 MHz,D_2O)



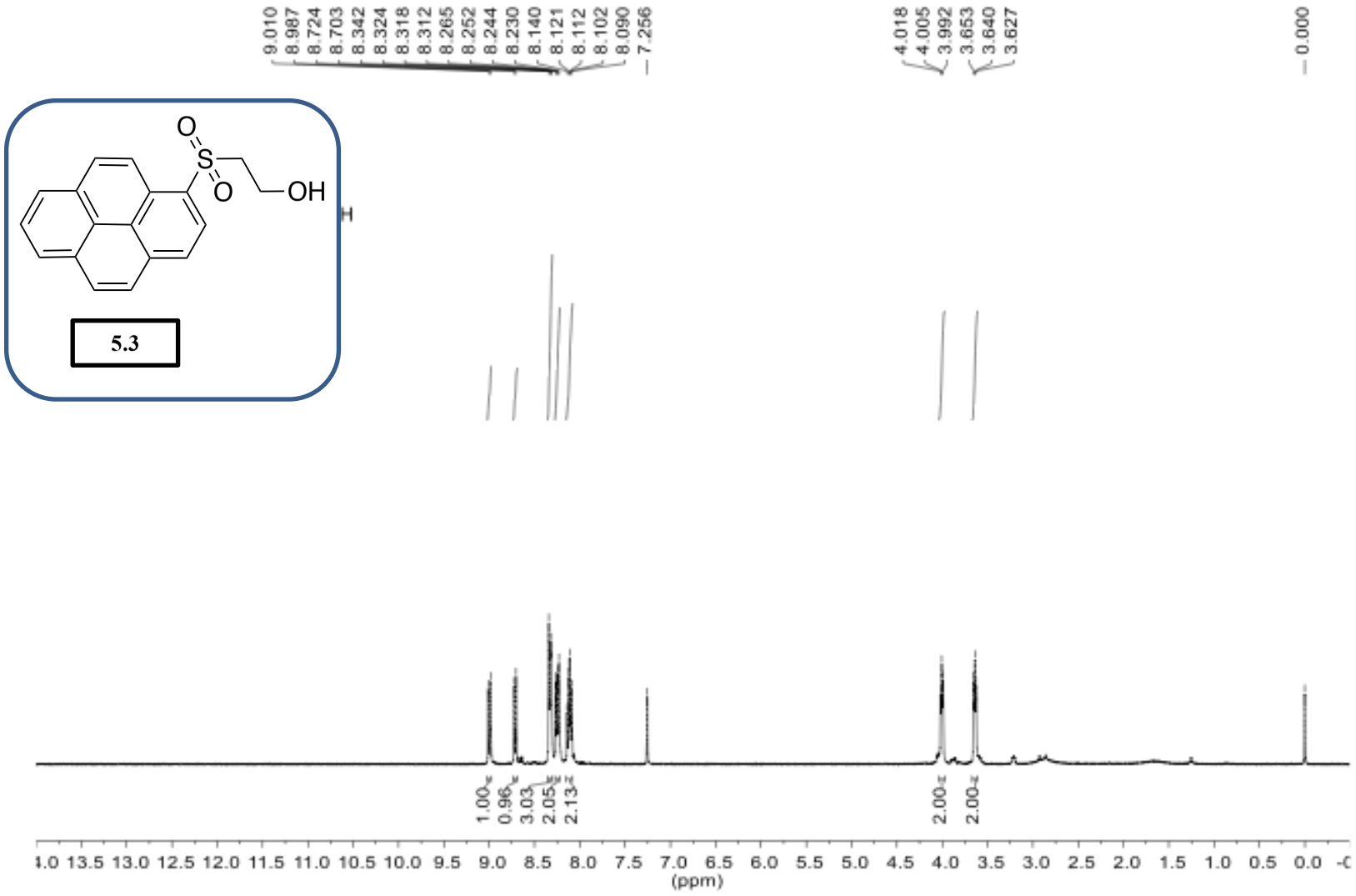
^{13}C NMR spectrum of compound **4.9d** (500 MHz,D_2O)



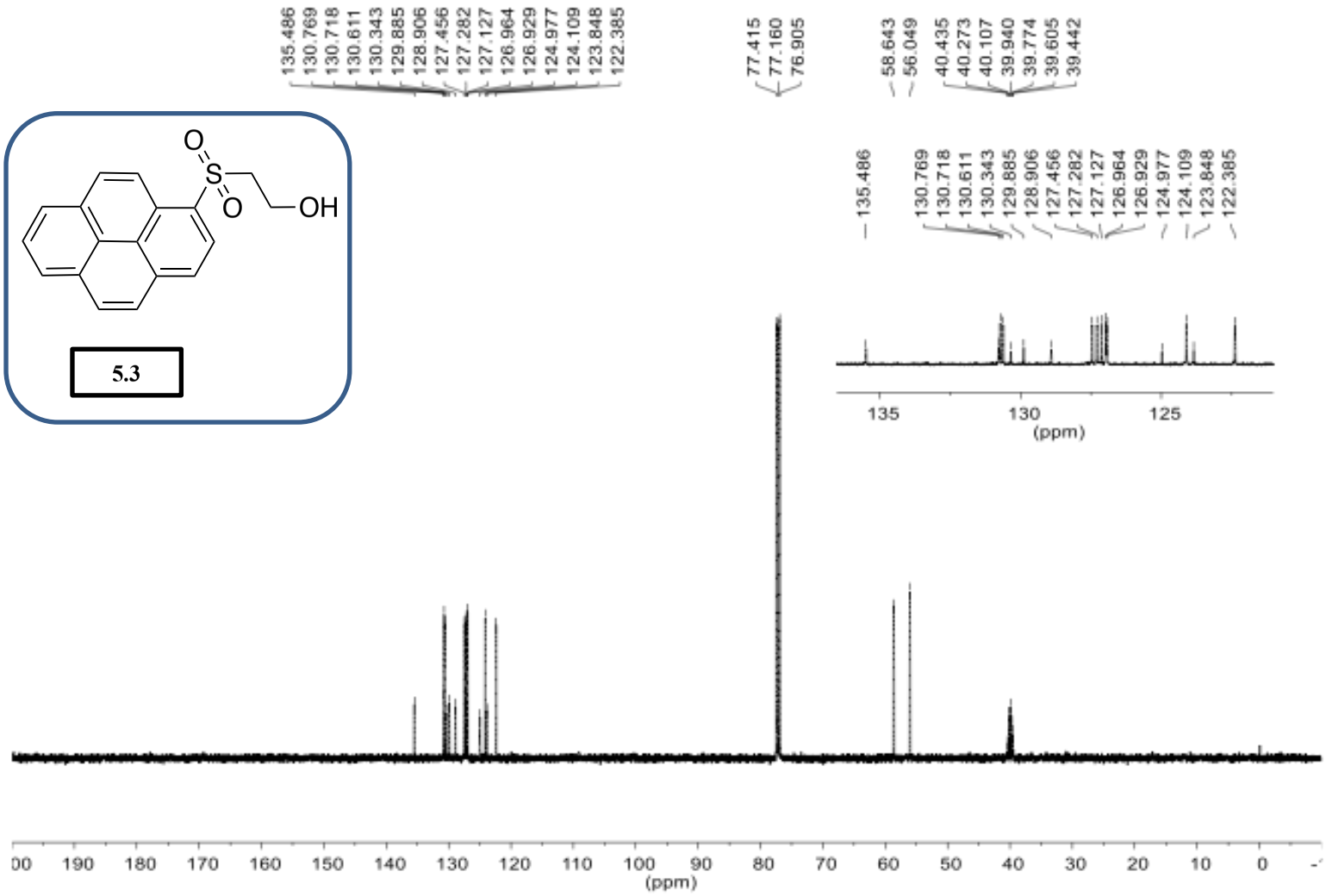
^{31}P NMR spectrum of compound **4.9d** (500 MHz,D_2O)



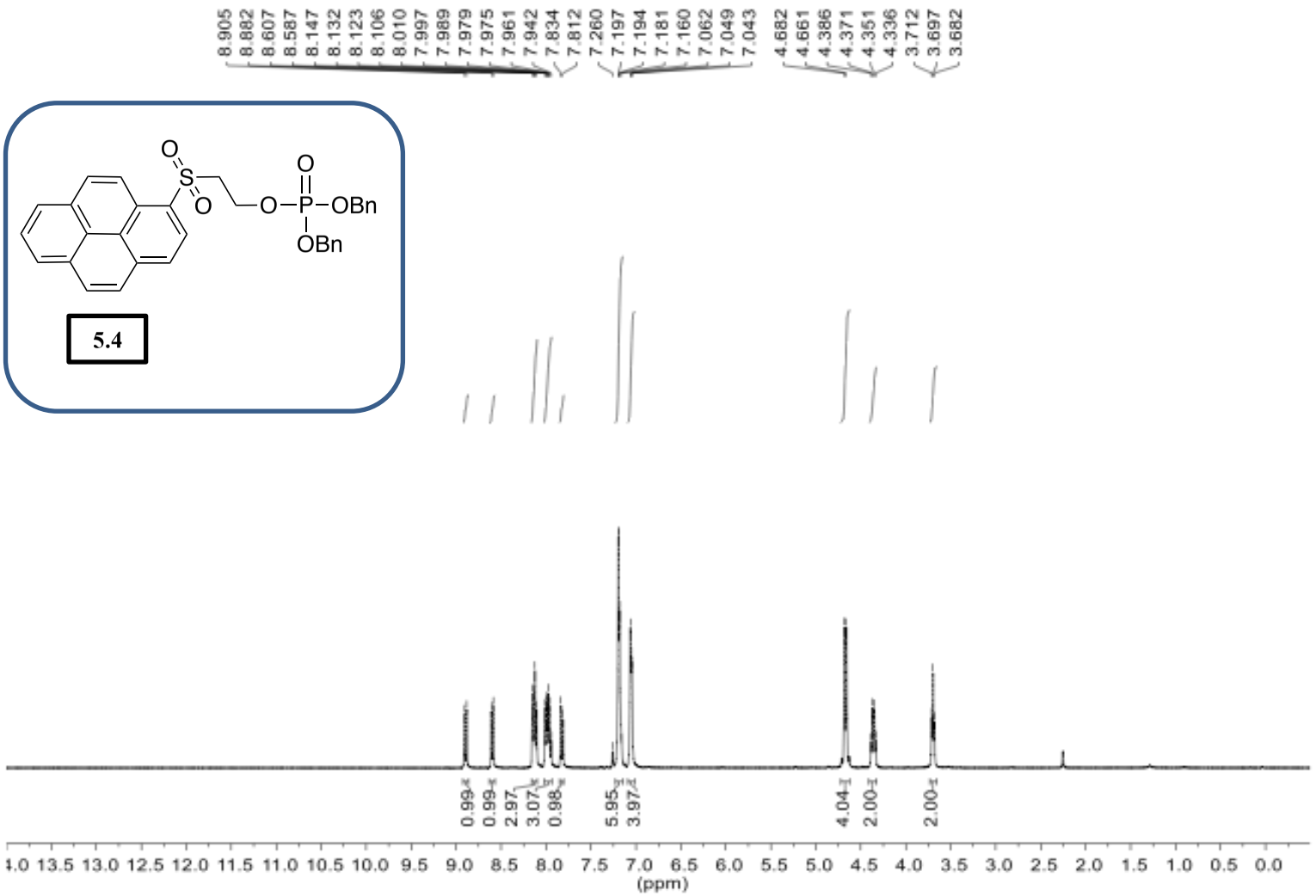
^{1}H NMR spectrum of compound **5.2** (400 MHz,$CDCl_3$)



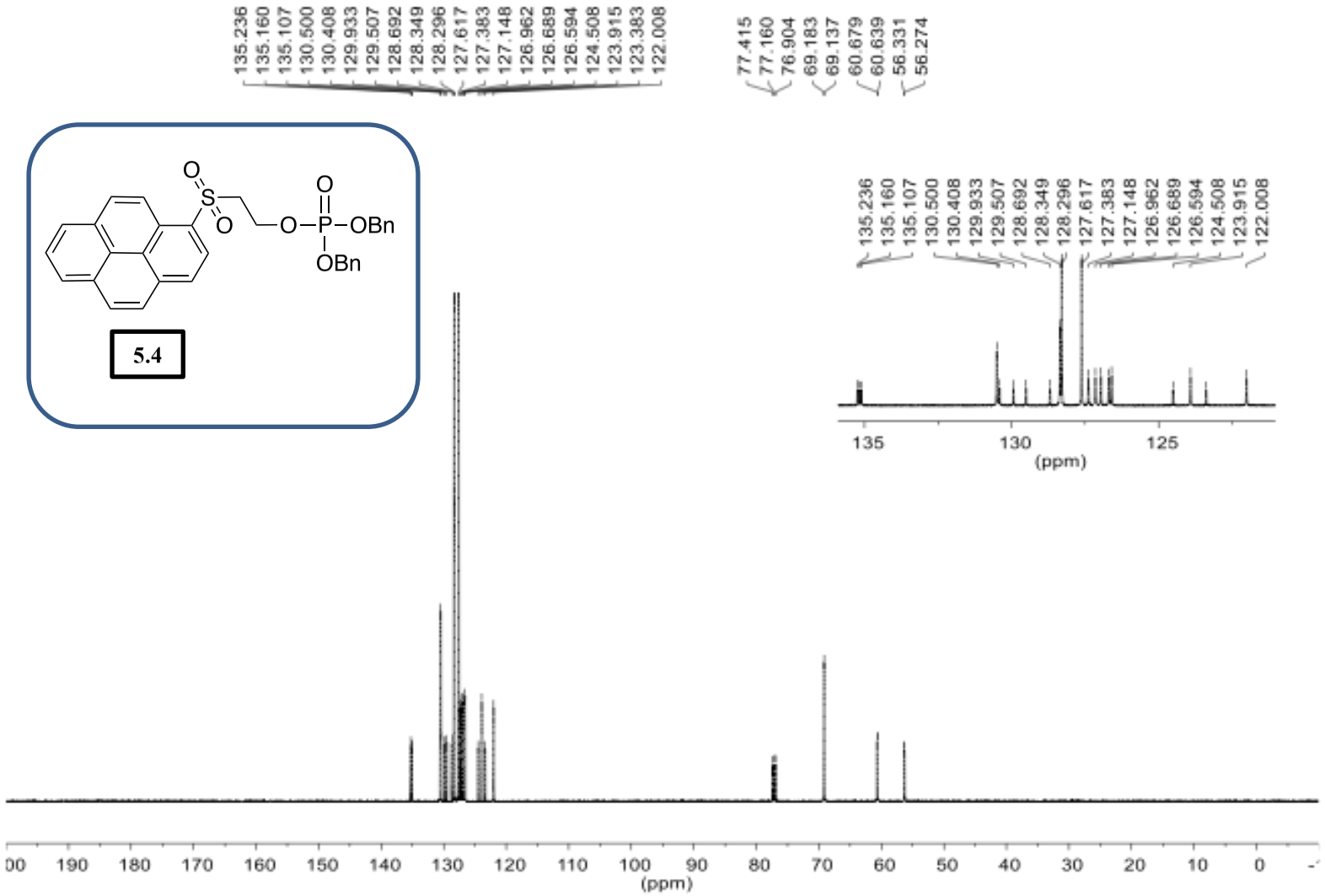
^{1}H NMR spectrum of compound **5.3** (400 MHz,$CDCl_3$)



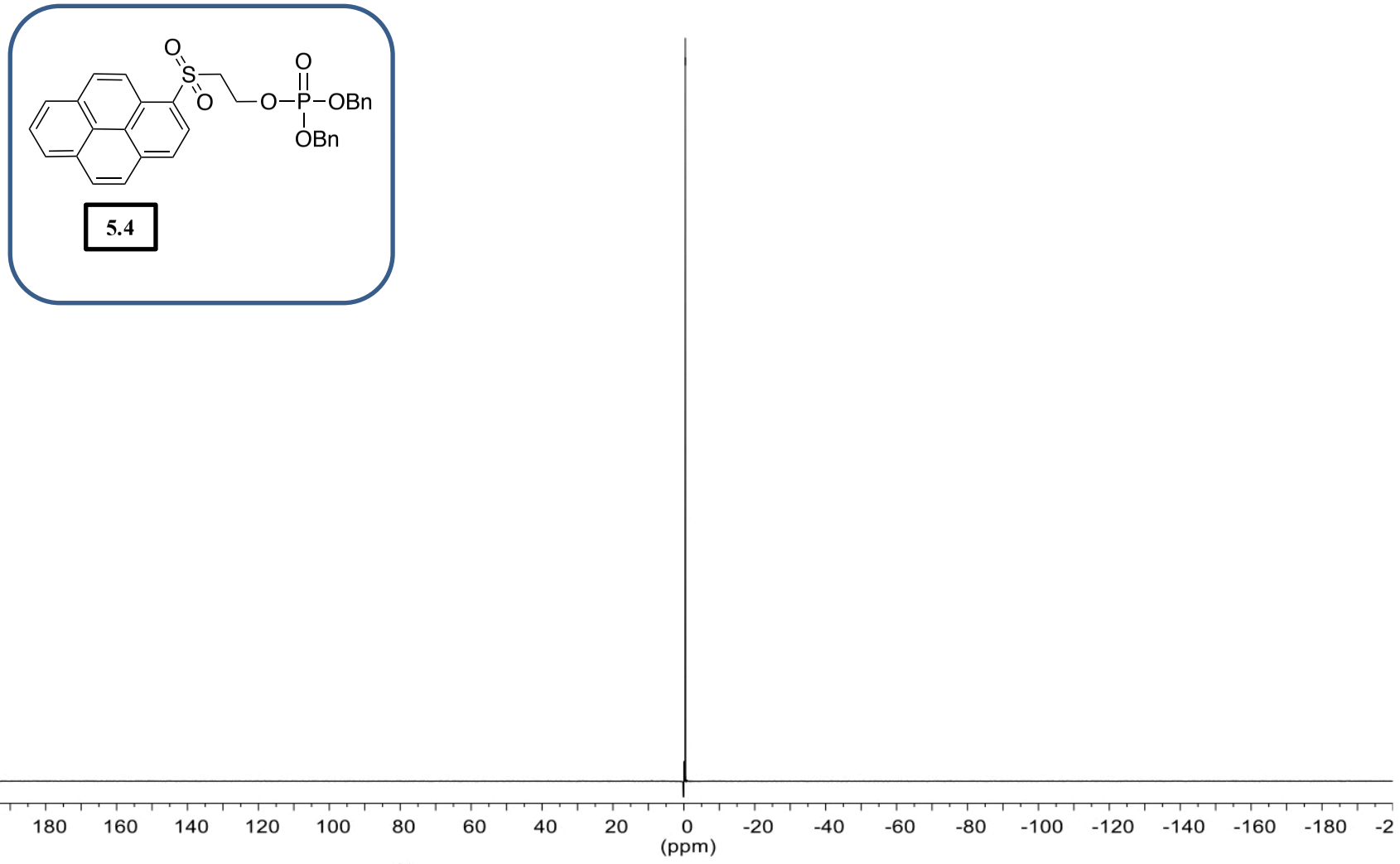
^{13}C NMR spectrum of compound **5.3** (400 MHz,$CDCl_3$)



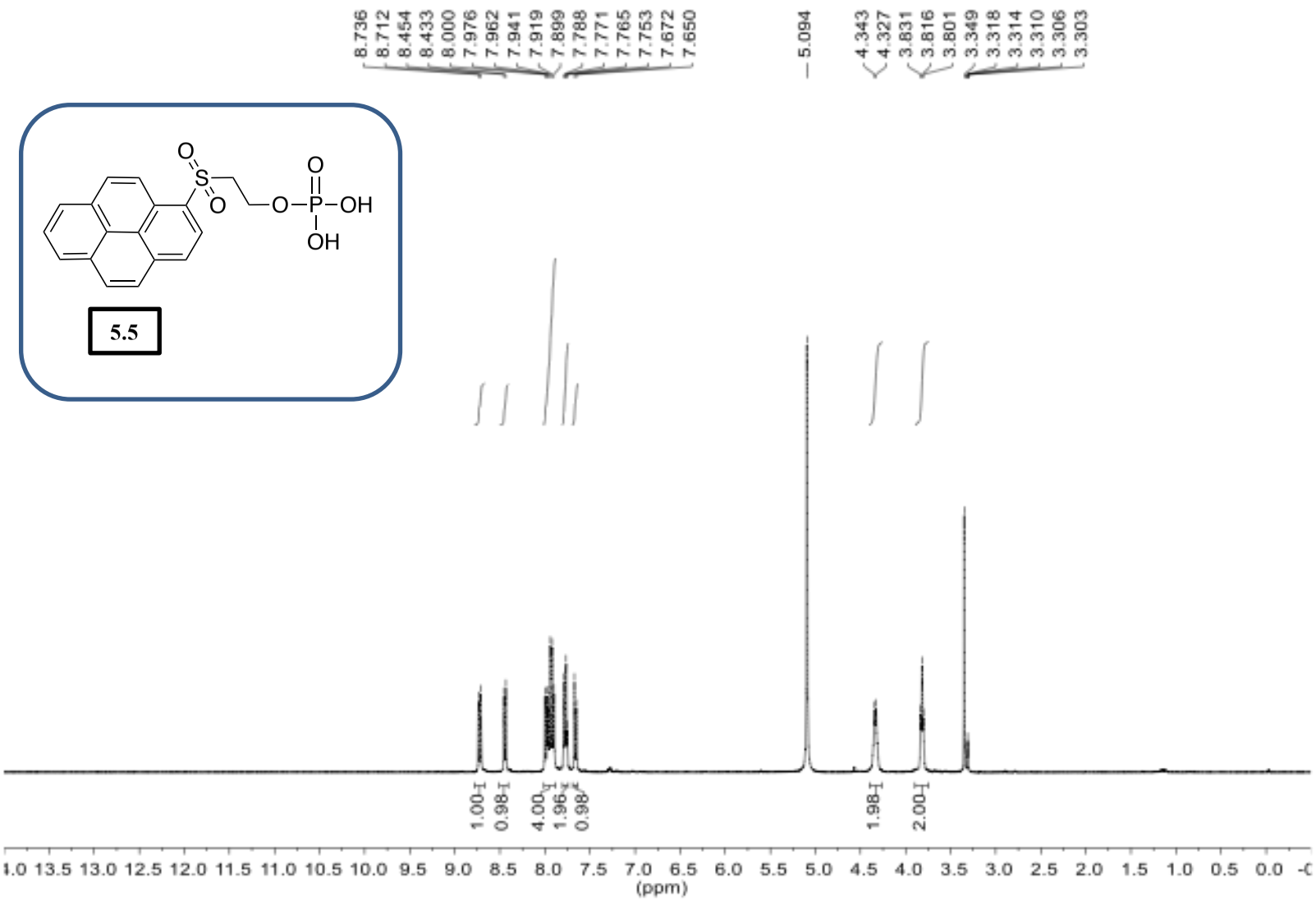
^{1}H NMR spectrum of compound **5.4** (400 MHz,$CDCl_3$)



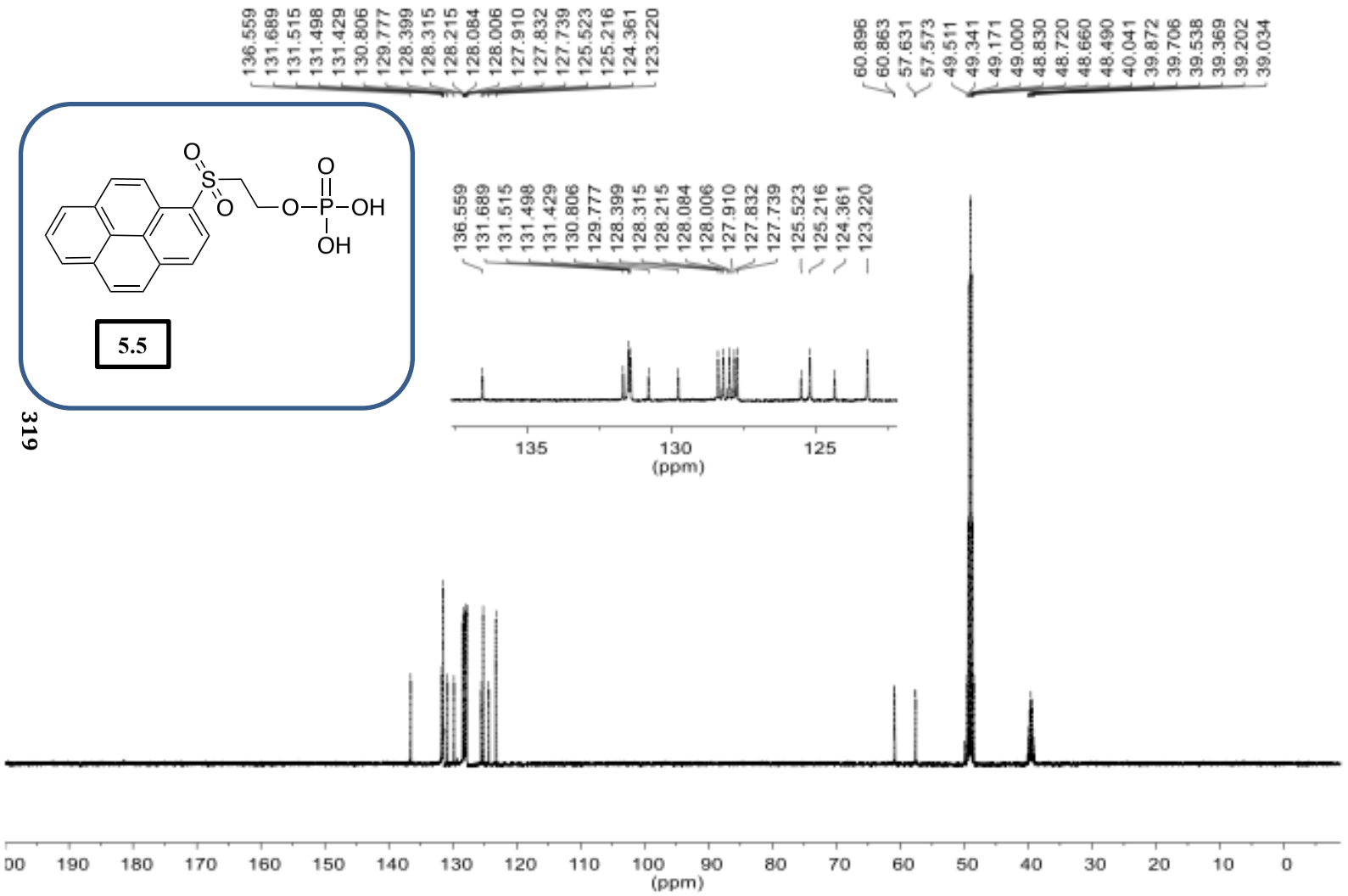
^{13}C NMR spectrum of compound **5.4** (400 MHz,$CDCl_3$)



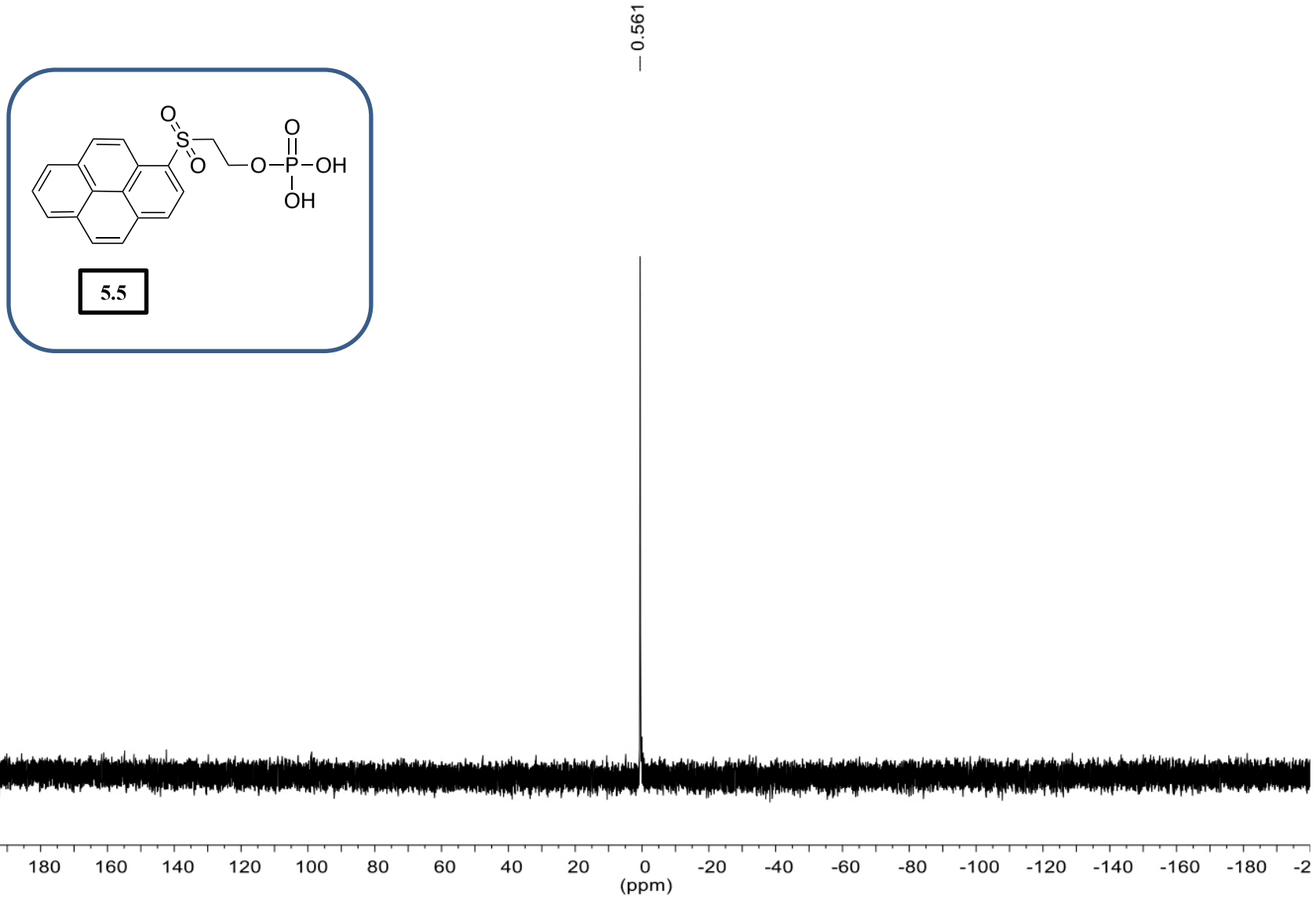
^{31}P NMR spectrum of compound **5.4** (400 MHz,$CDCl_3$)



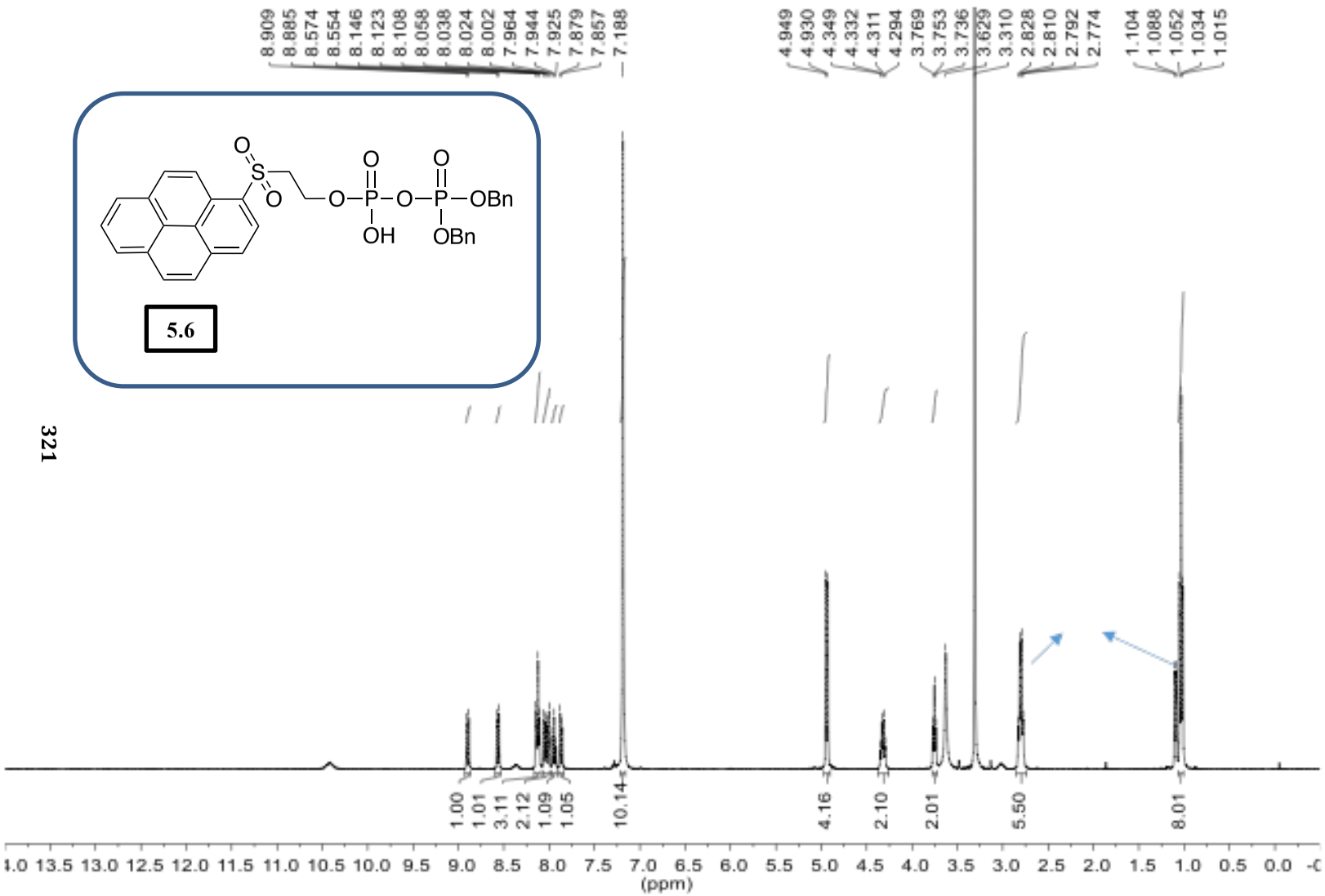
^{1}H NMR spectrum of compound **5.5** (400 MHz,CD_3OD)



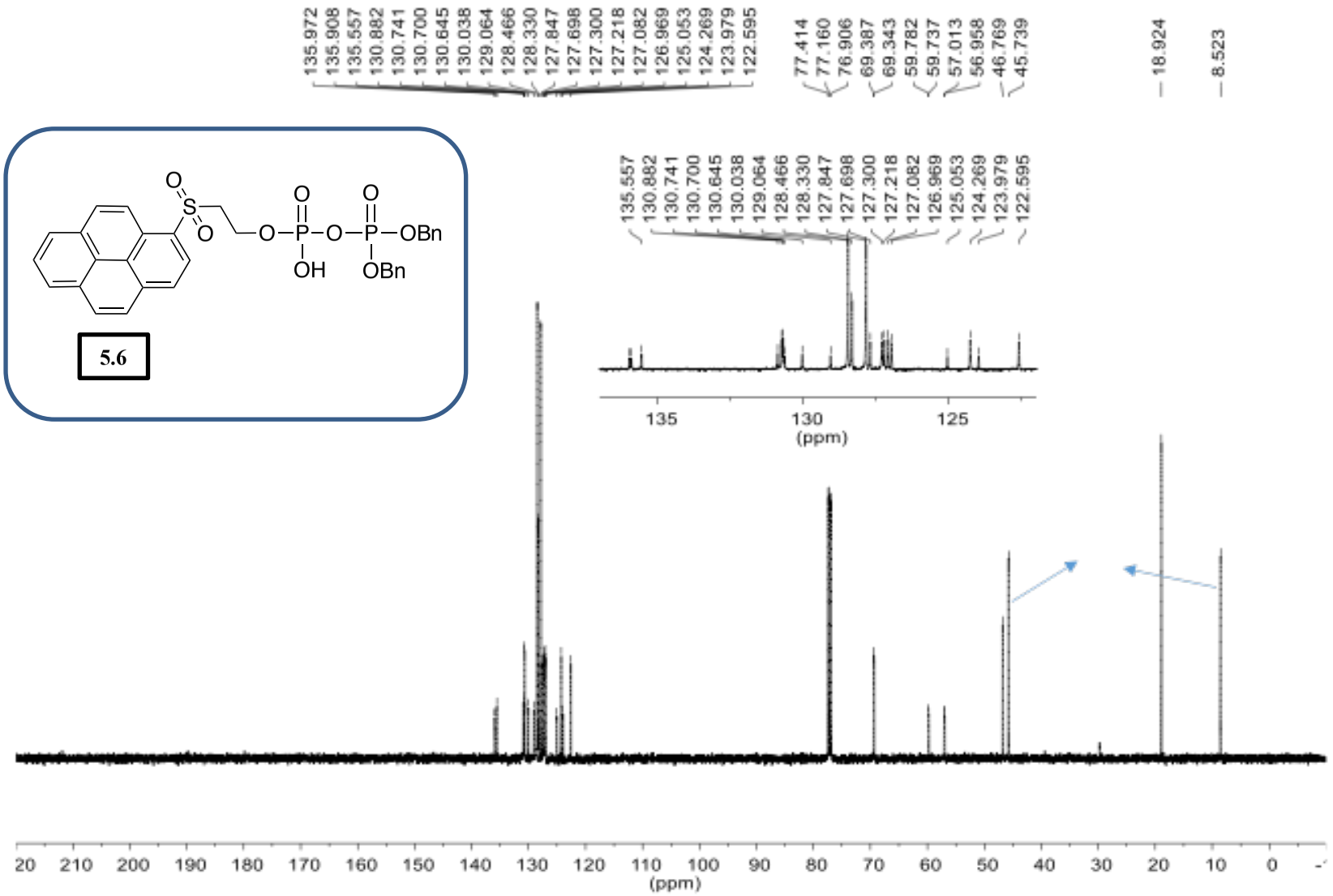
^{13}C NMR spectrum of compound **5.5** (400 MHz,CD_3OD)



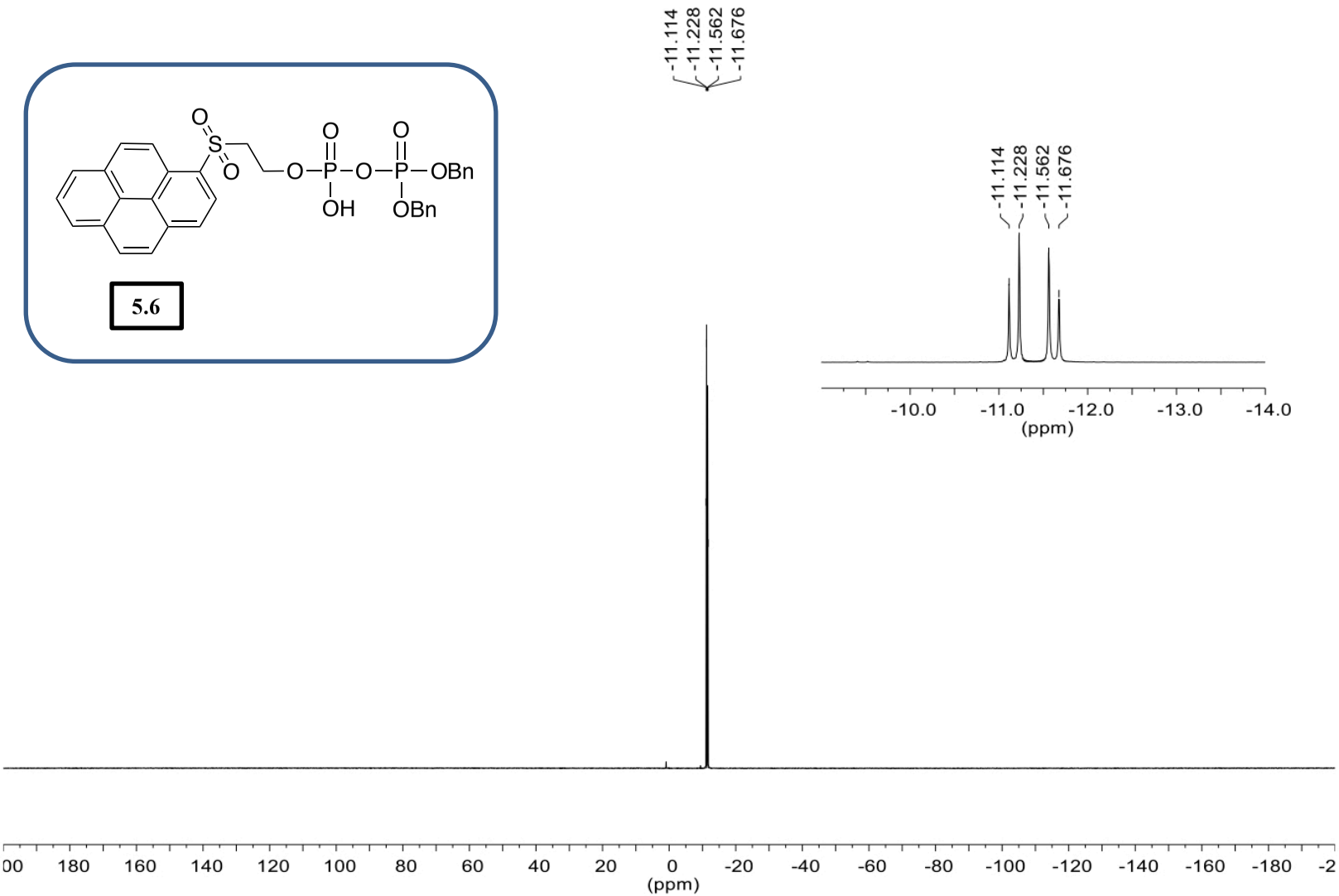
^{31}P NMR spectrum of compound **5.5** (400 MHz,CD_3OD)



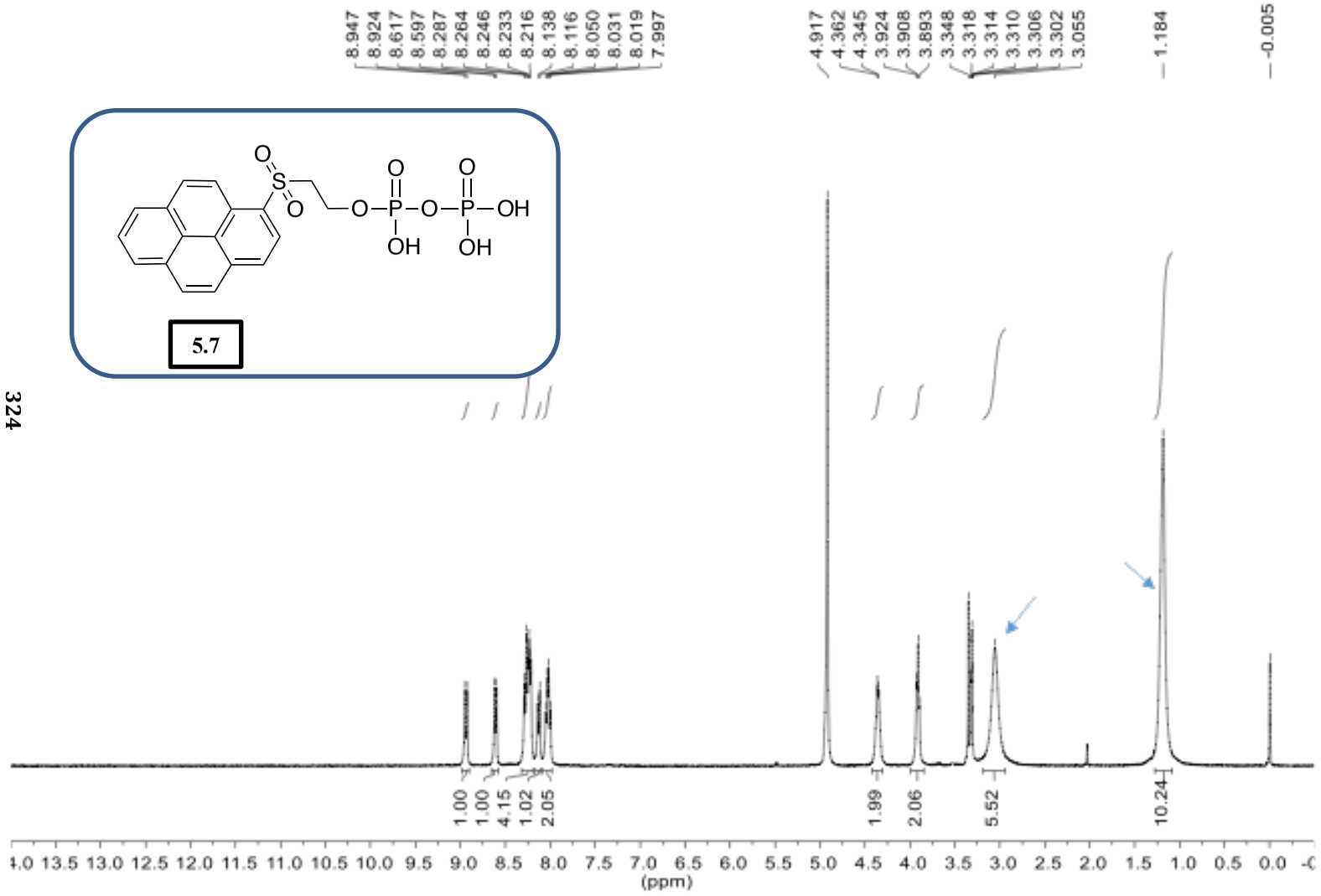
^{1}H NMR spectrum of compound **5.6** (400 MHz,CD_3OD)



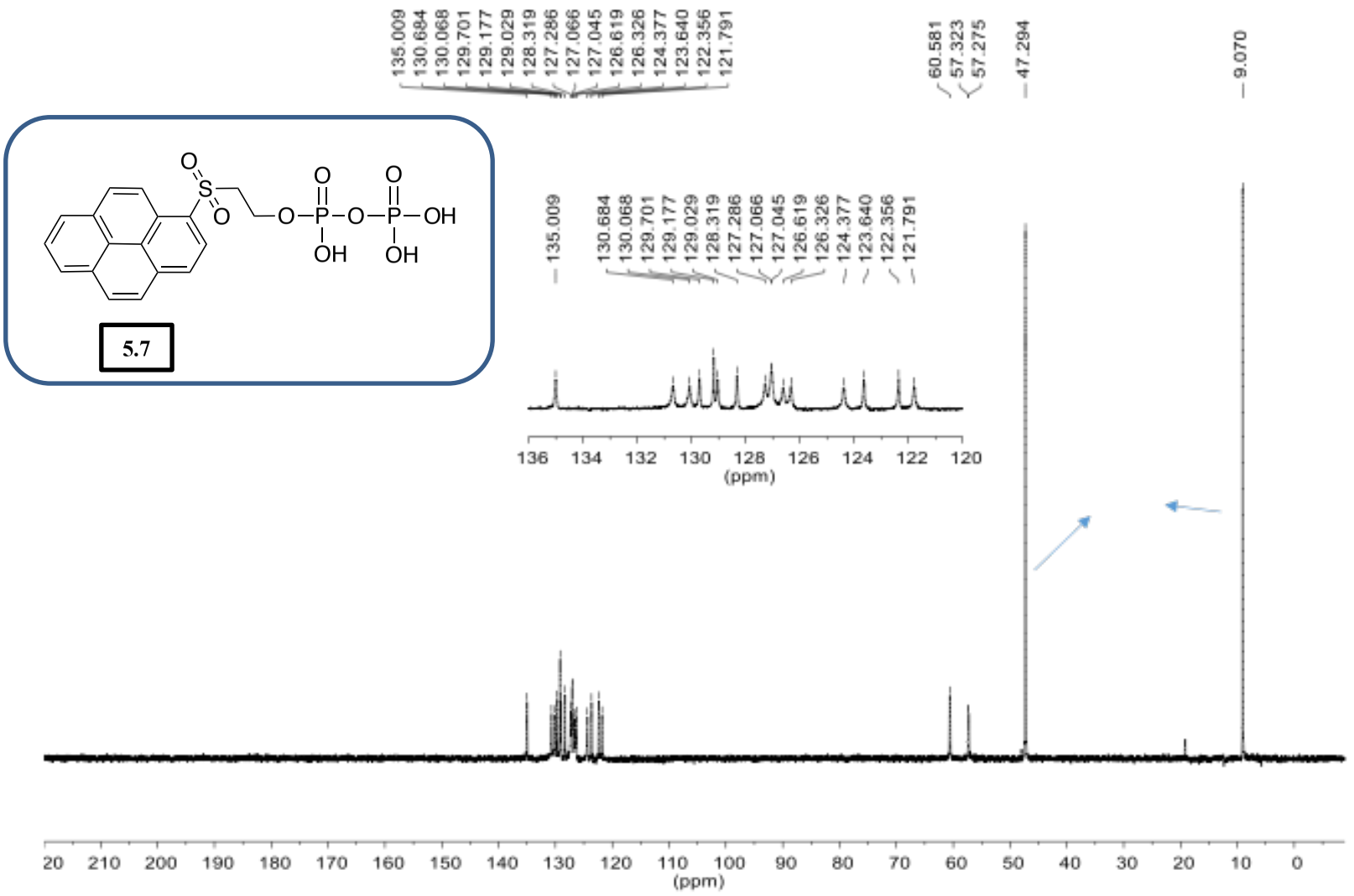
^{13}C NMR spectrum of compound **5.6** (500 MHz,CD_3OD)



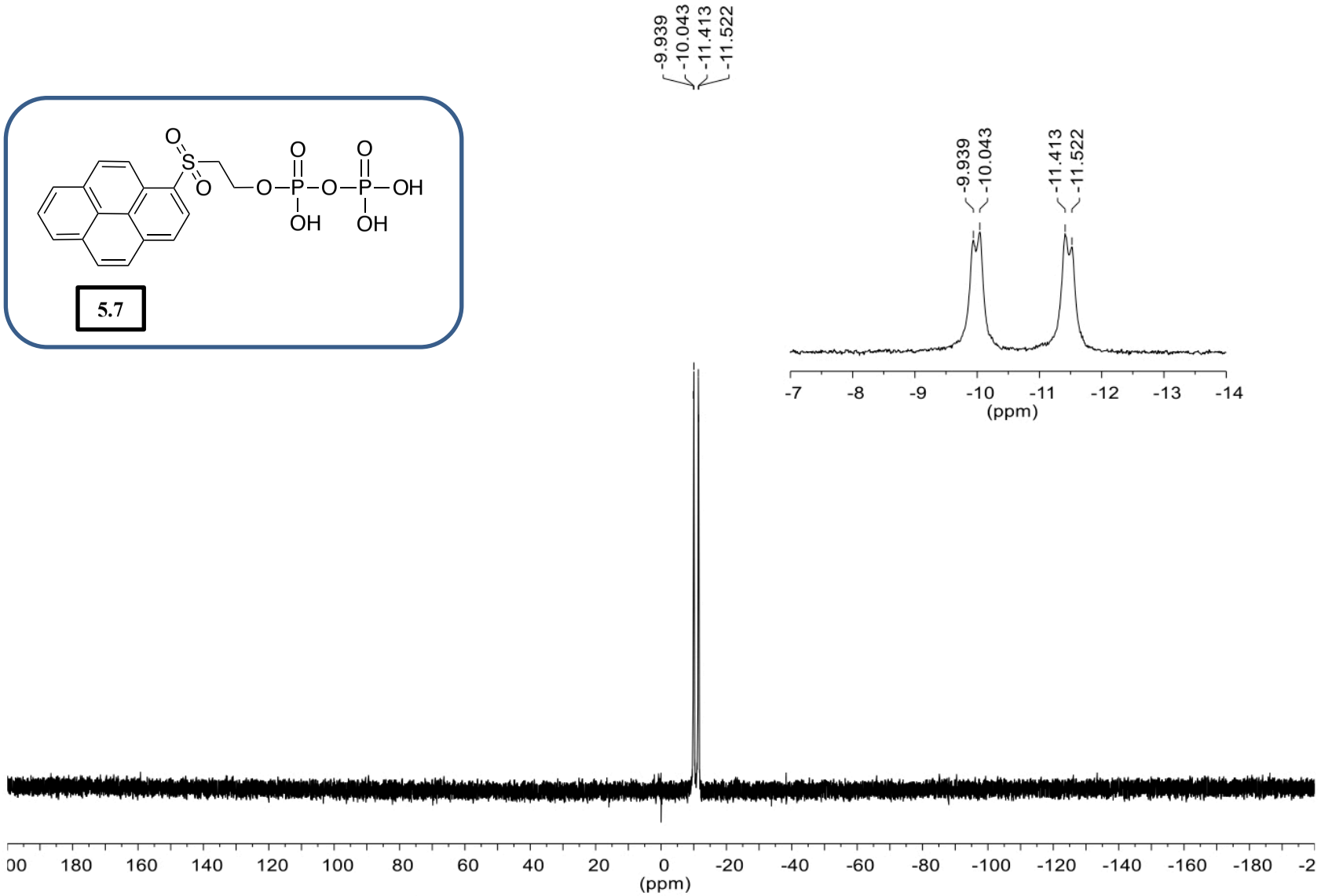
^{31}P NMR spectrum of compound **5.6** (500 MHz,CD_3OD)



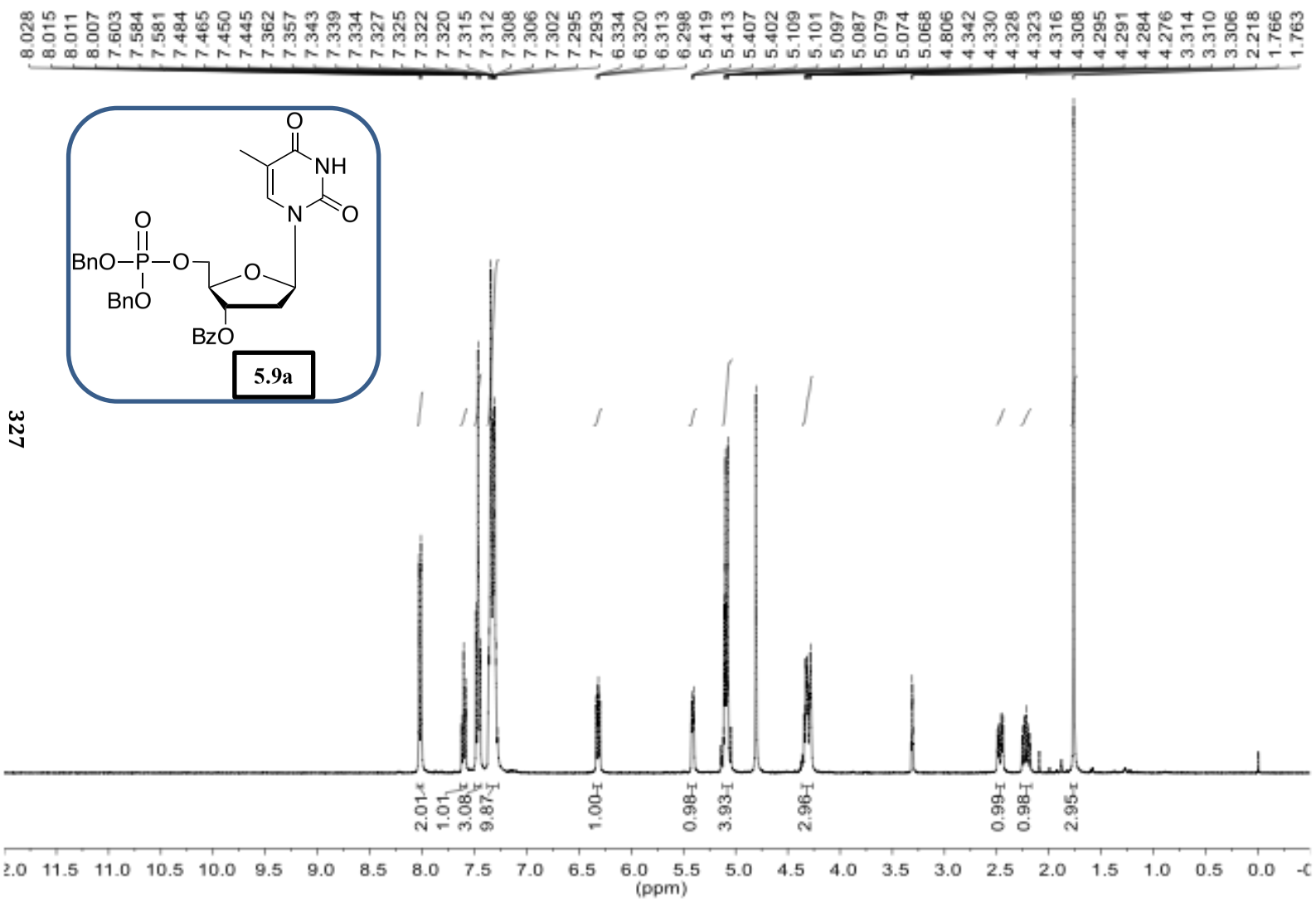
^{1}H NMR spectrum of compound **5.7** (400 MHz,CD_3OD)



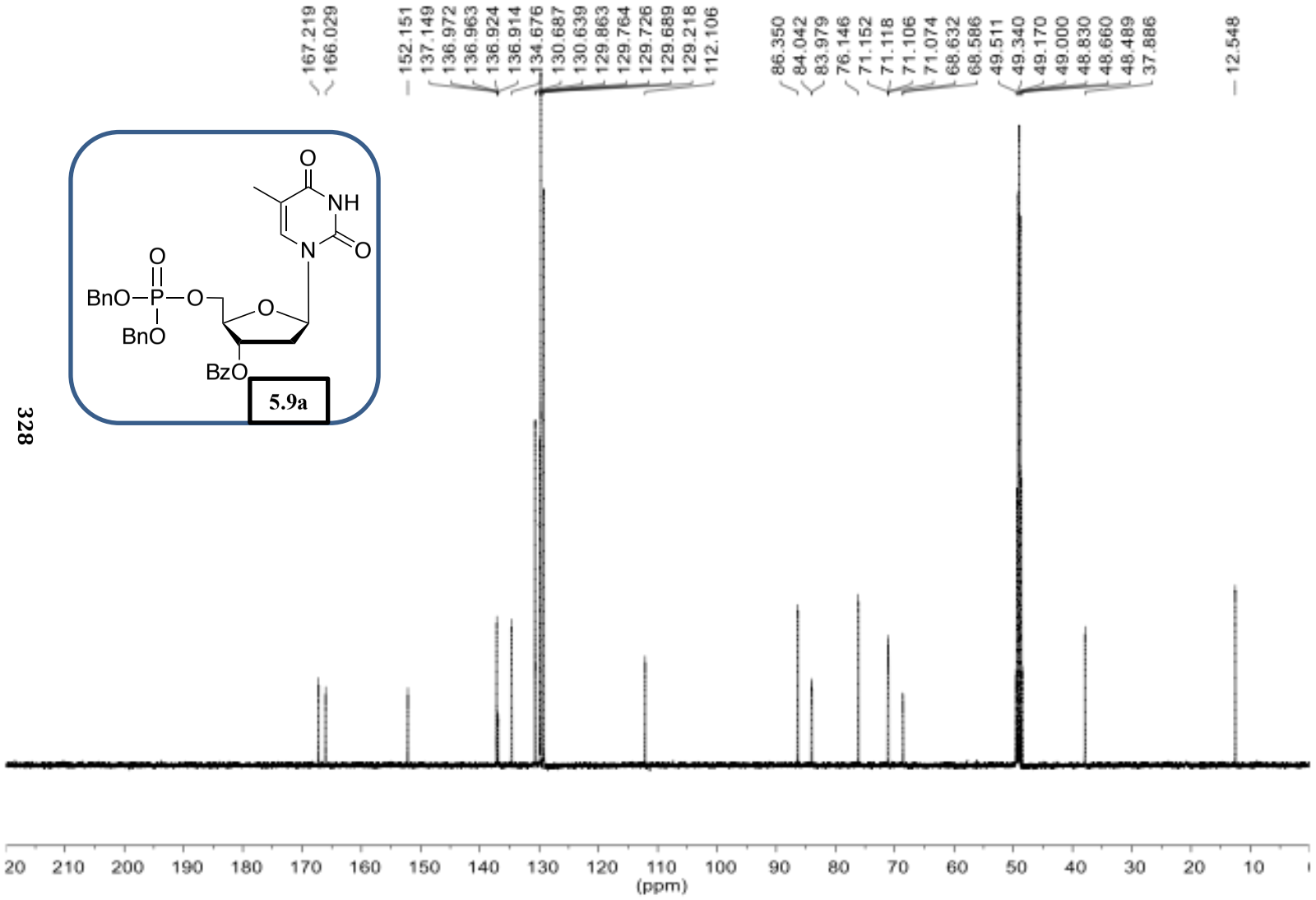
^{13}C NMR spectrum of compound **5.7** (400 MHz,CD_3OD)



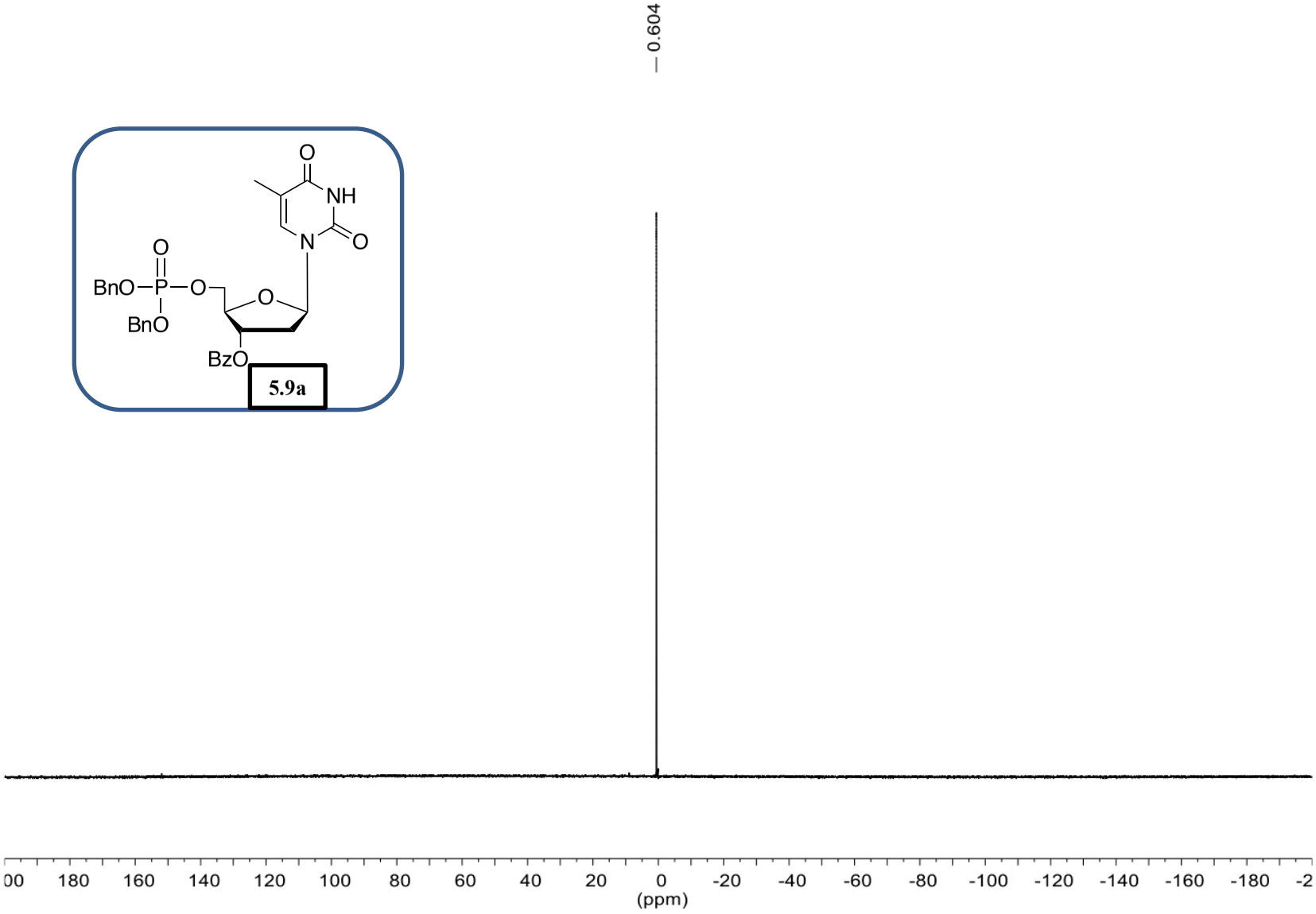
^{31}P NMR spectrum of compound **5.7** (400 MHz,CD_3OD)



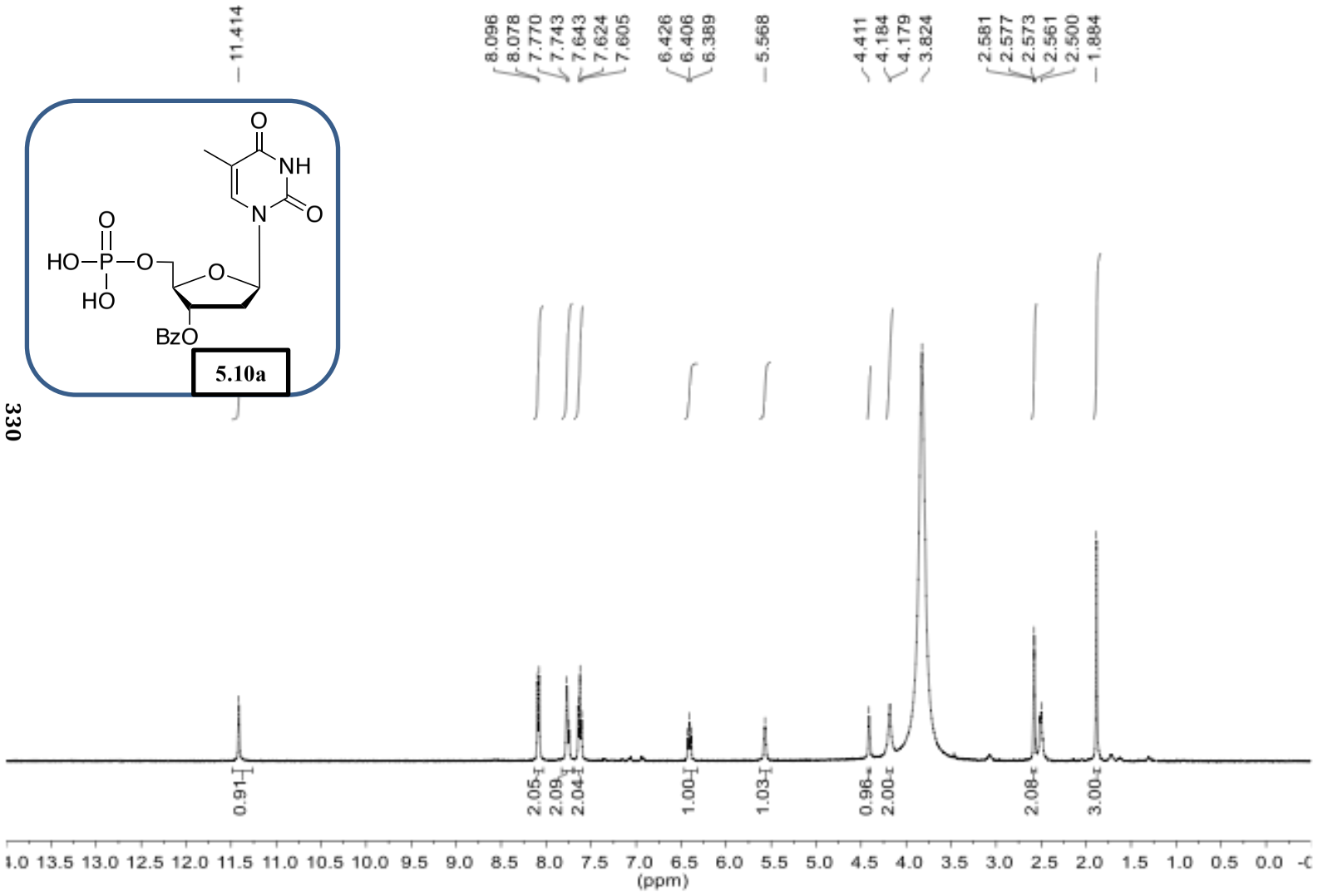
^{1}H NMR spectrum of compound **5.9a** (400 MHz,$CDCl_3$)



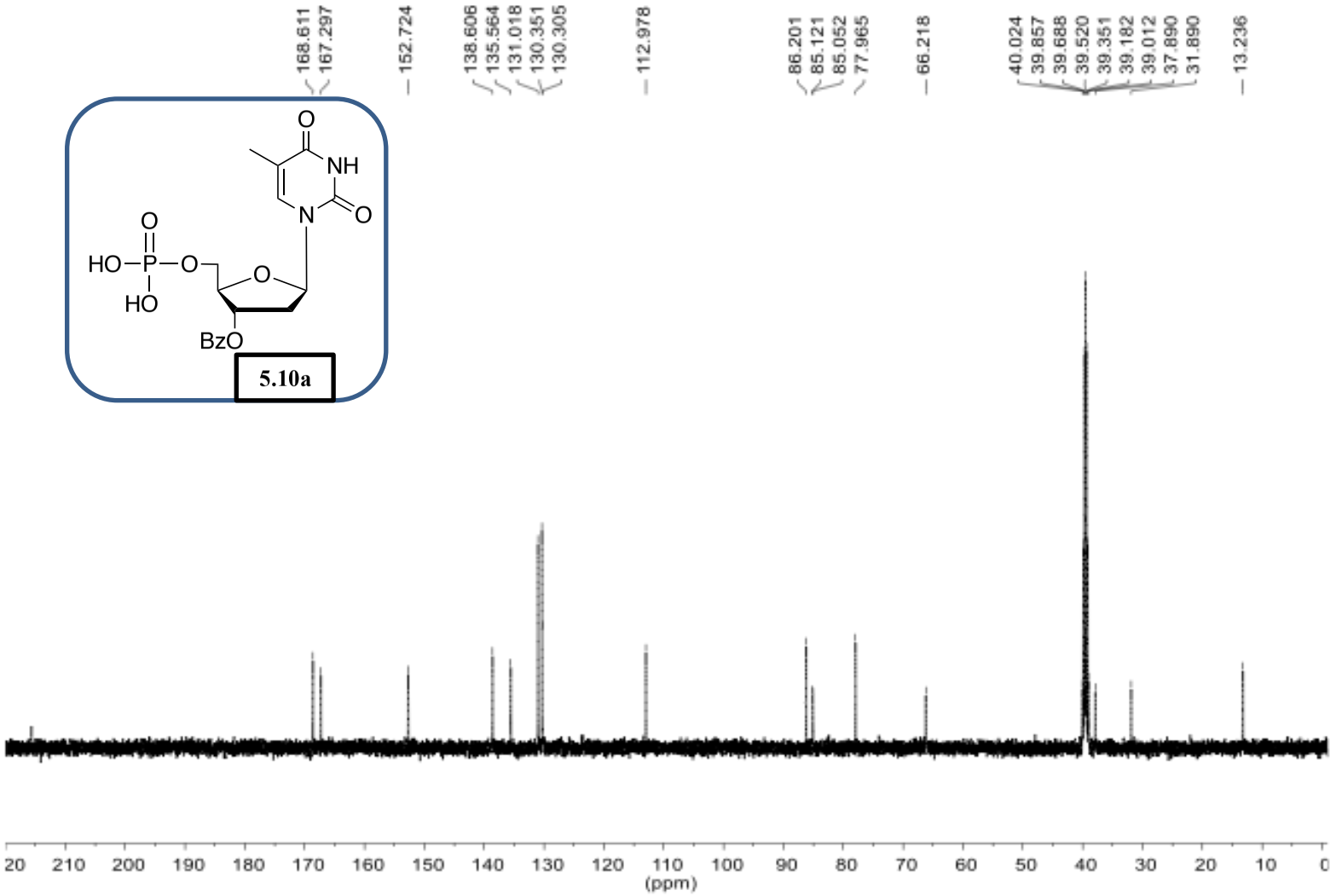
^{13}C NMR spectrum of compound **5.9a** (500 MHz,$CDCl_3$)



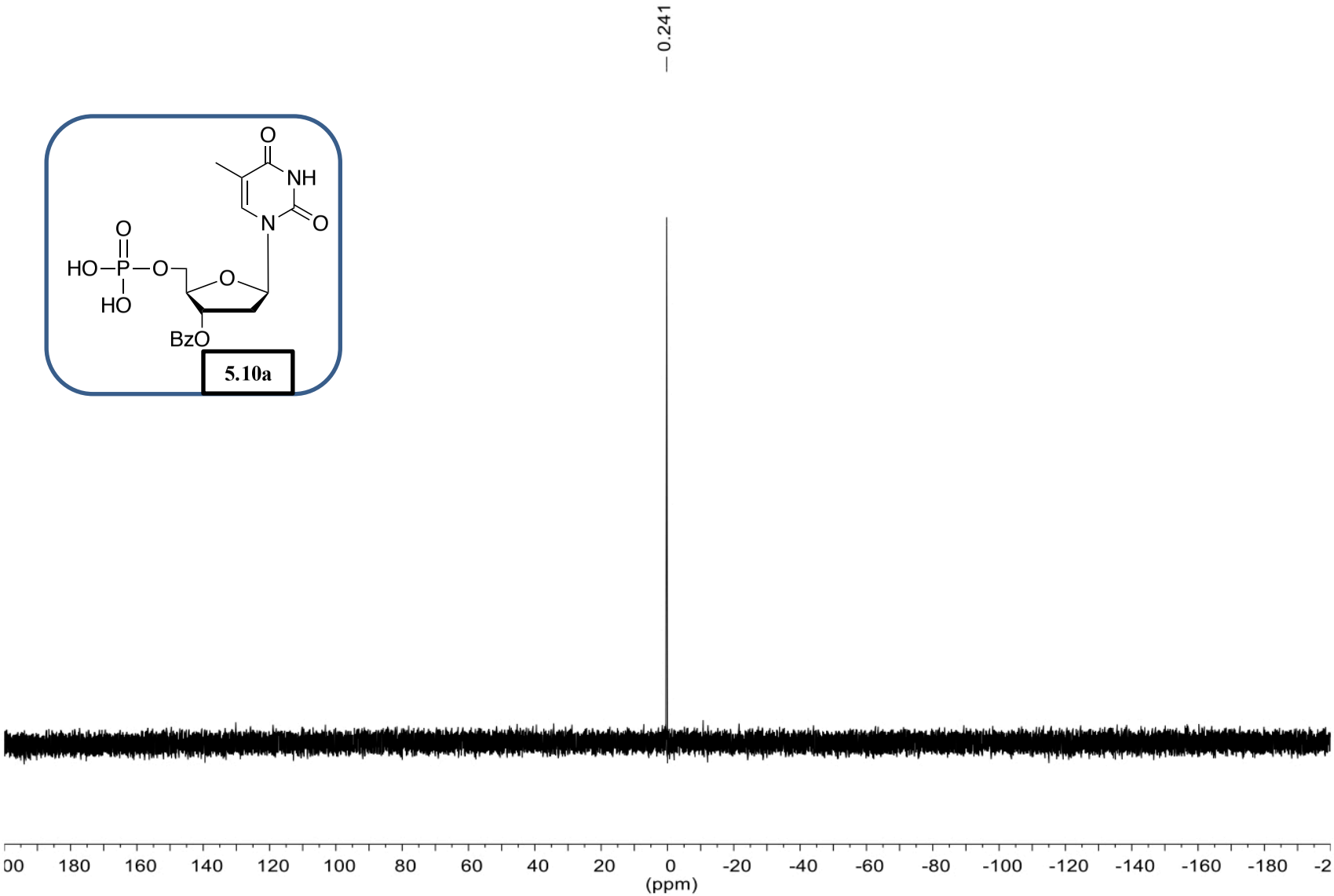
^{31}P NMR spectrum of compound **5.9a** (500 MHz,$CDCl_3$)



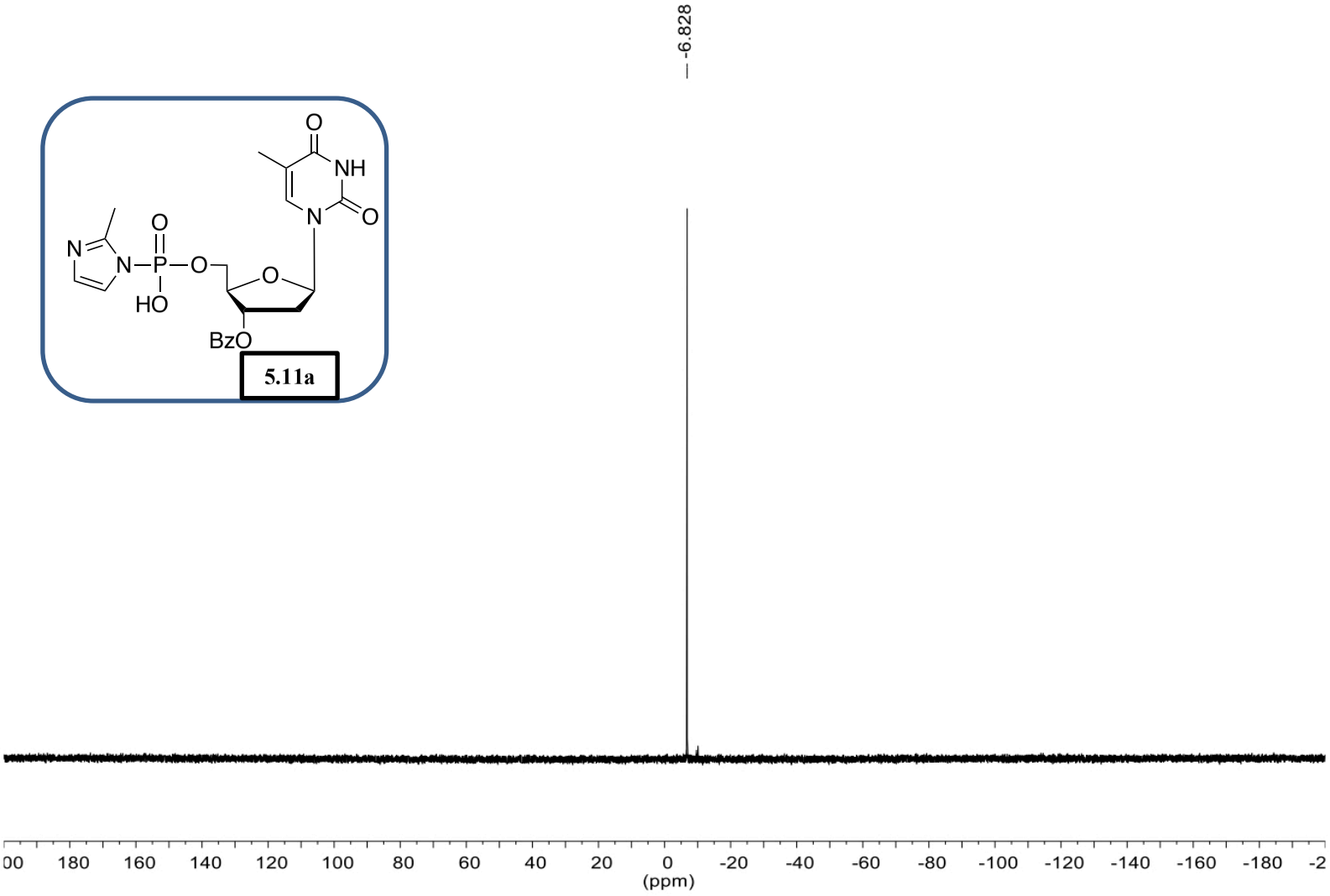
^{1}H NMR spectrum of compound **5.10a** (500 MHz,CD_3OD)



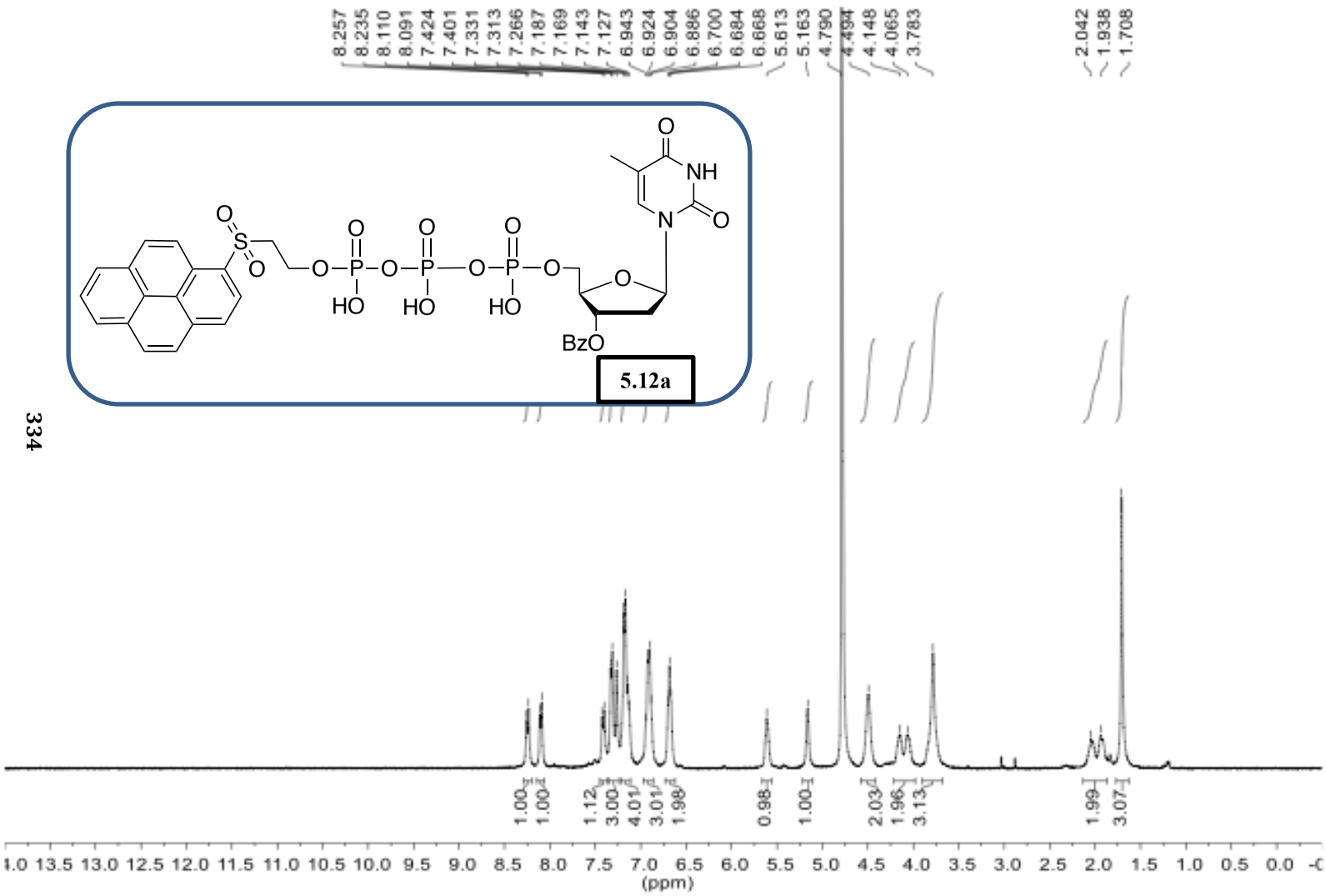
^{13}C NMR spectrum of compound **5.10a** (500 MHz,CD_3OD)



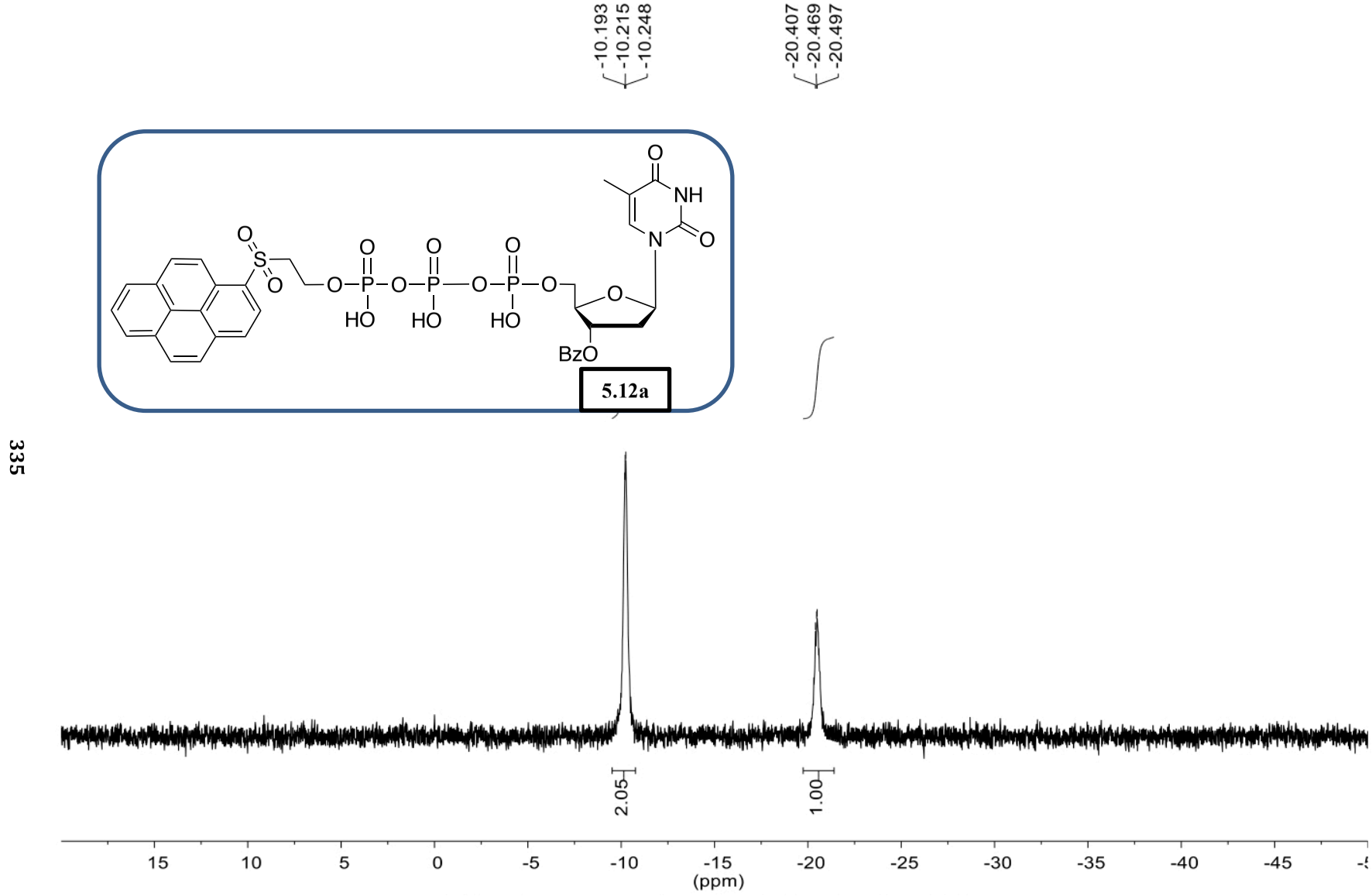
^{31}P NMR spectrum of compound **5.10a** (500 MHz,CD_3OD)



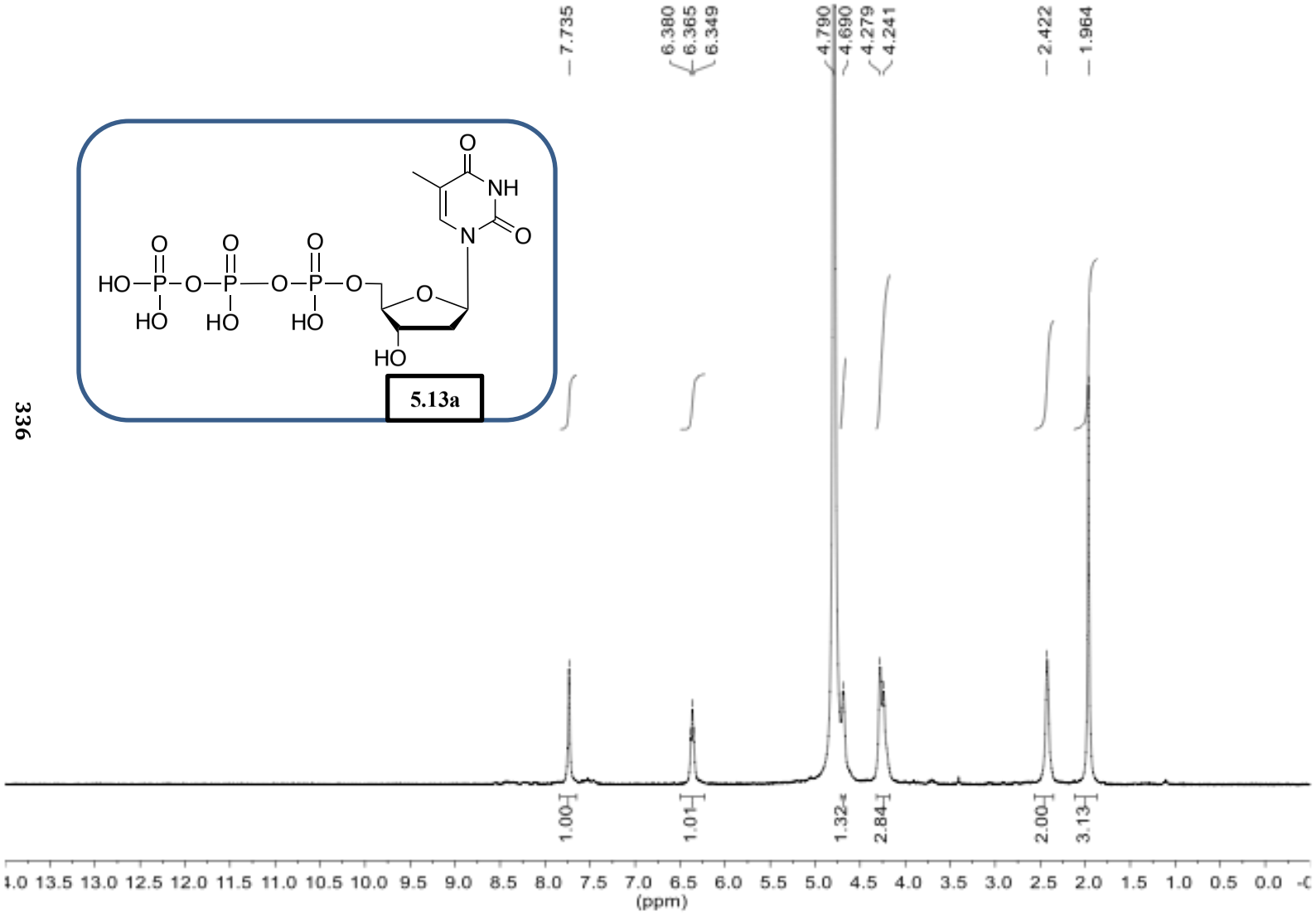
^{31}P NMR spectrum of compound **5.11a** (400 MHz,CD_3OD)



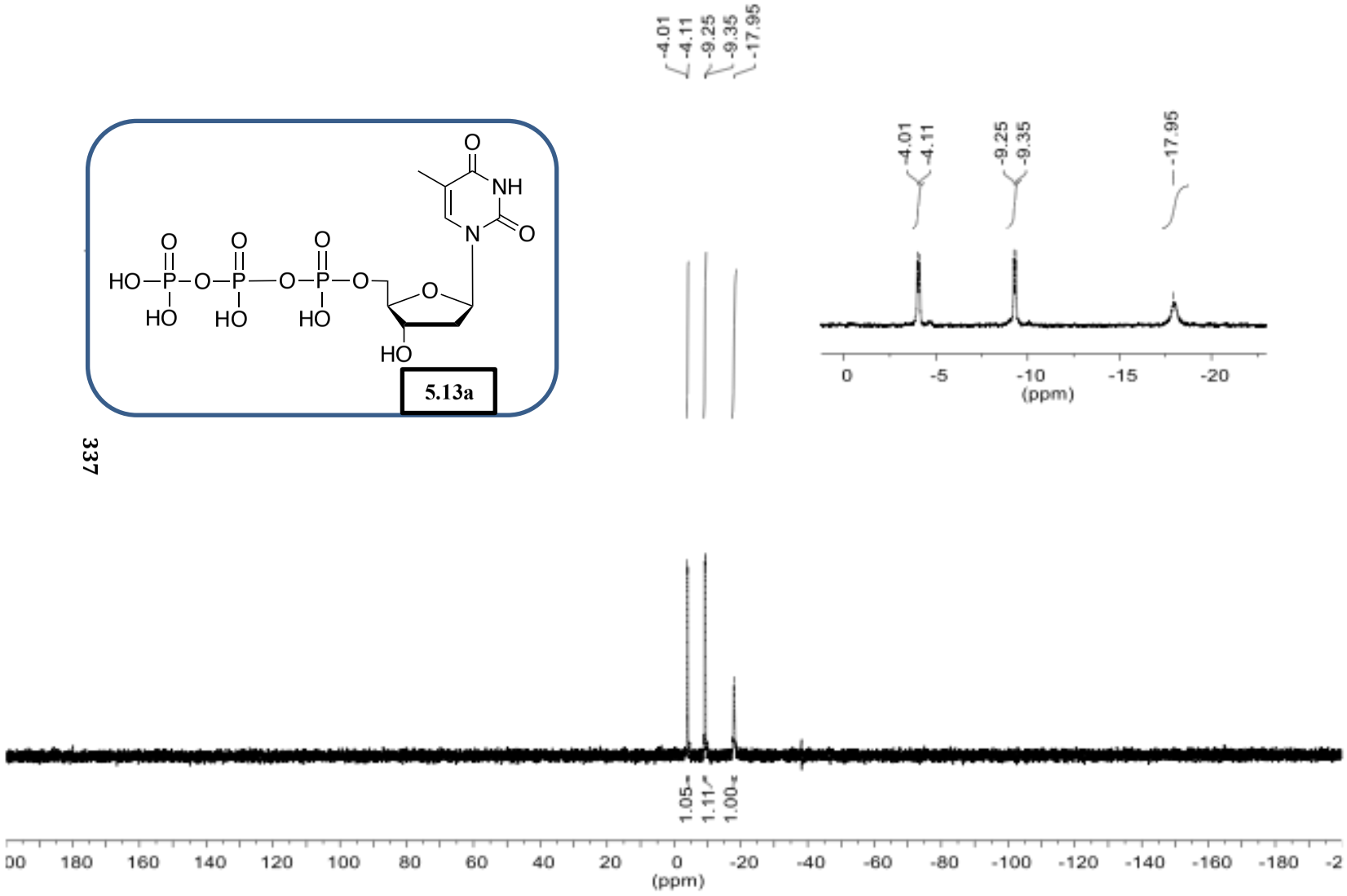
^{1}H NMR spectrum of compound **5.12a** (400 MHz,D_2O)



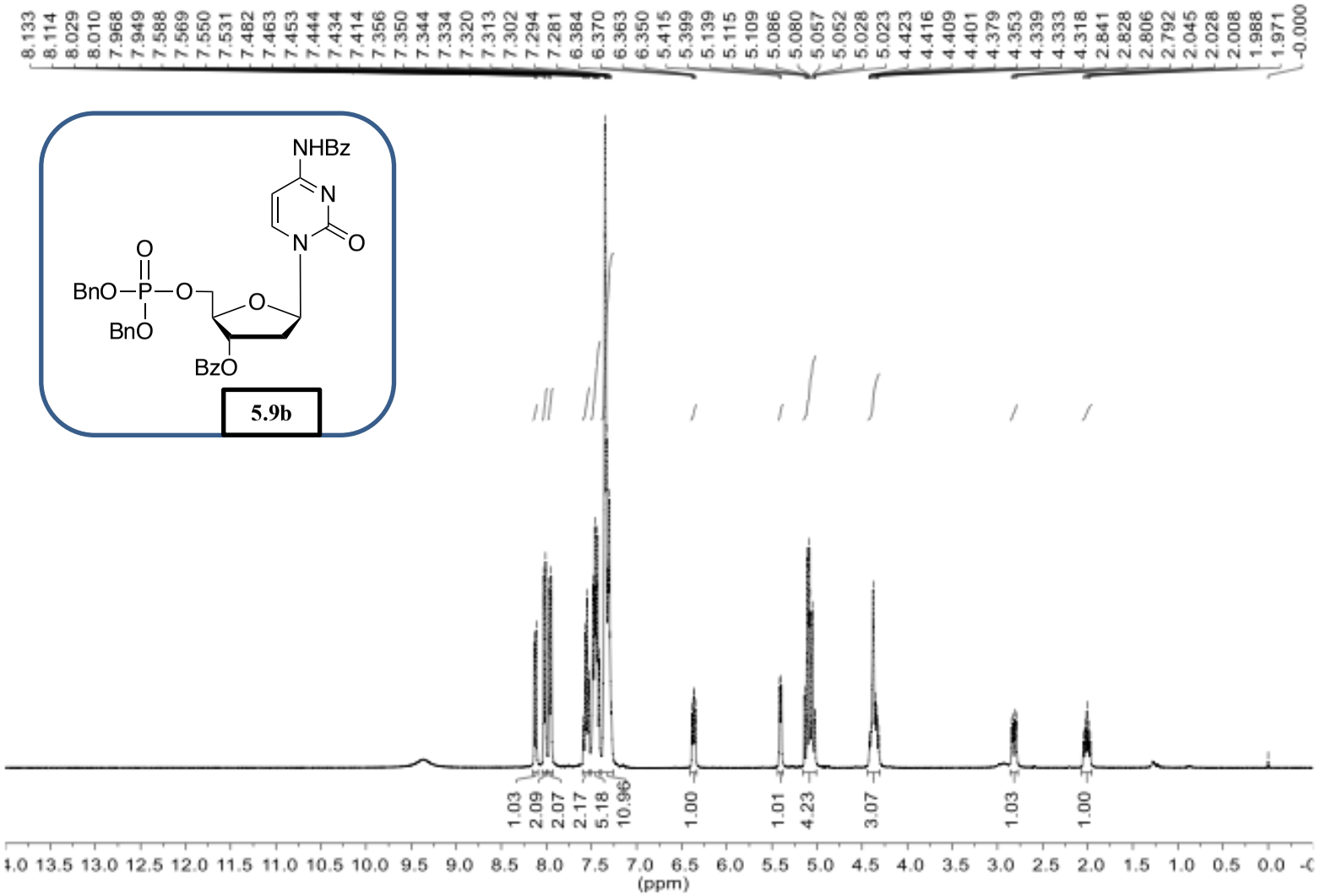
^{31}P NMR spectrum of compound **5.12a** (400 MHz,D_2O)



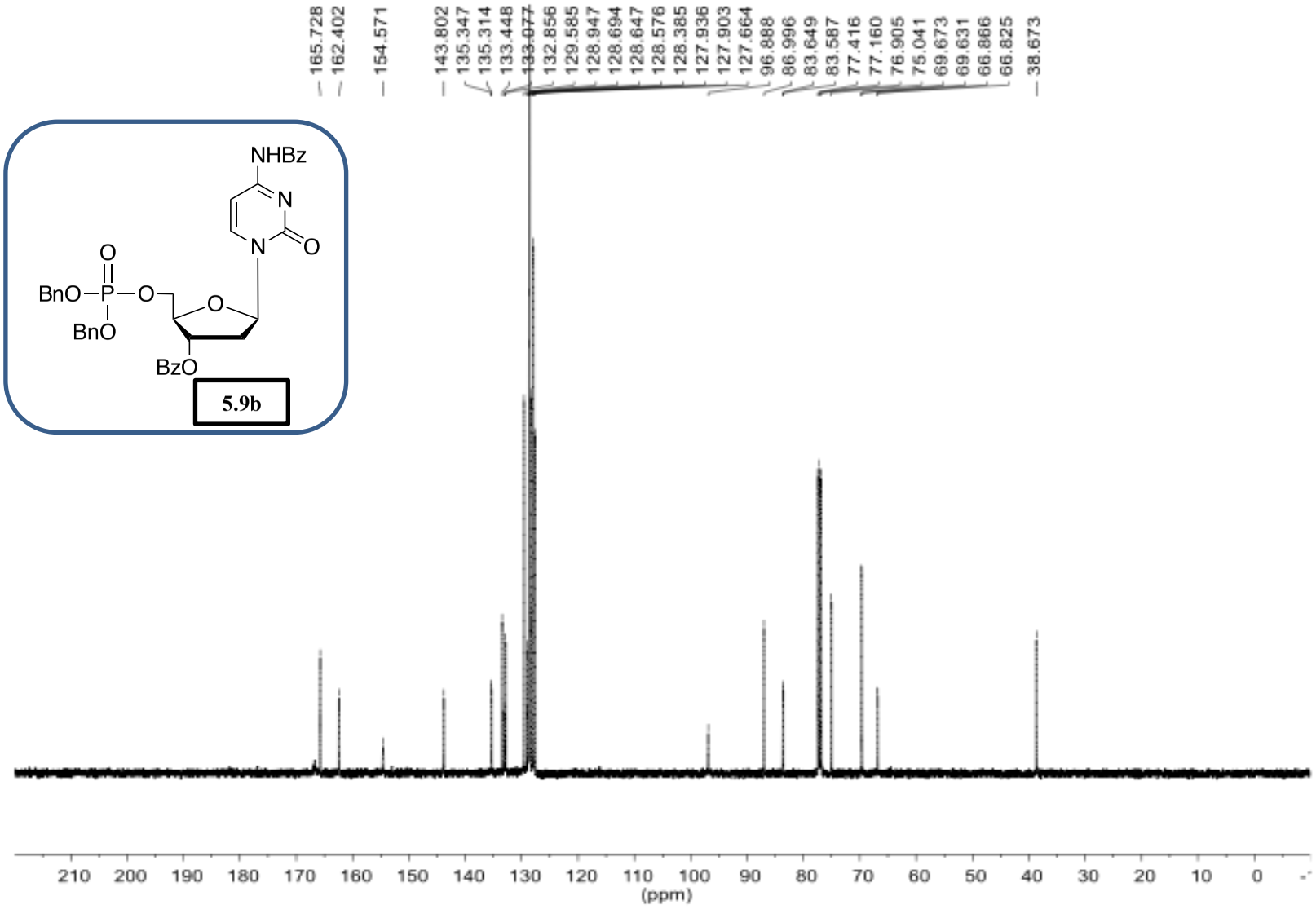
^{1}H NMR spectrum of compound **5.13a** (400 MHz,D_2O)



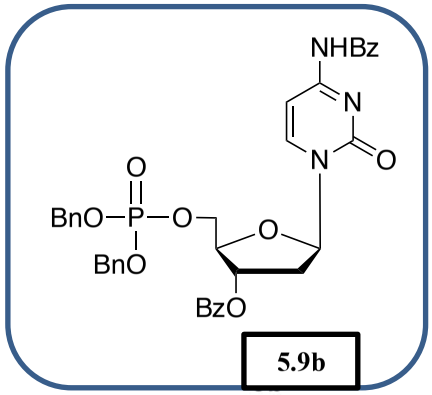
^{31}P NMR spectrum of compound **5.13a** (500 MHz,D_2O)



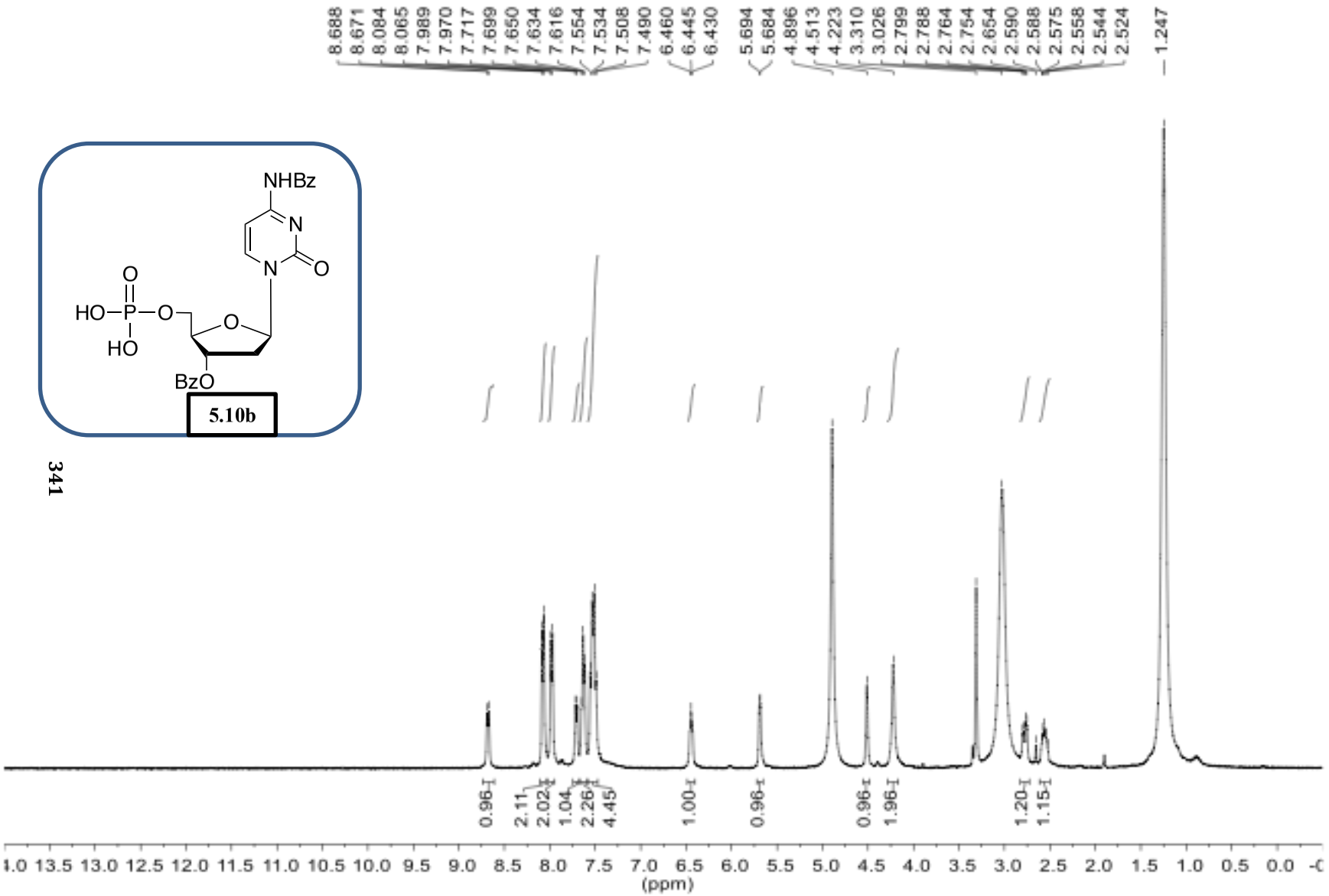
^{1}H NMR spectrum of compound **5.9b** (500 MHz,$CDCl_3$)



^{13}C NMR spectrum of compound **5.9b** (500 MHz,$CDCl_3$)

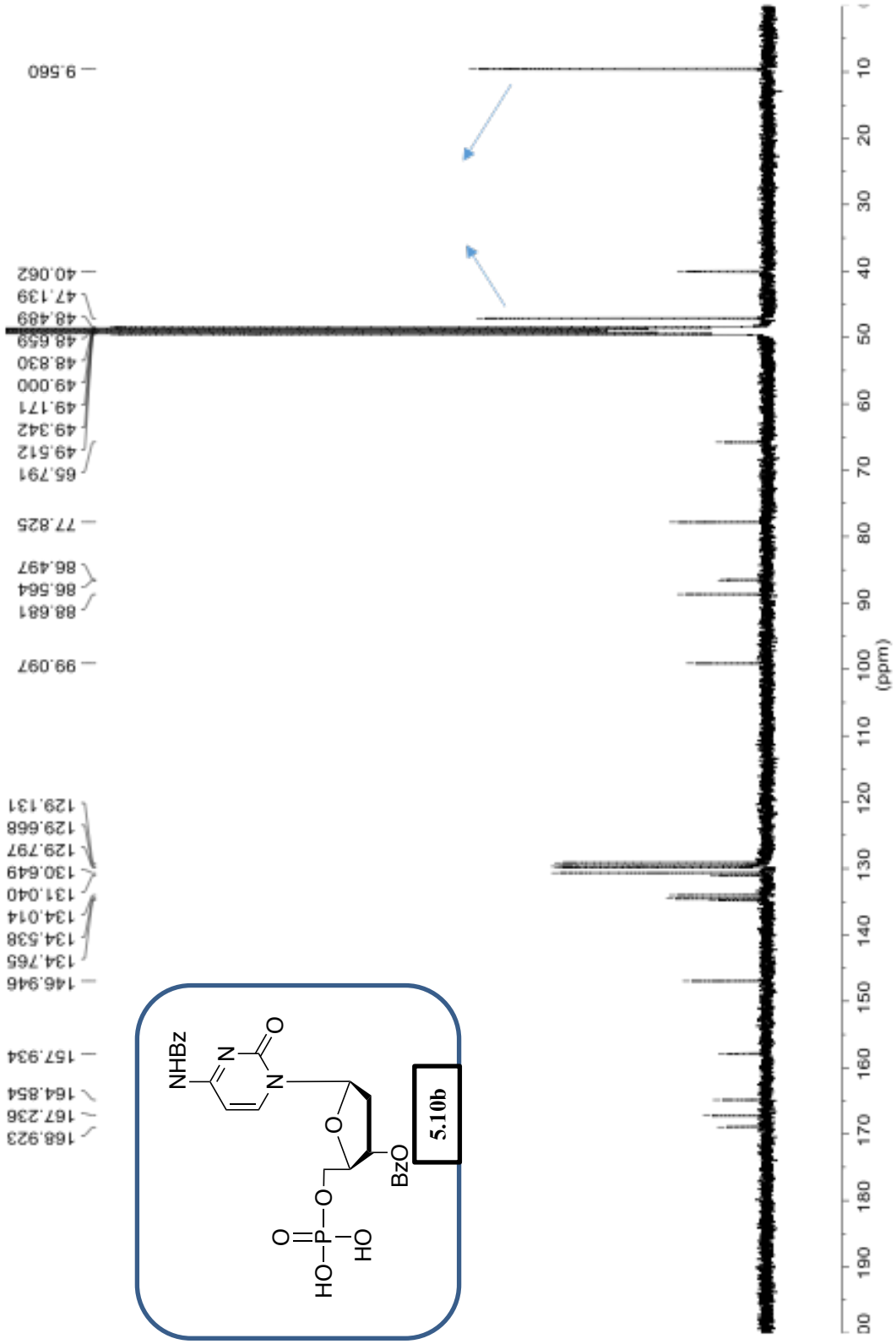


^{31}P NMR spectrum of compound **5.9b** (500 MHz,$CDCl_3$)

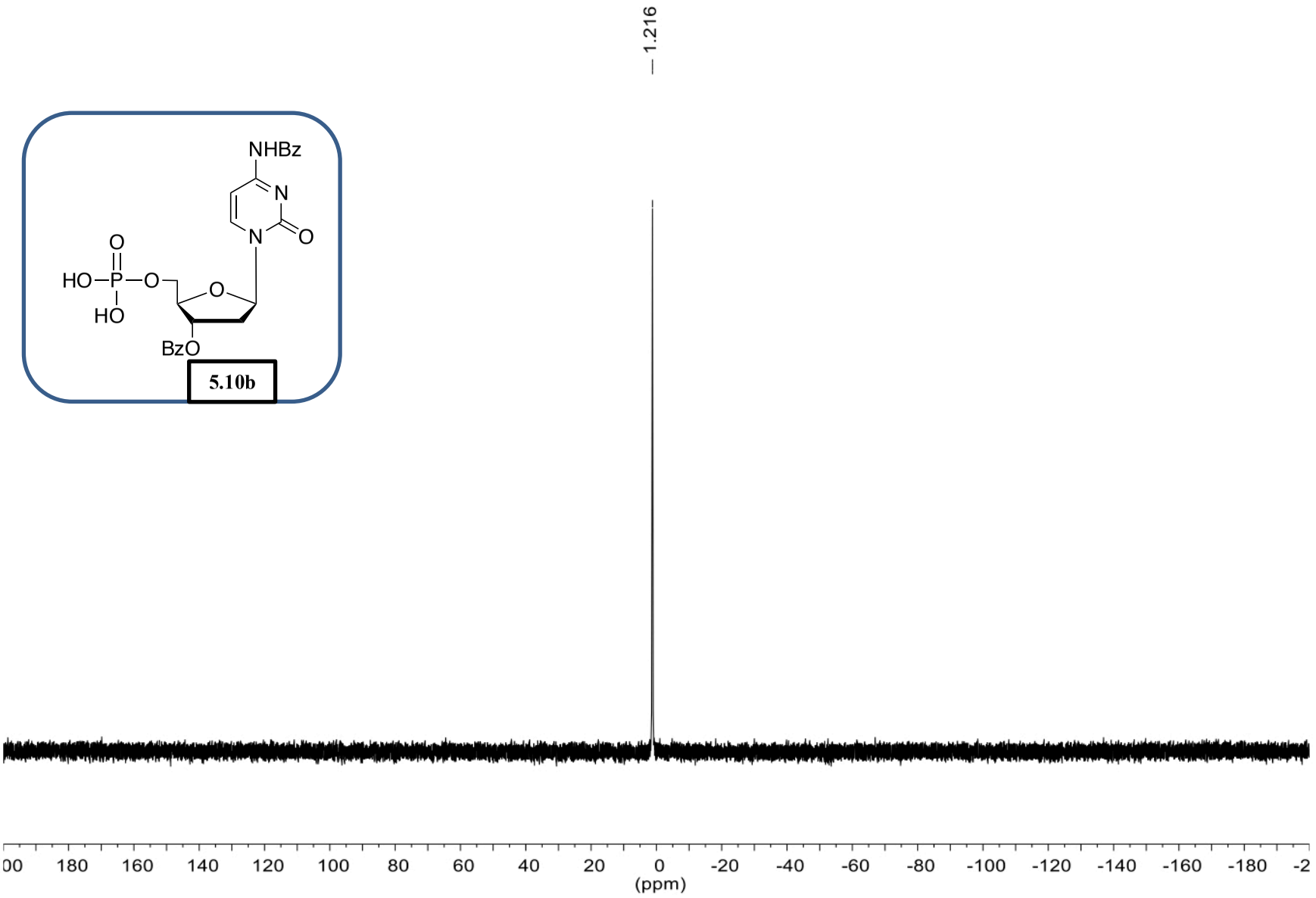


341

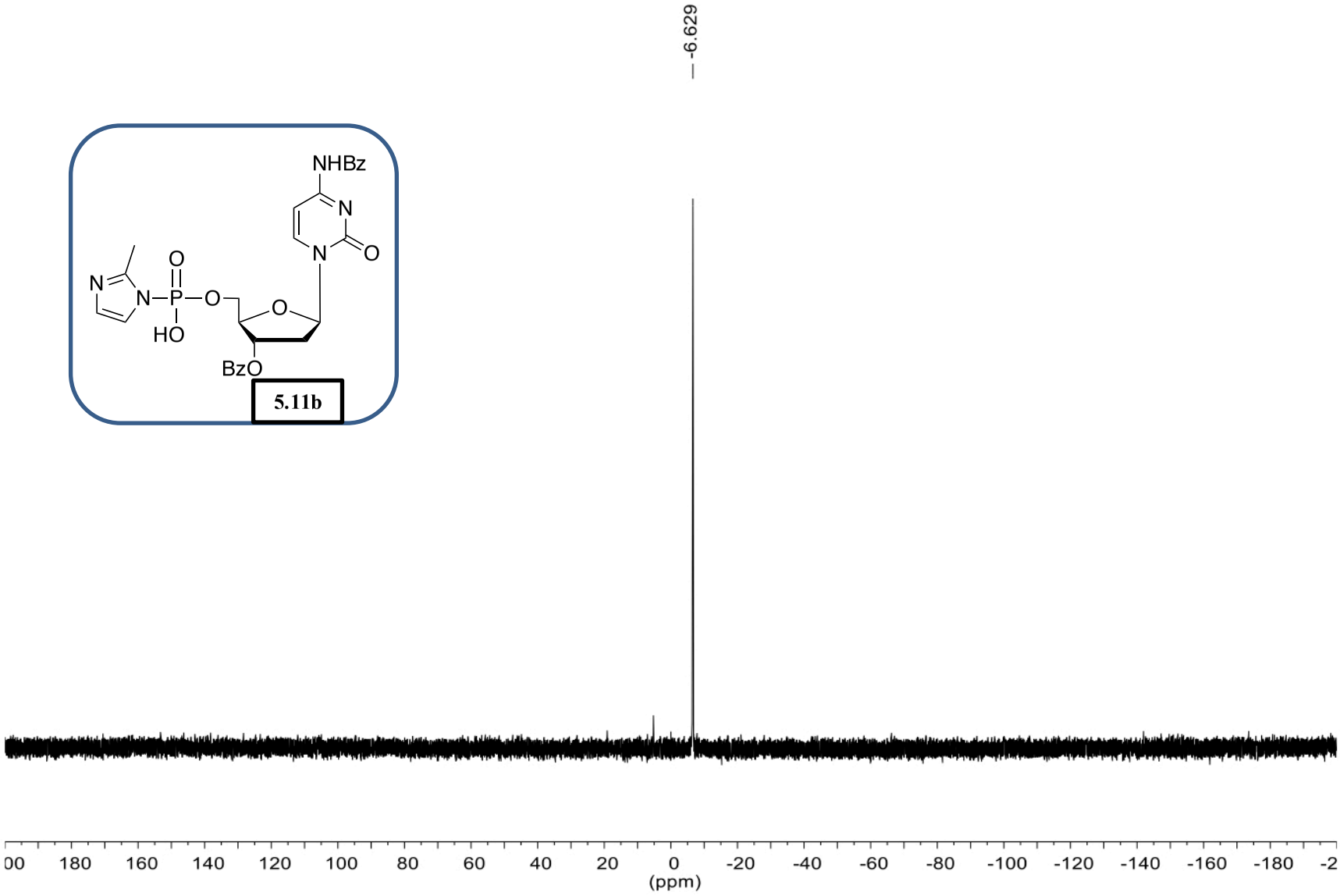
^{1}H NMR spectrum of compound **5.10b** (500 MHz,CD_3OD)



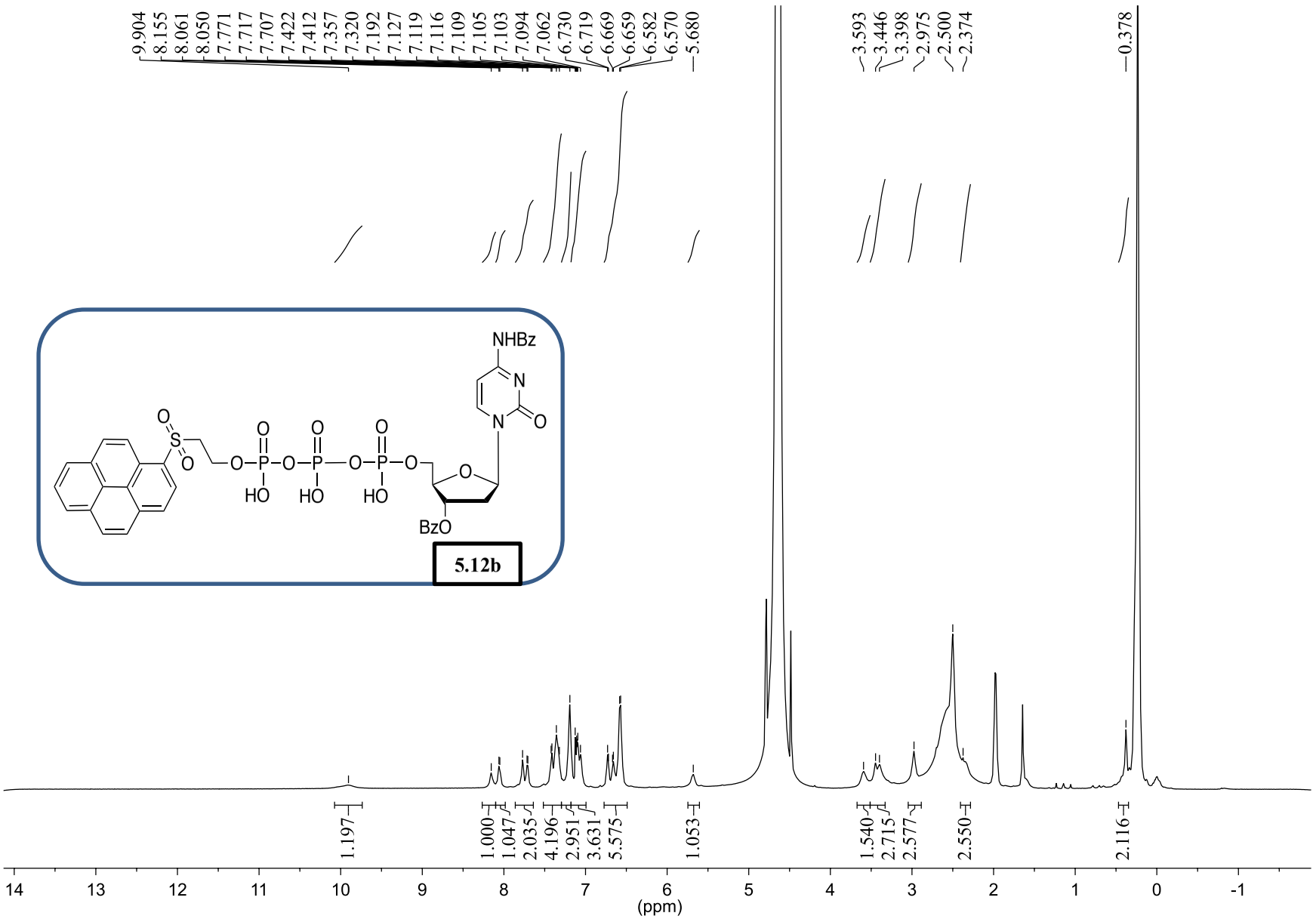
^{13}C NMR spectrum of compound **5.10b** (500 MHz,CD_3OD)



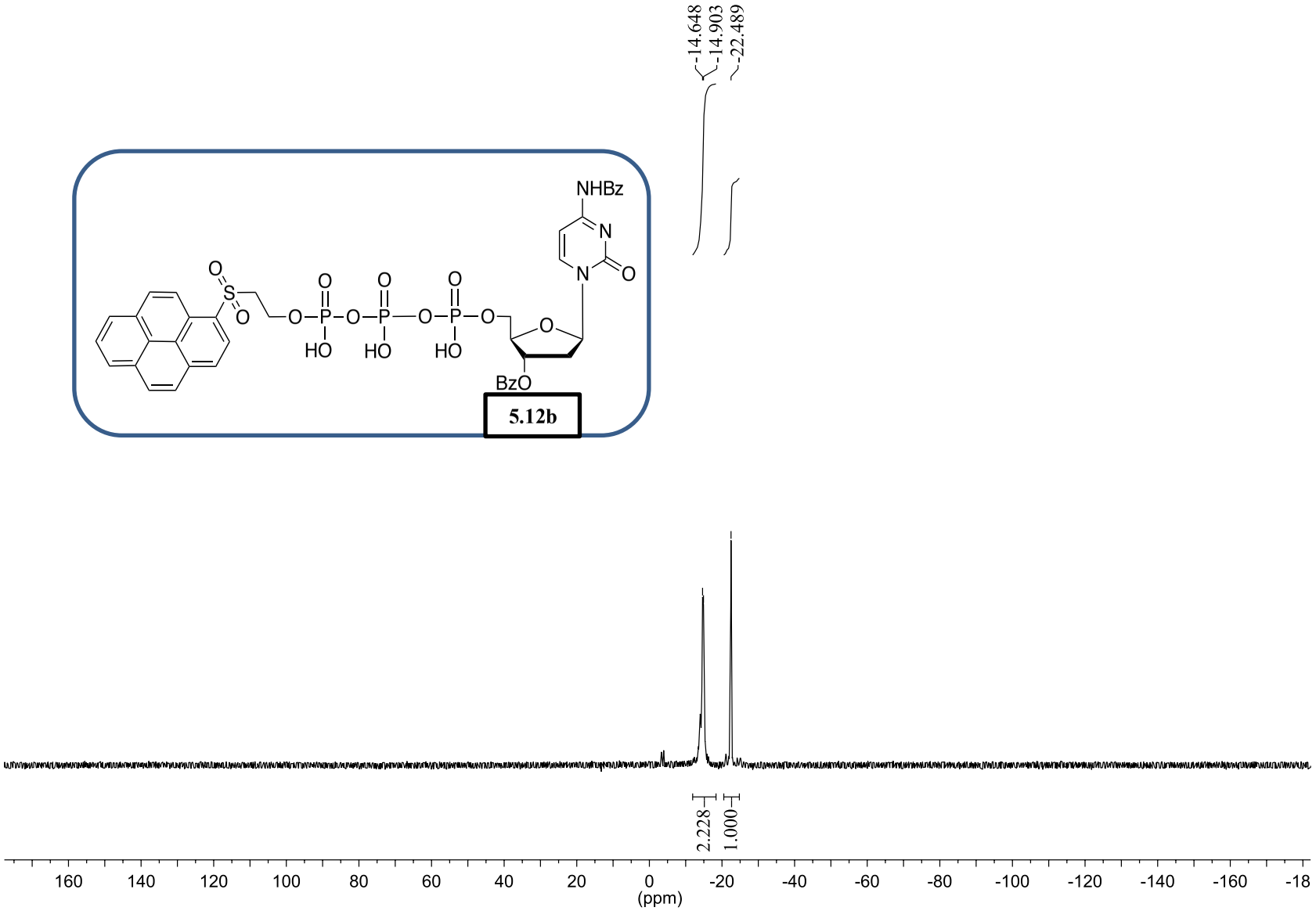
^{31}P NMR spectrum of compound **5.10b** (500 MHz,CD_3OD)



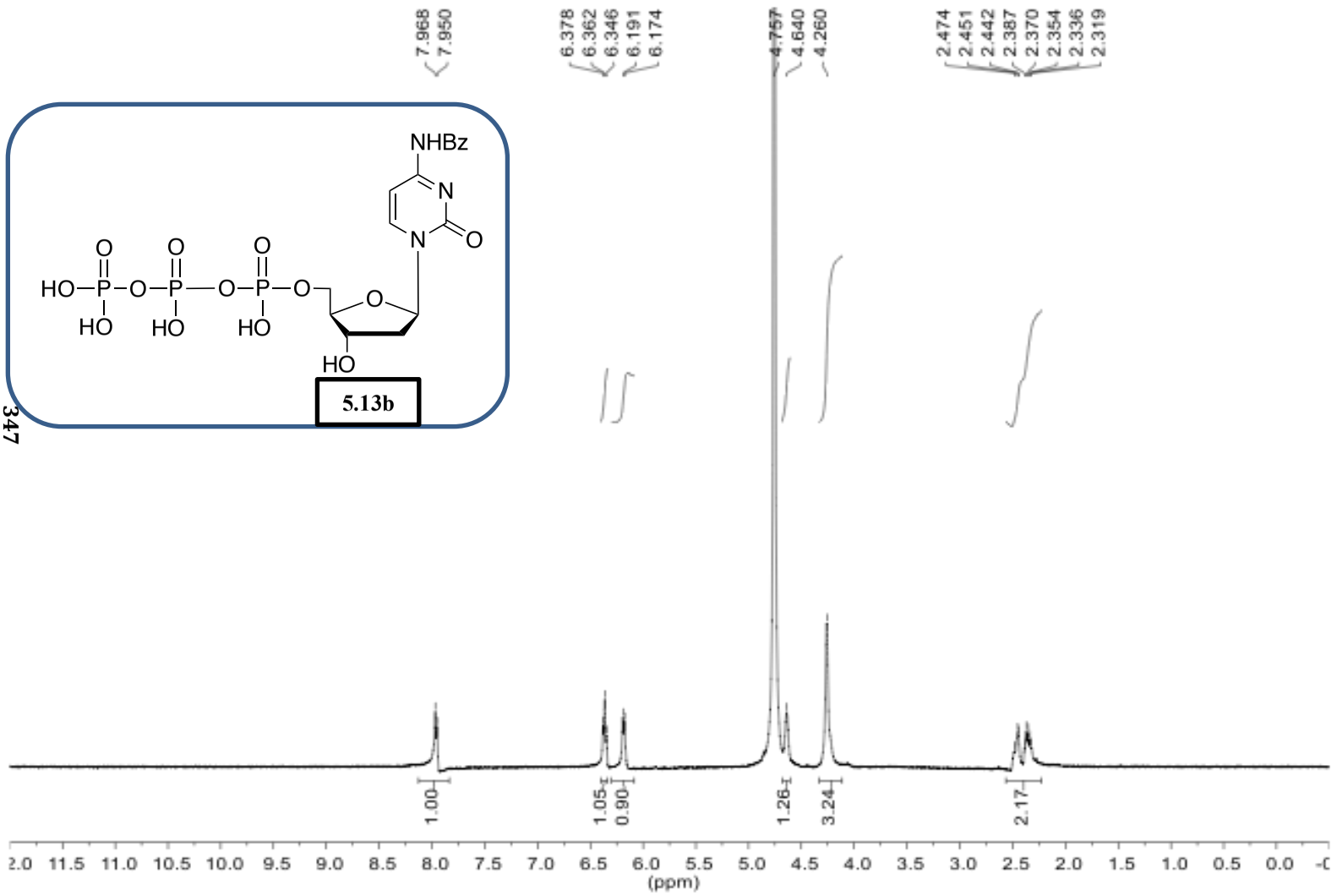
^{31}P NMR spectrum of compound **5.11b** (500 MHz,CD_3OD)



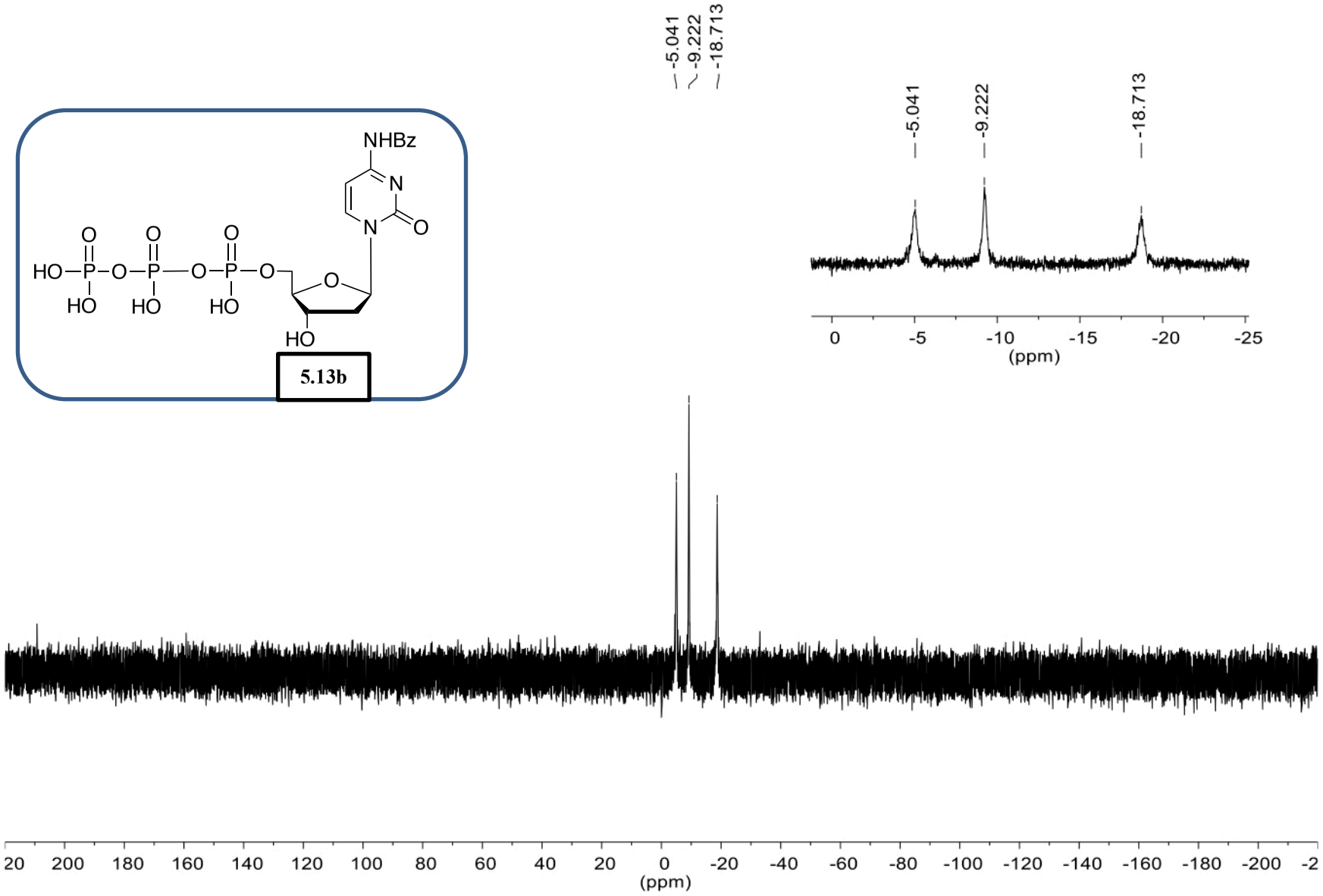
^{1}H NMR spectrum of compound **5.12b** (500 MHz,D_2O)



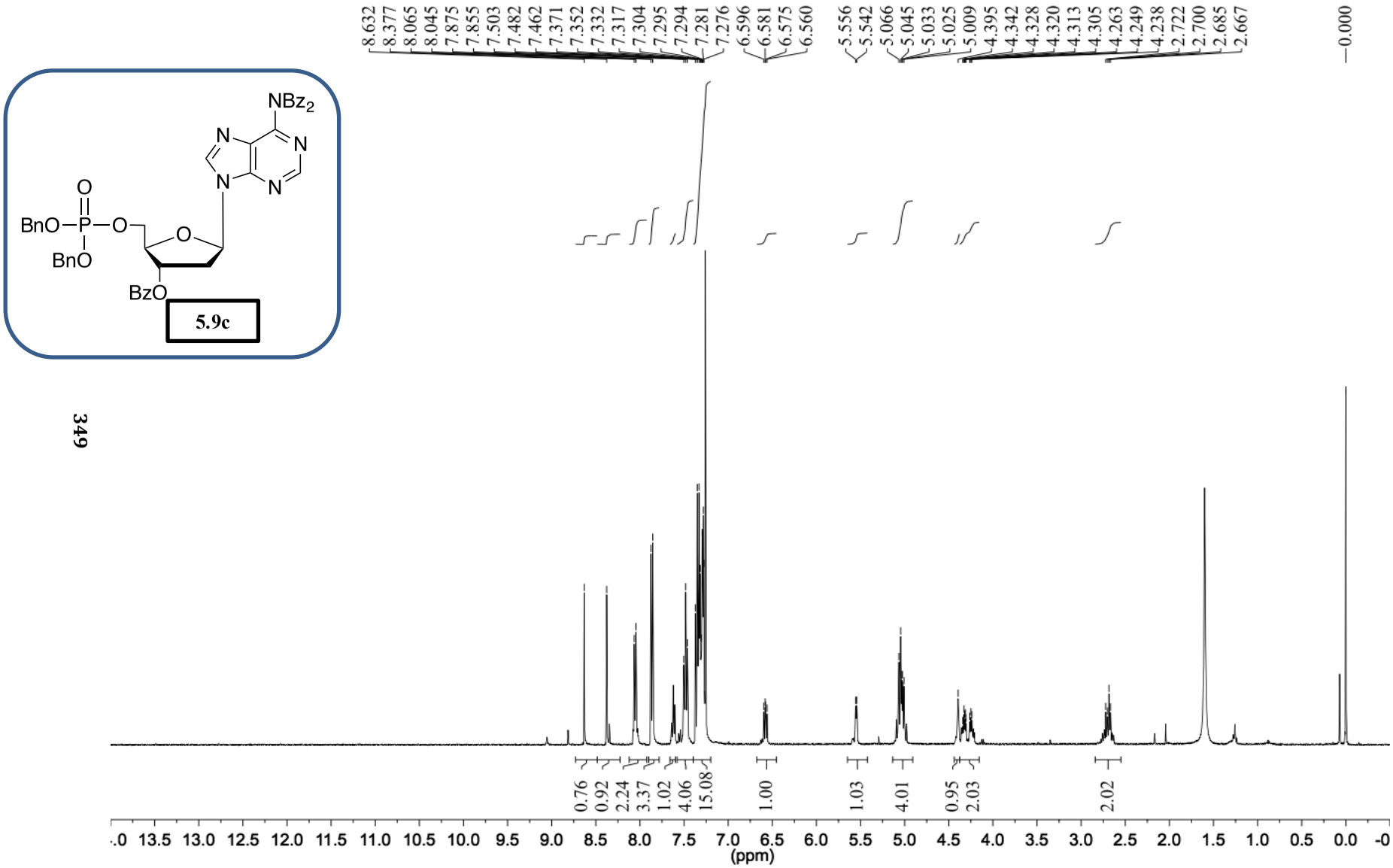
^{31}P NMR spectrum of compound **5.12b** (500 MHz,D_2O)



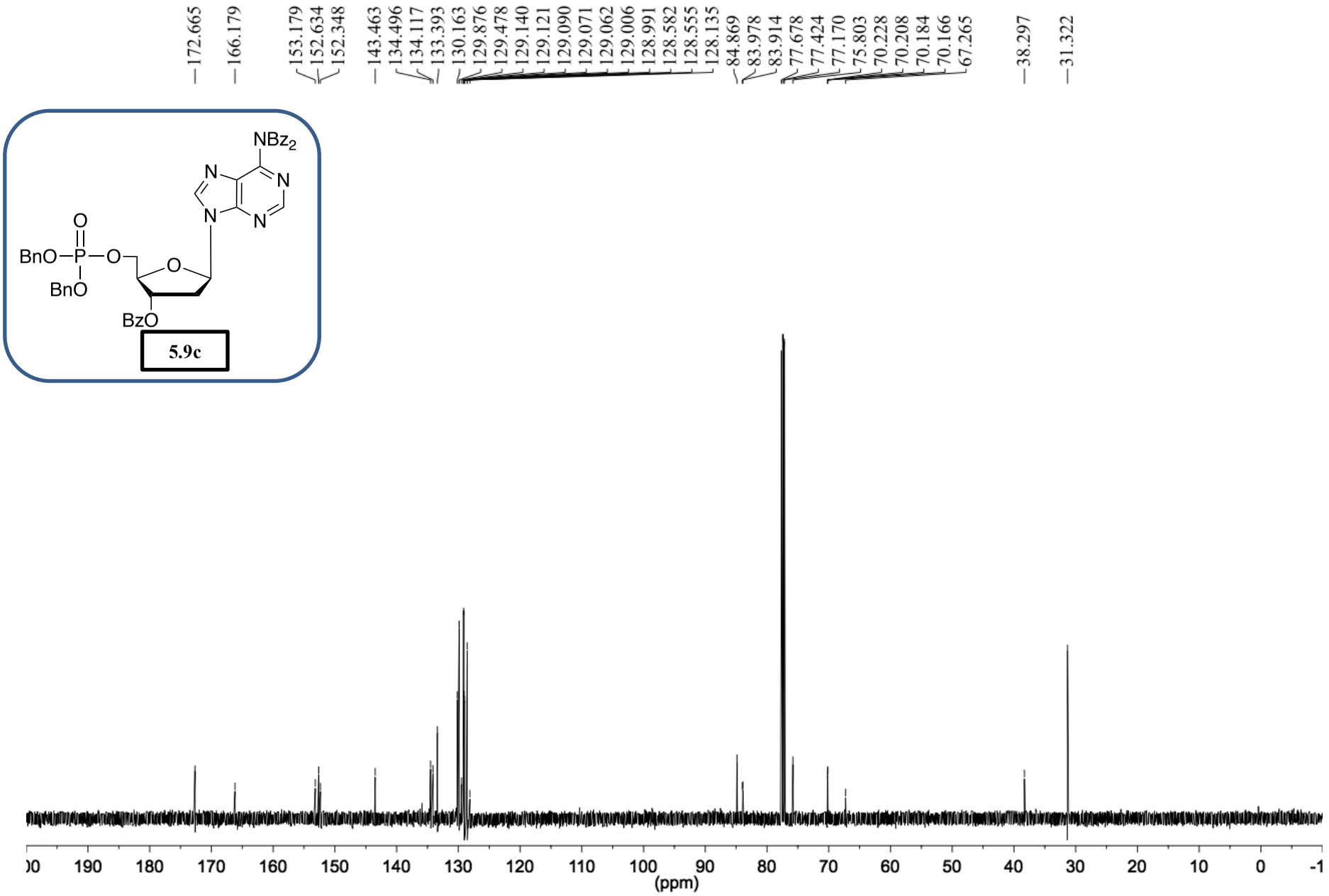
^{1}H NMR spectrum of compound **5.13b** (500 MHz,D_2O)



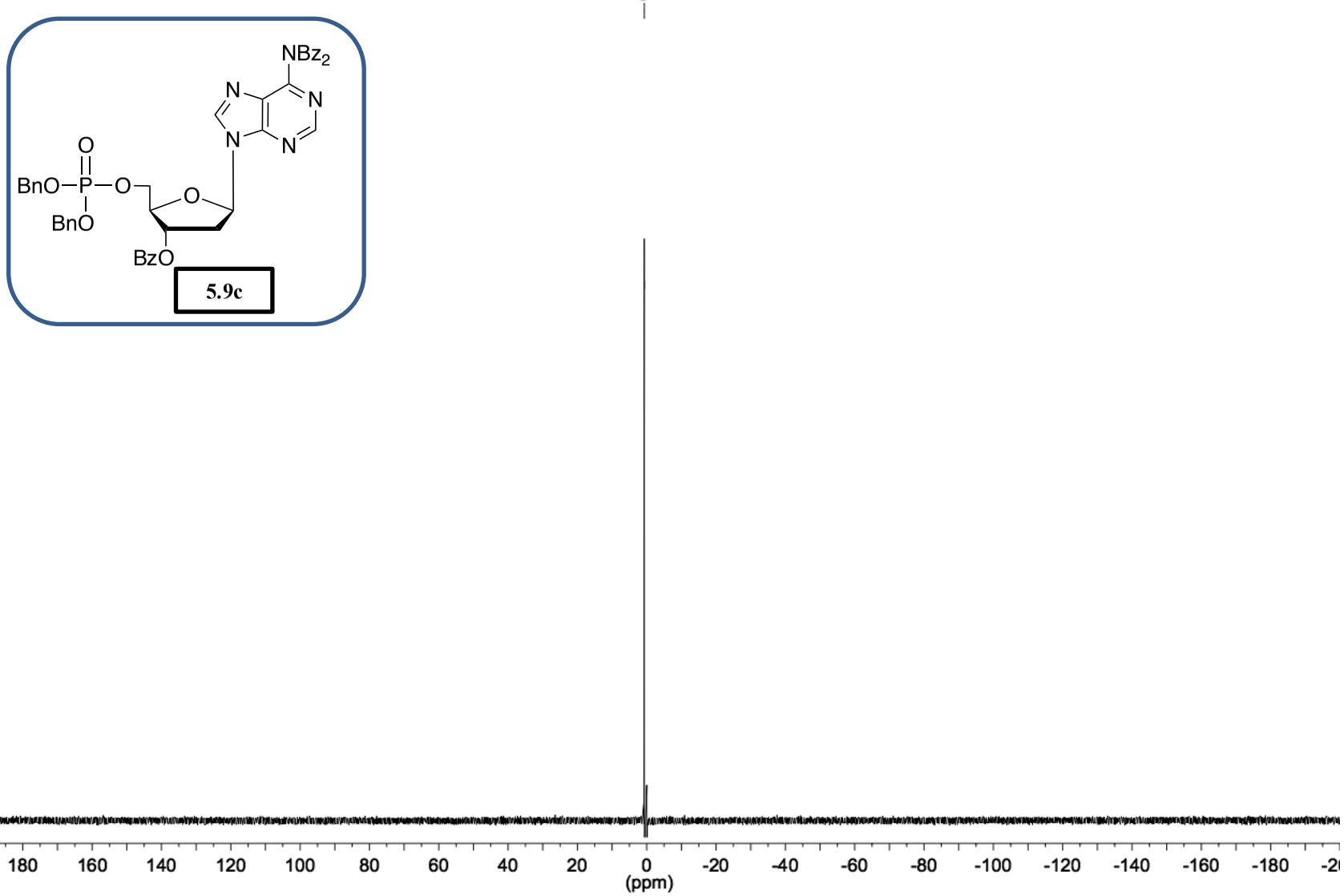
^{31}P NMR spectrum of compound **5.13b** (500 MHz,D_2O)



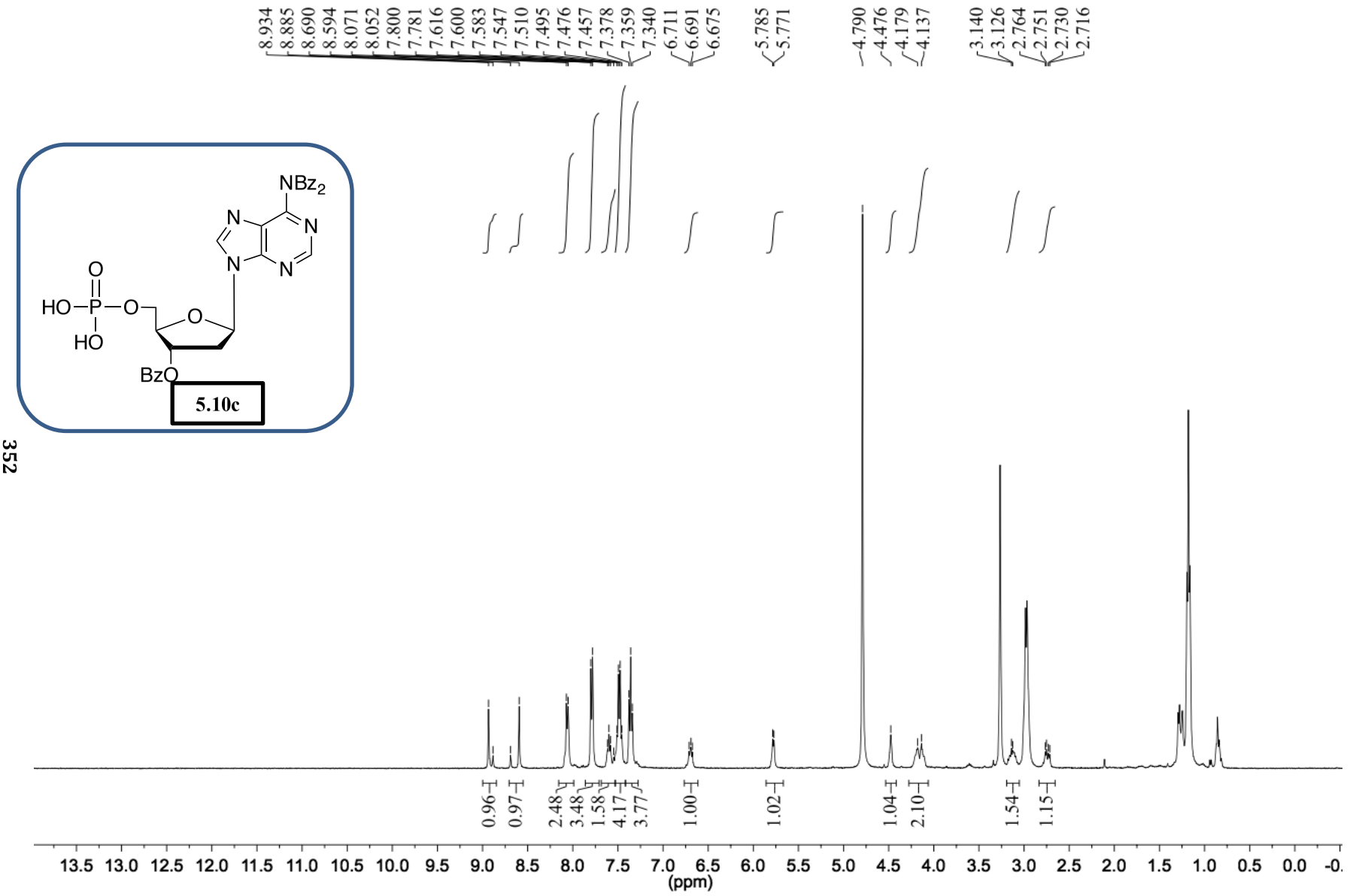
^{1}H NMR spectrum of compound **5.9c** (500 MHz,$CDCl_3$)



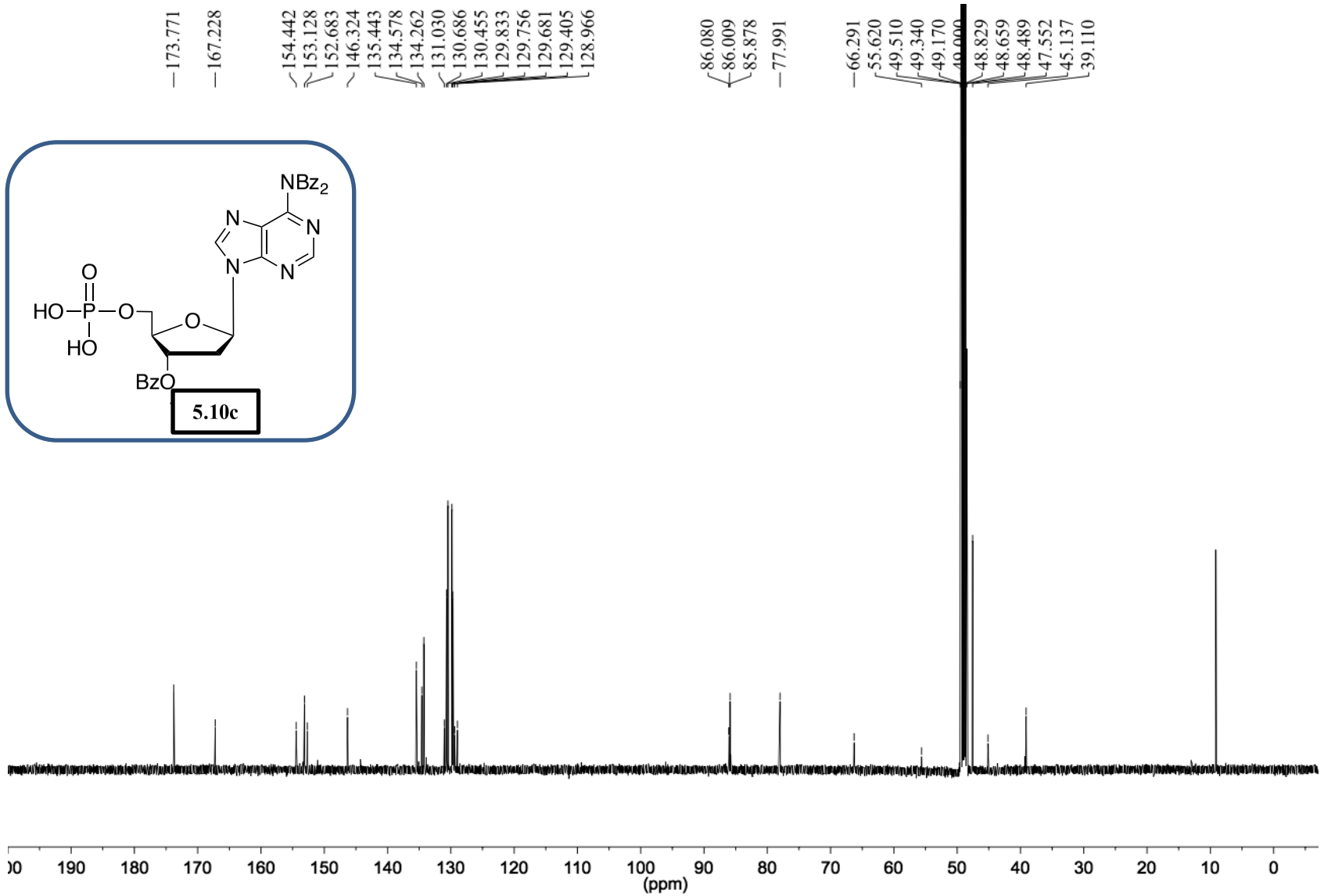
^{13}C NMR spectrum of compound **4.9c** (500 MHz,$CDCl_3$)



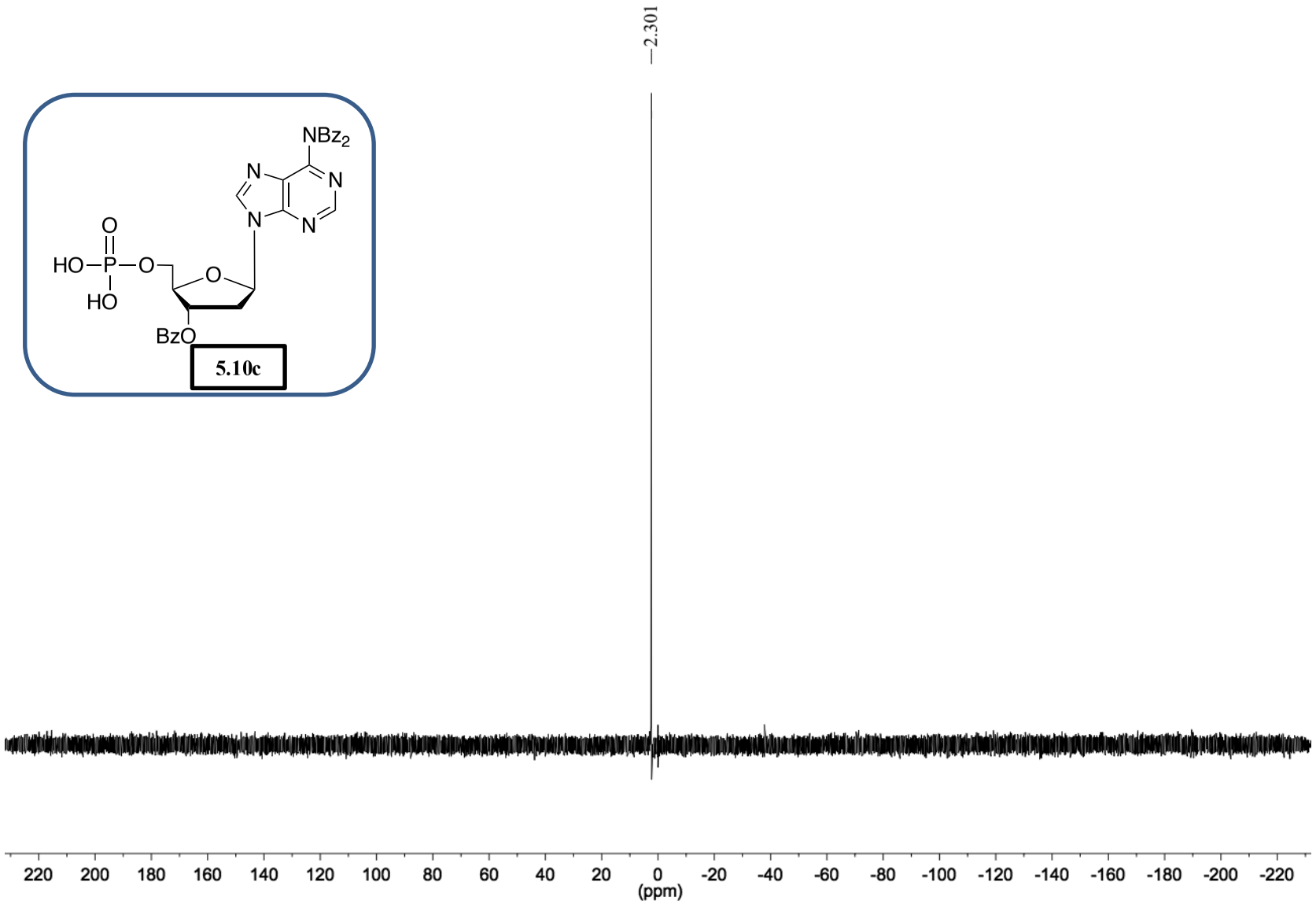
^{31}P NMR spectrum of compound **5.9c** (500 MHz,$CDCl_3$)



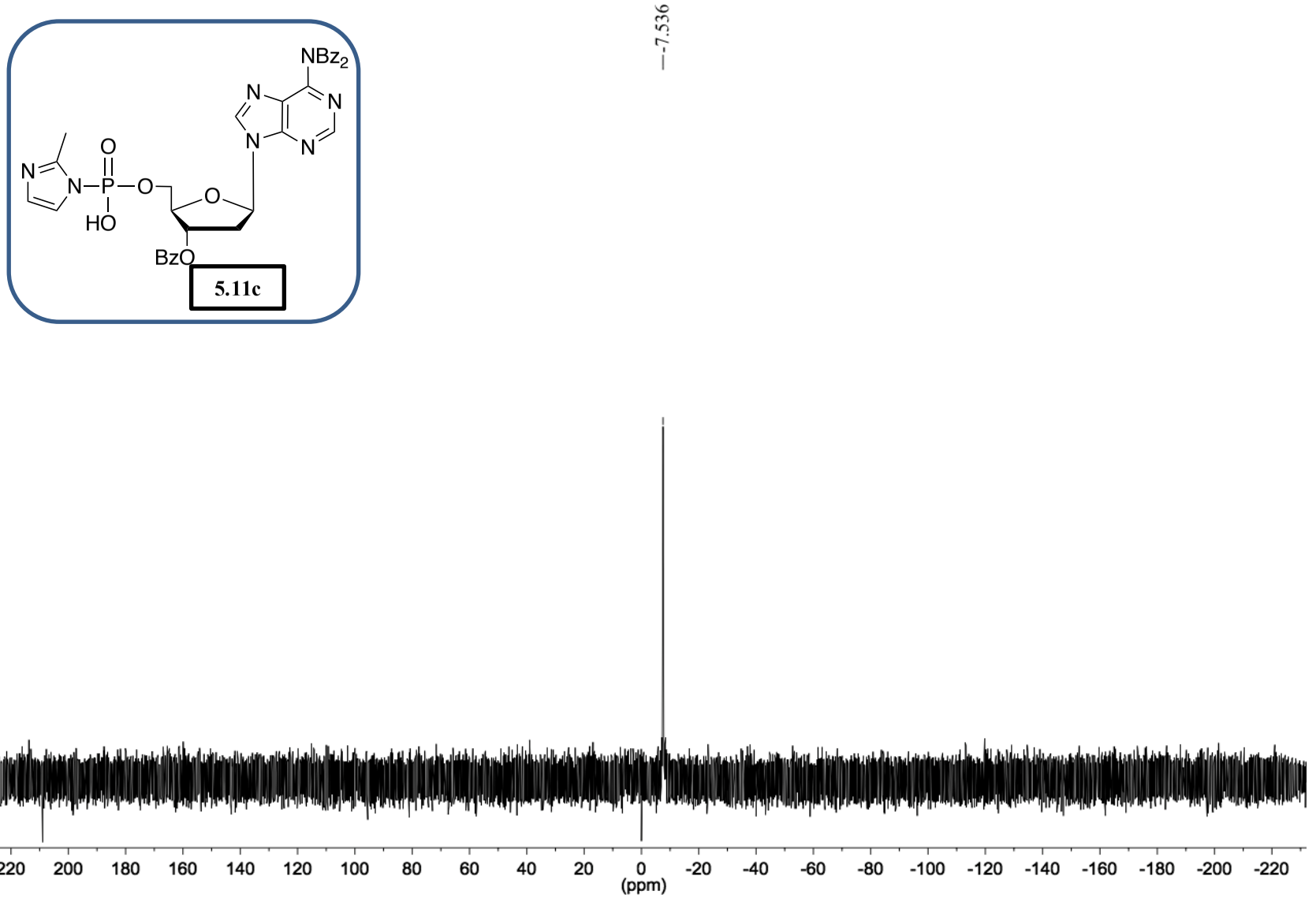
^{1}H NMR spectrum of compound **5.10c** (500 MHz,CD_3OD)



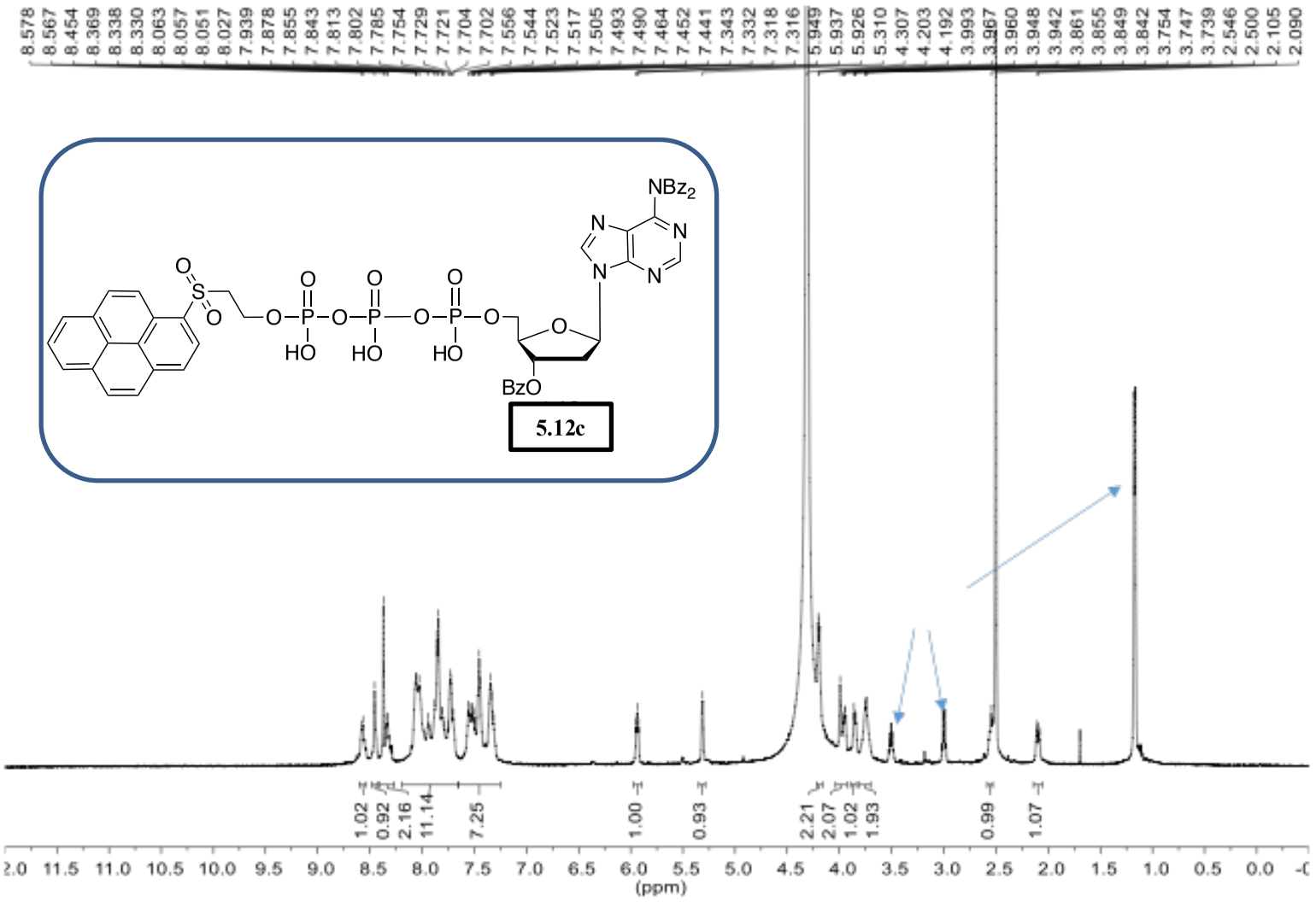
^{13}C NMR spectrum of compound **5.10c** (500 MHz,CD_3OD)



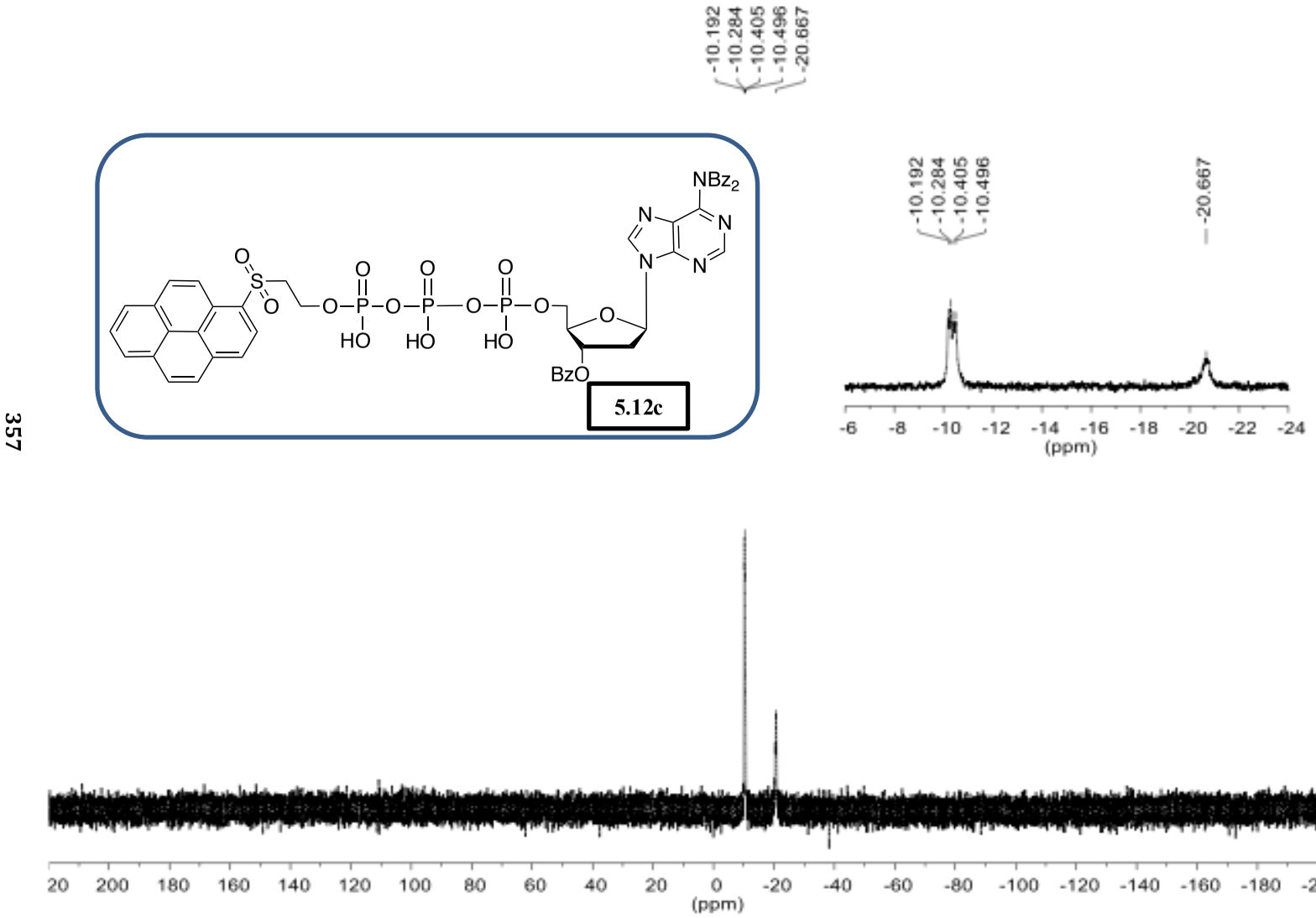
^{31}P NMR spectrum of compound **5.10c** (500 MHz,CD_3OD)



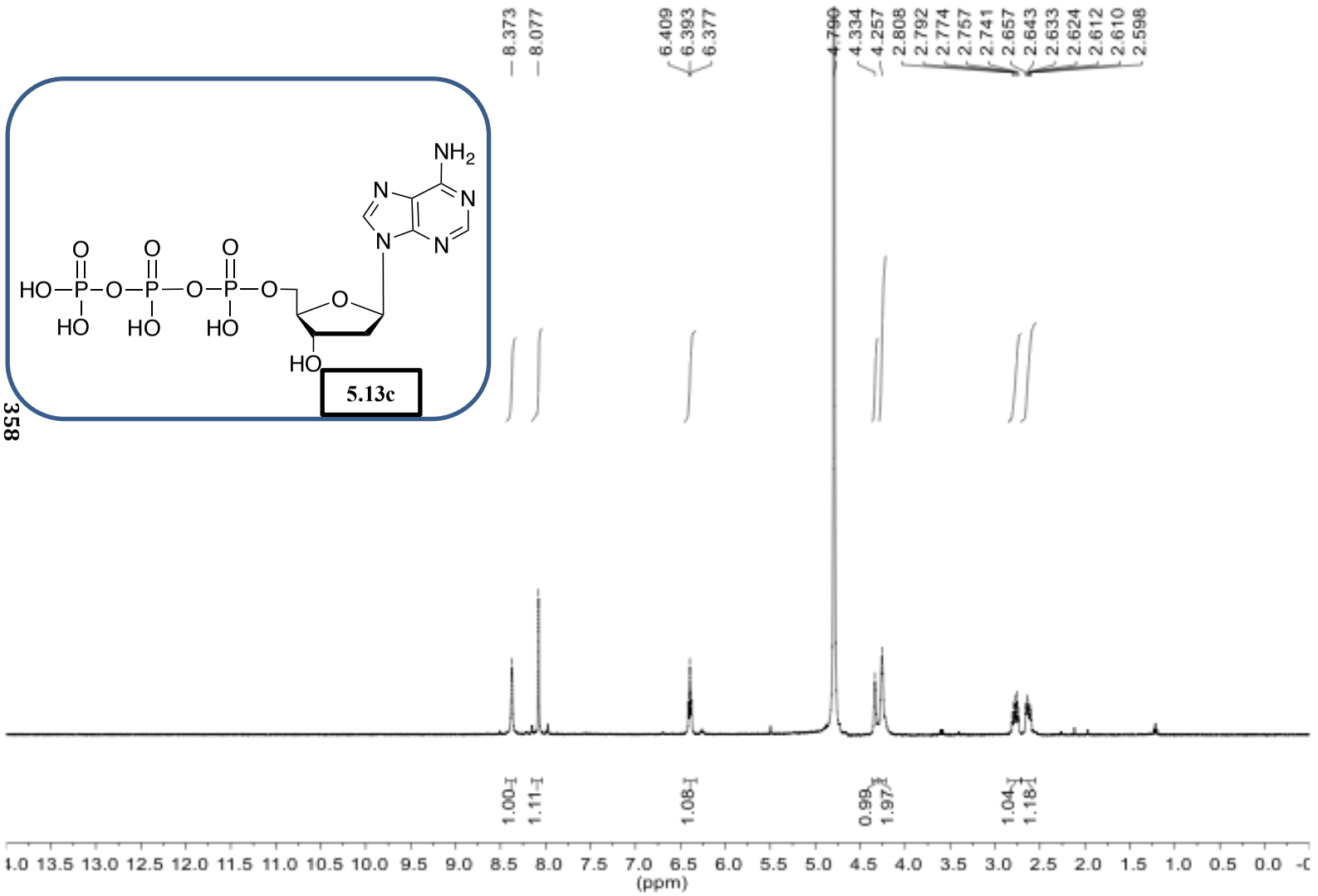
^{31}P NMR spectrum of compound **5.11c** (500 MHz,CD_3OD)



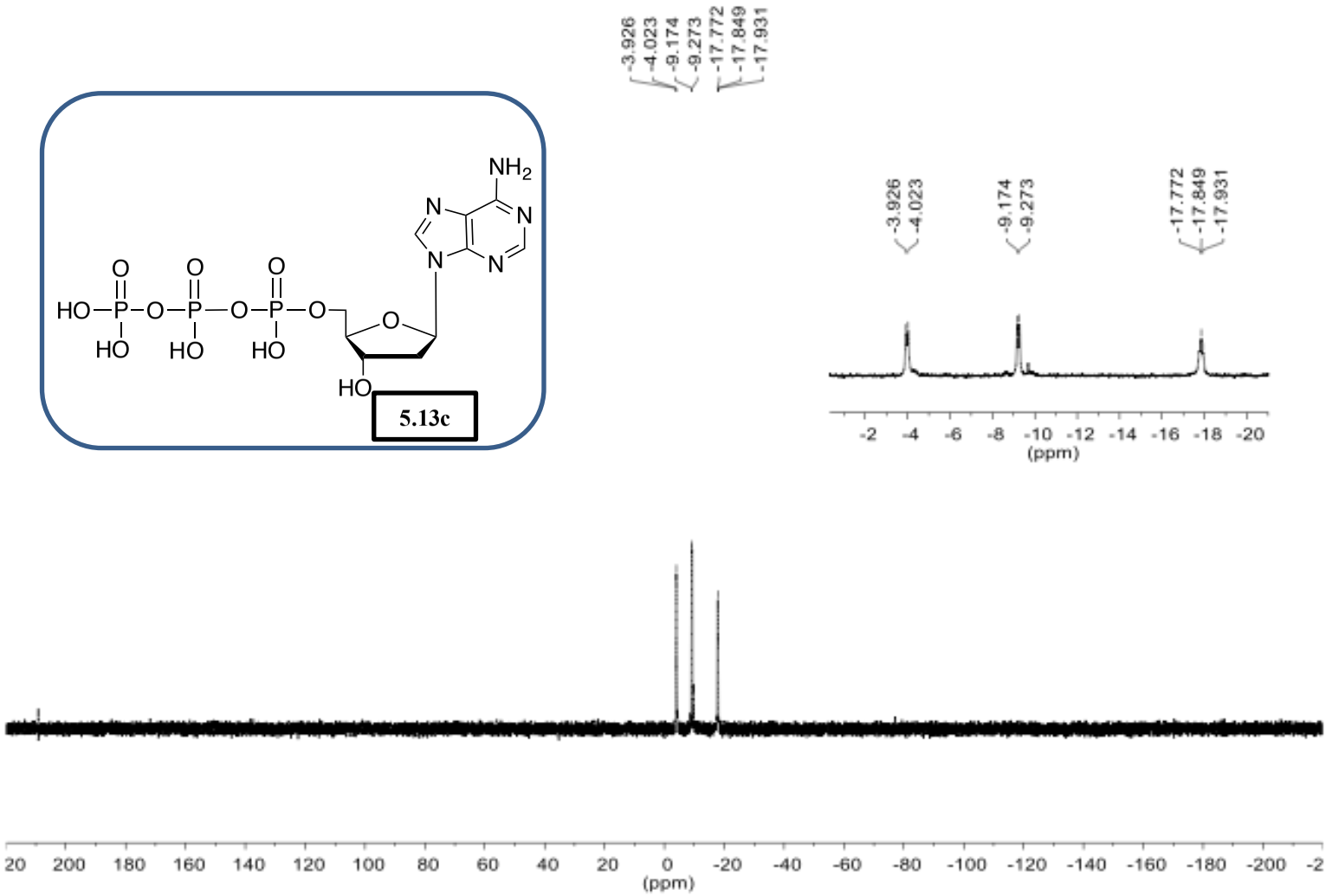
^{1}H NMR spectrum of compound **5.12c** (500 MHz,D_2O)



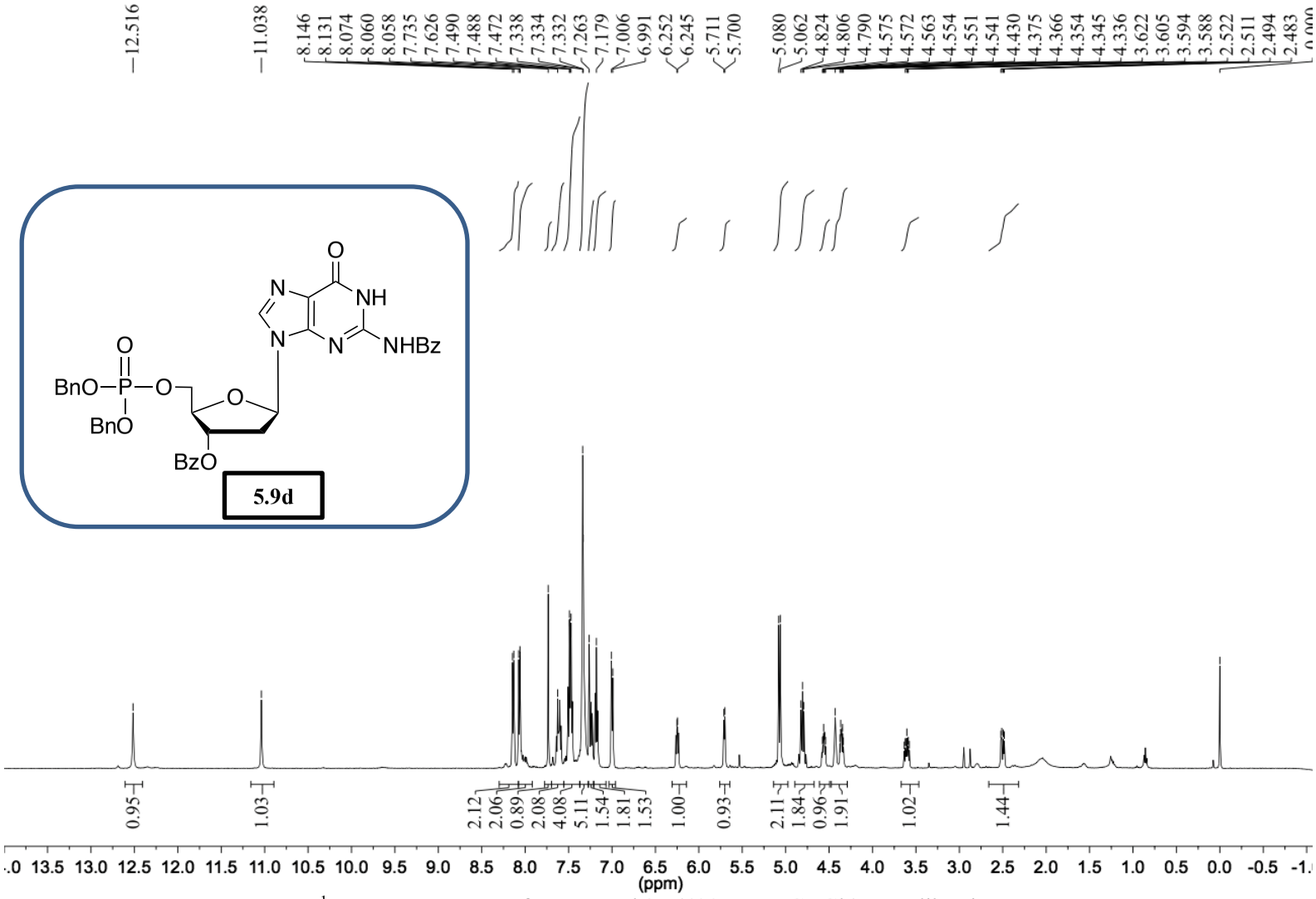
^{31}P NMR spectrum of compound **5.12c** (500 MHz,D_2O)



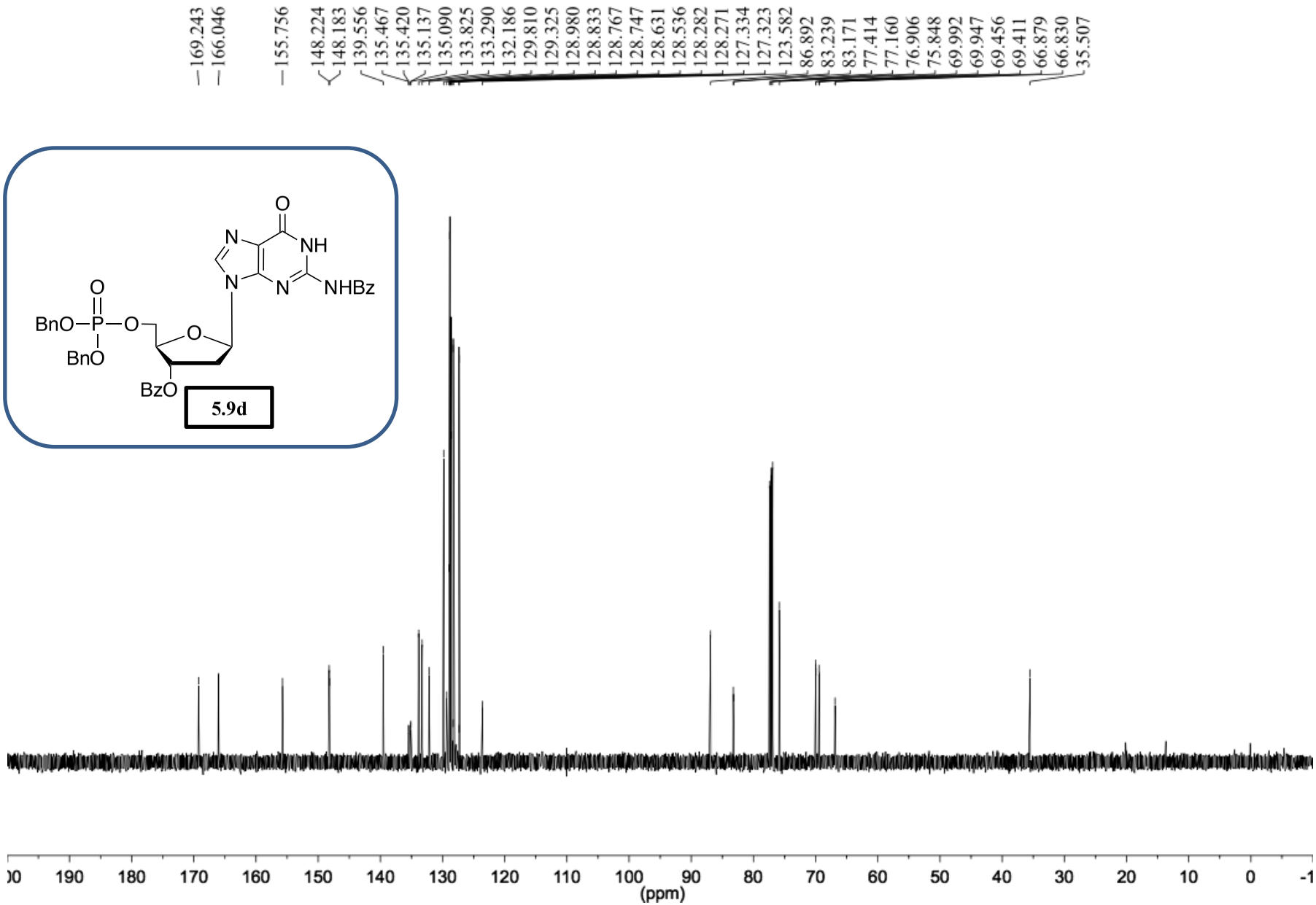
^{1}H NMR spectrum of compound **5.13c** (500 MHz,D_2O)



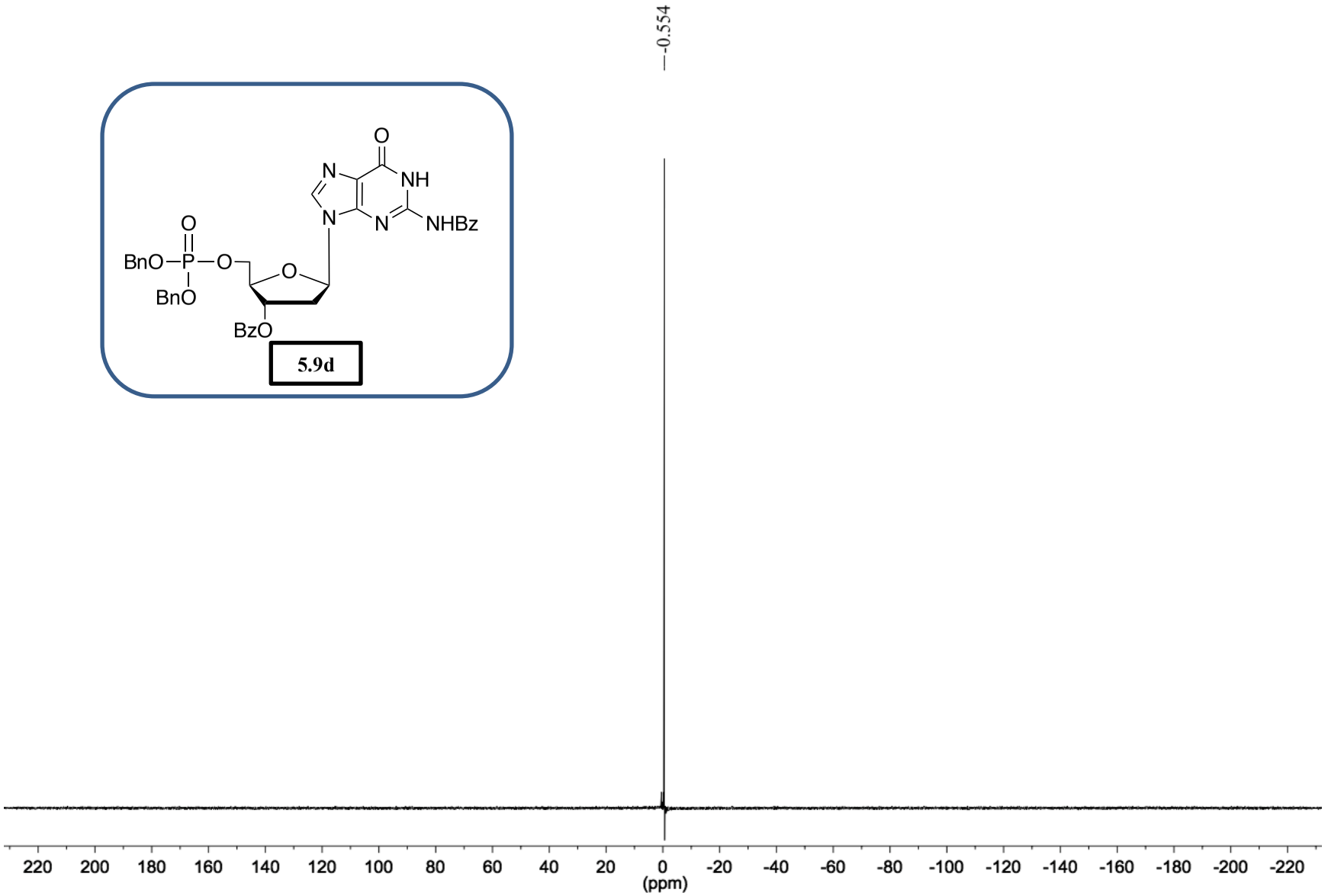
^{31}P NMR spectrum of compound **5.13c** (500 MHz,D_2O)



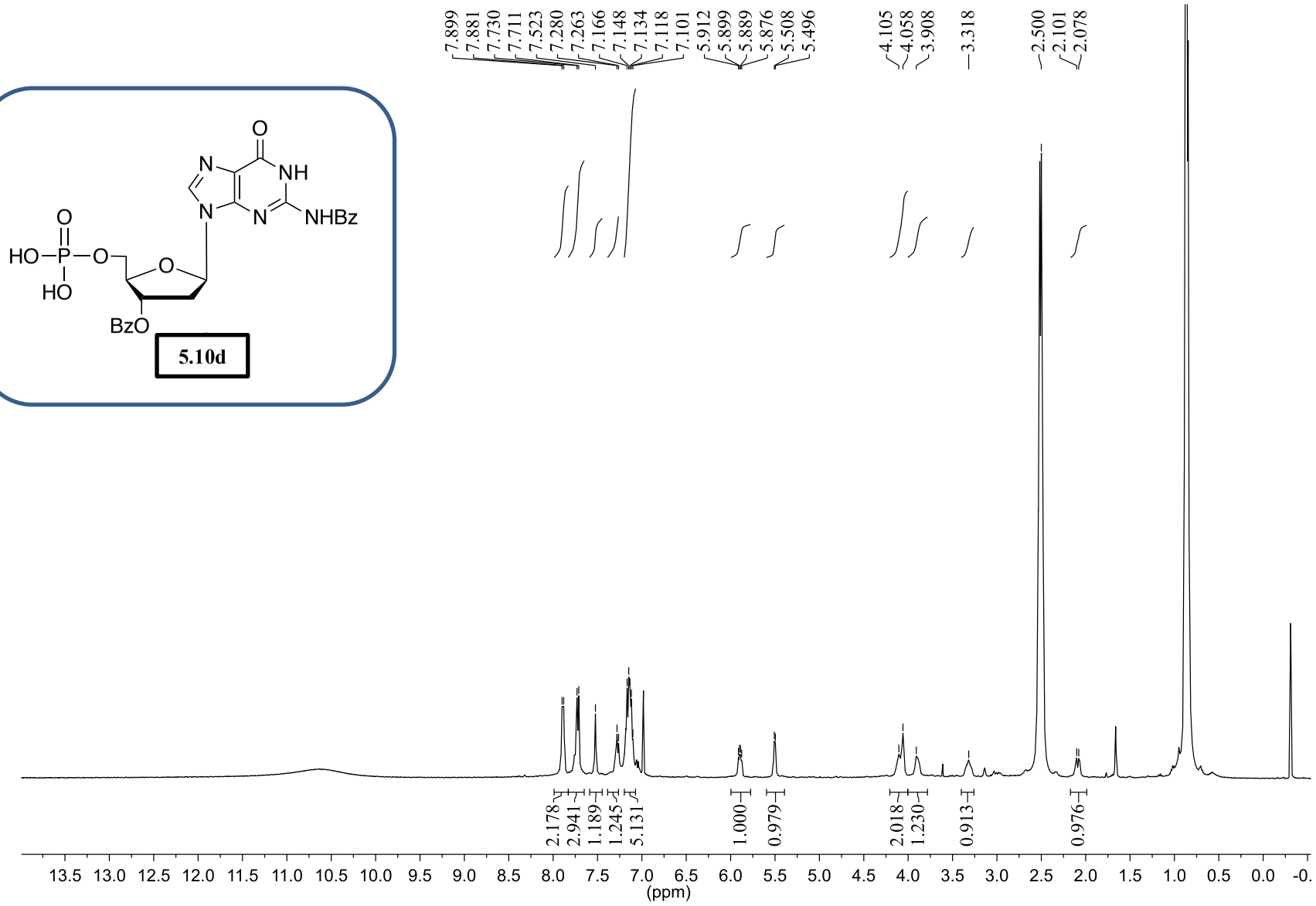
^{1}H NMR spectrum of compound **9d** (400 MHz, $CDCl_3$) no calibration



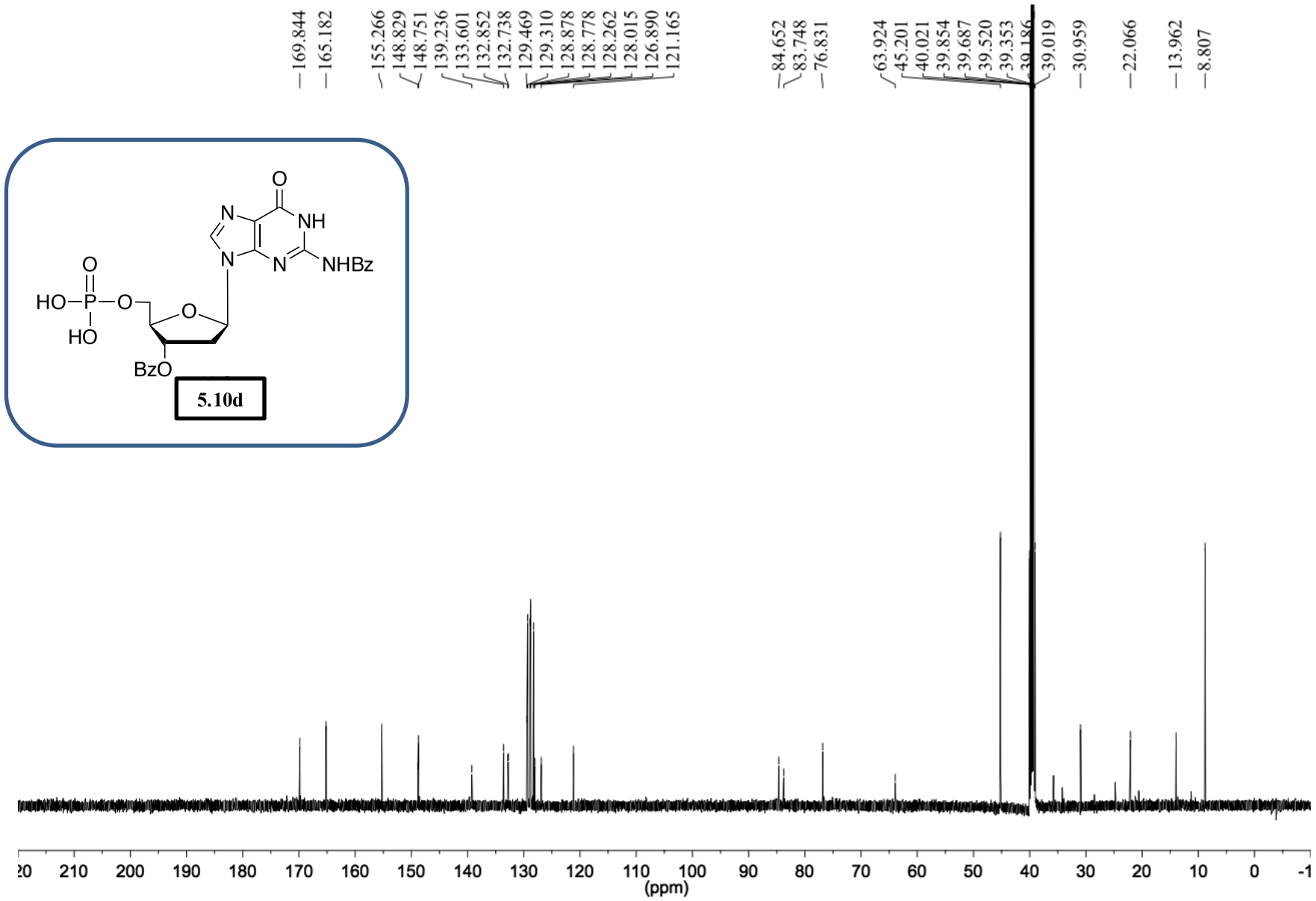
^{1}H NMR spectrum of compound **5.9d** (500 MHz,$CDCl_3$)



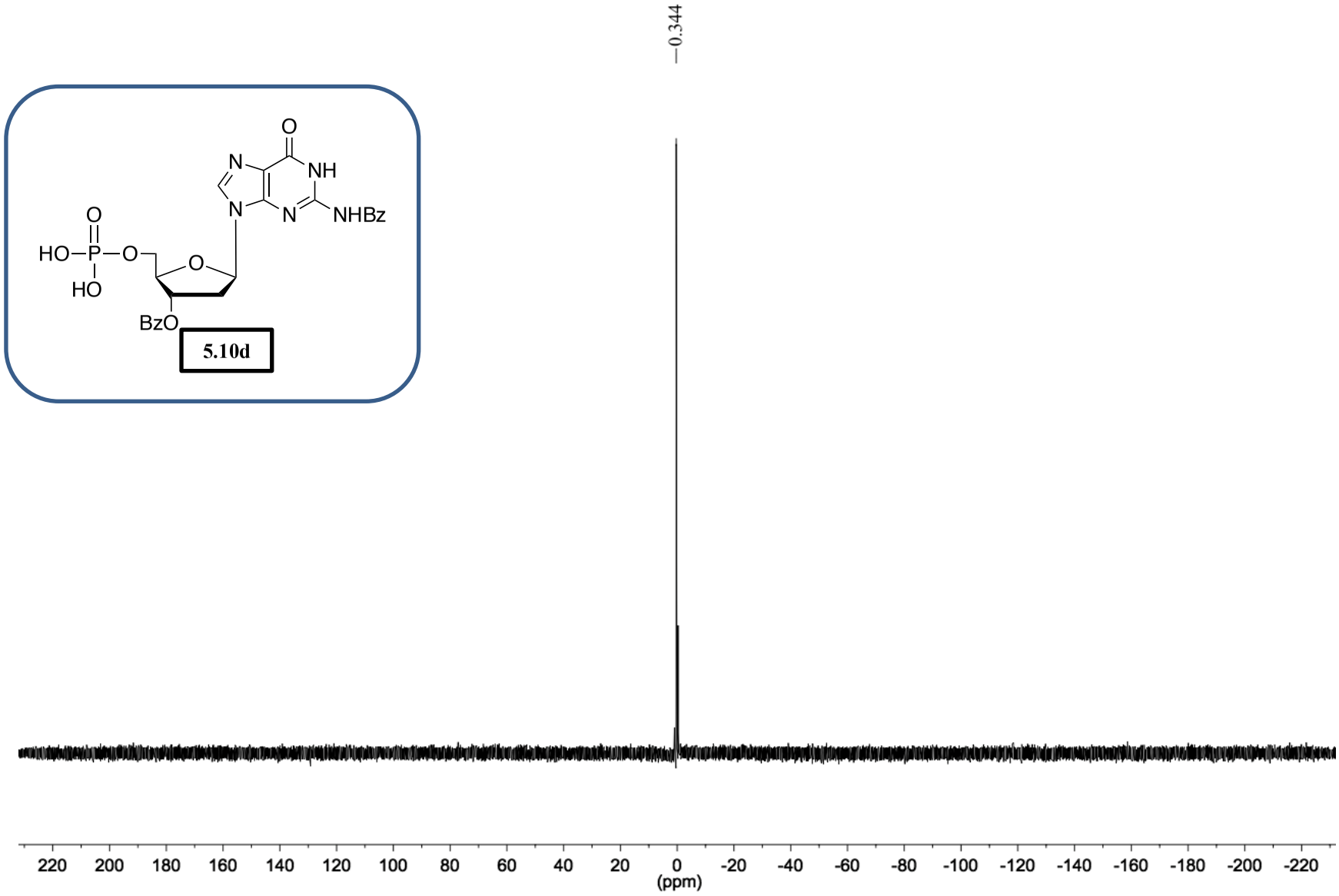
^{31}P NMR spectrum of compound **5.9d** (500 MHz,$CDCl_3$)



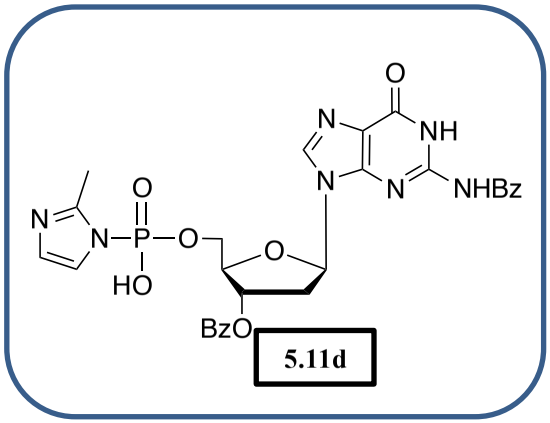
^{1}H NMR spectrum of compound **5.10d** (500 MHz,CD_3OD)



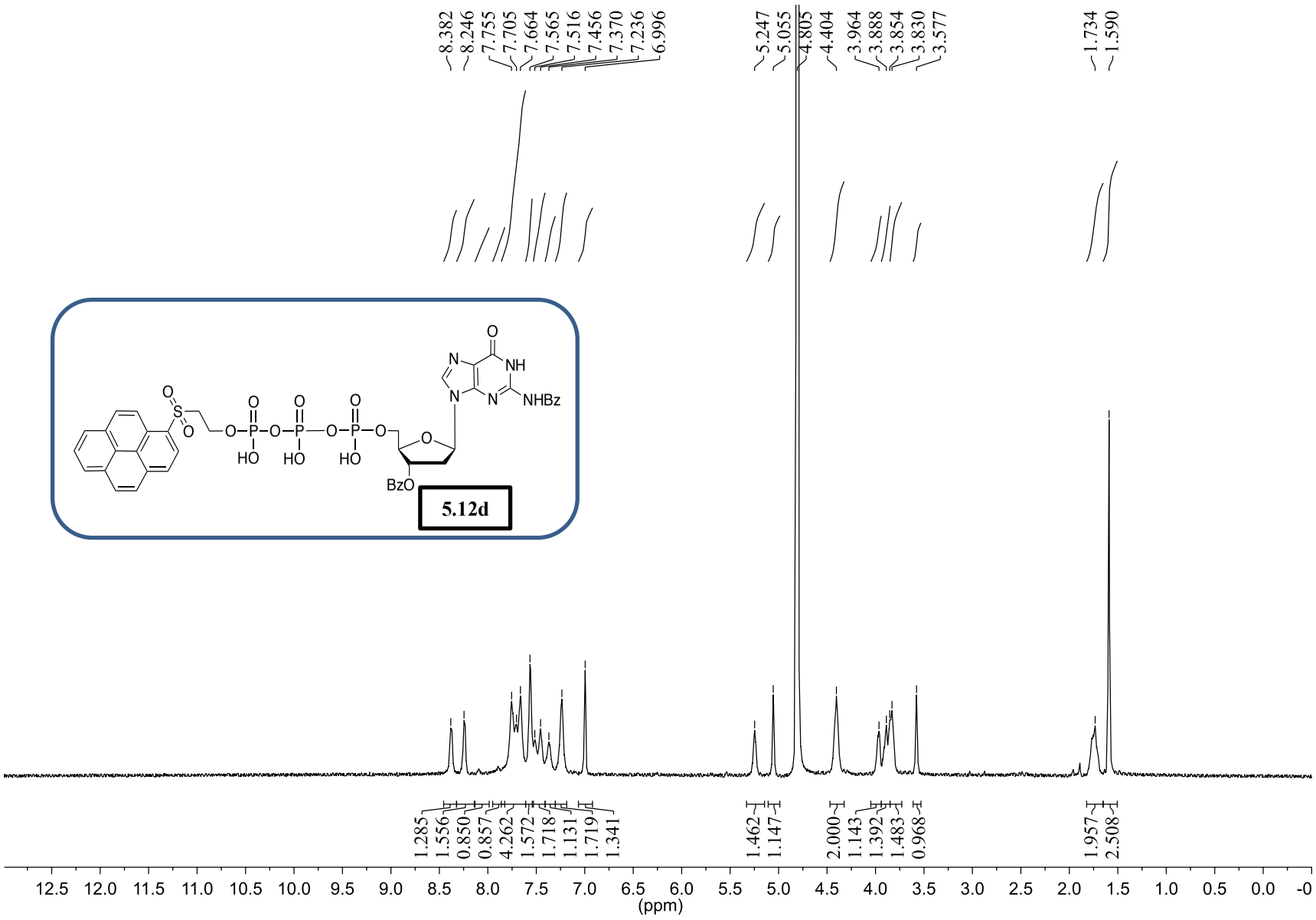
^{13}C NMR spectrum of compound **5.10d** (500 MHz,CD_3OD)



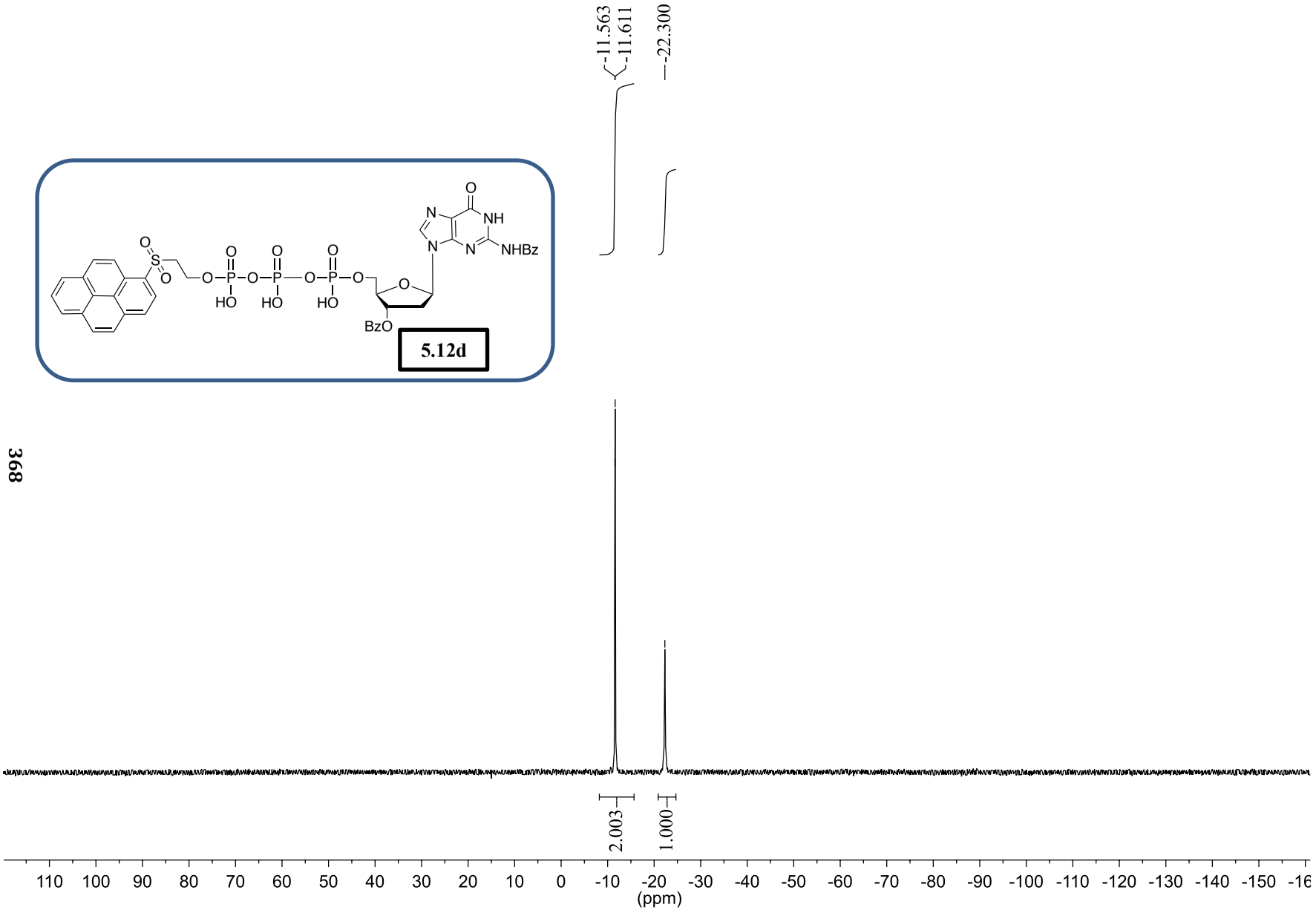
^{31}P NMR spectrum of compound **5.10d** (500 MHz,CD_3OD)



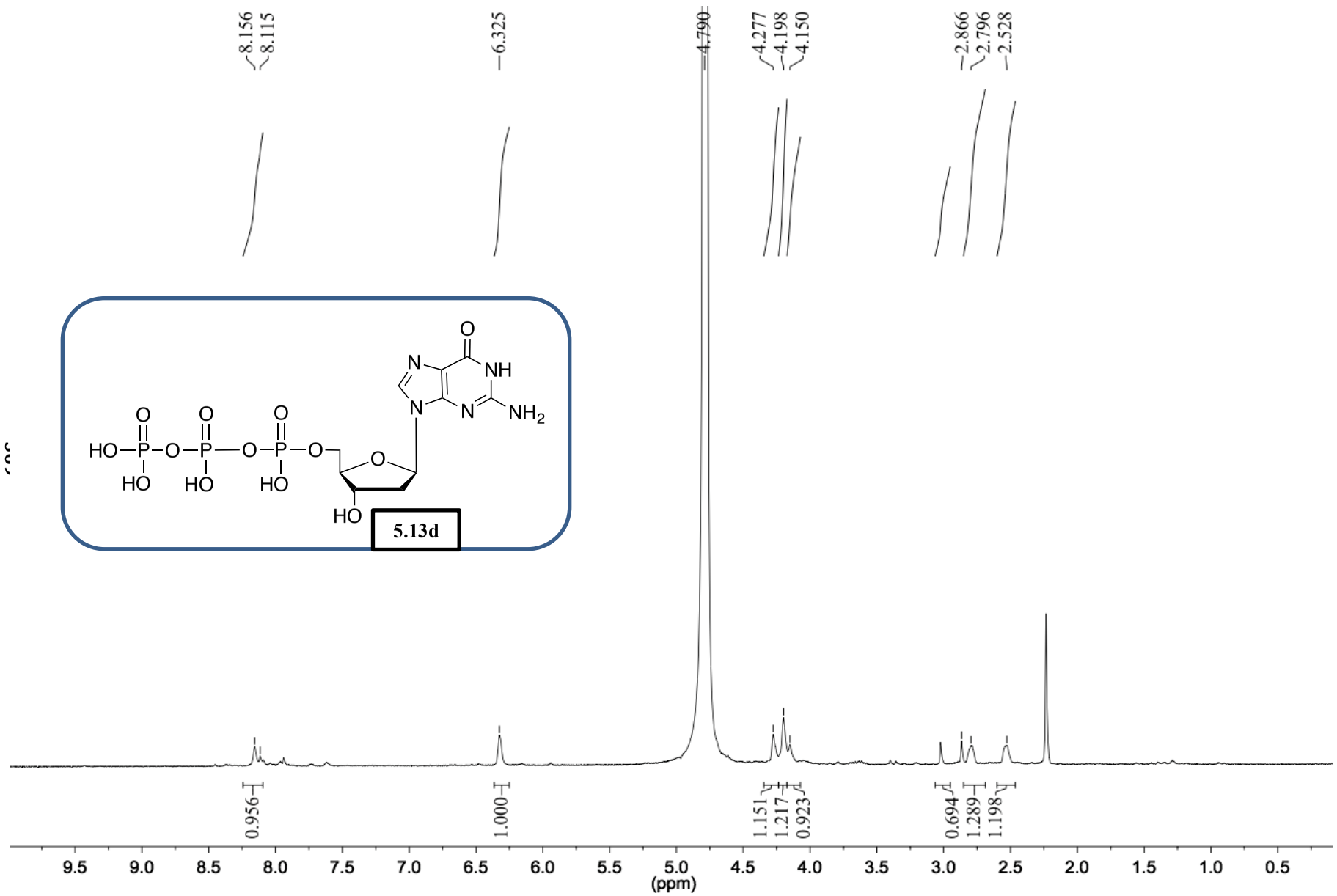
^{31}P NMR spectrum of compound **5.11d** (500 MHz,D_2O)



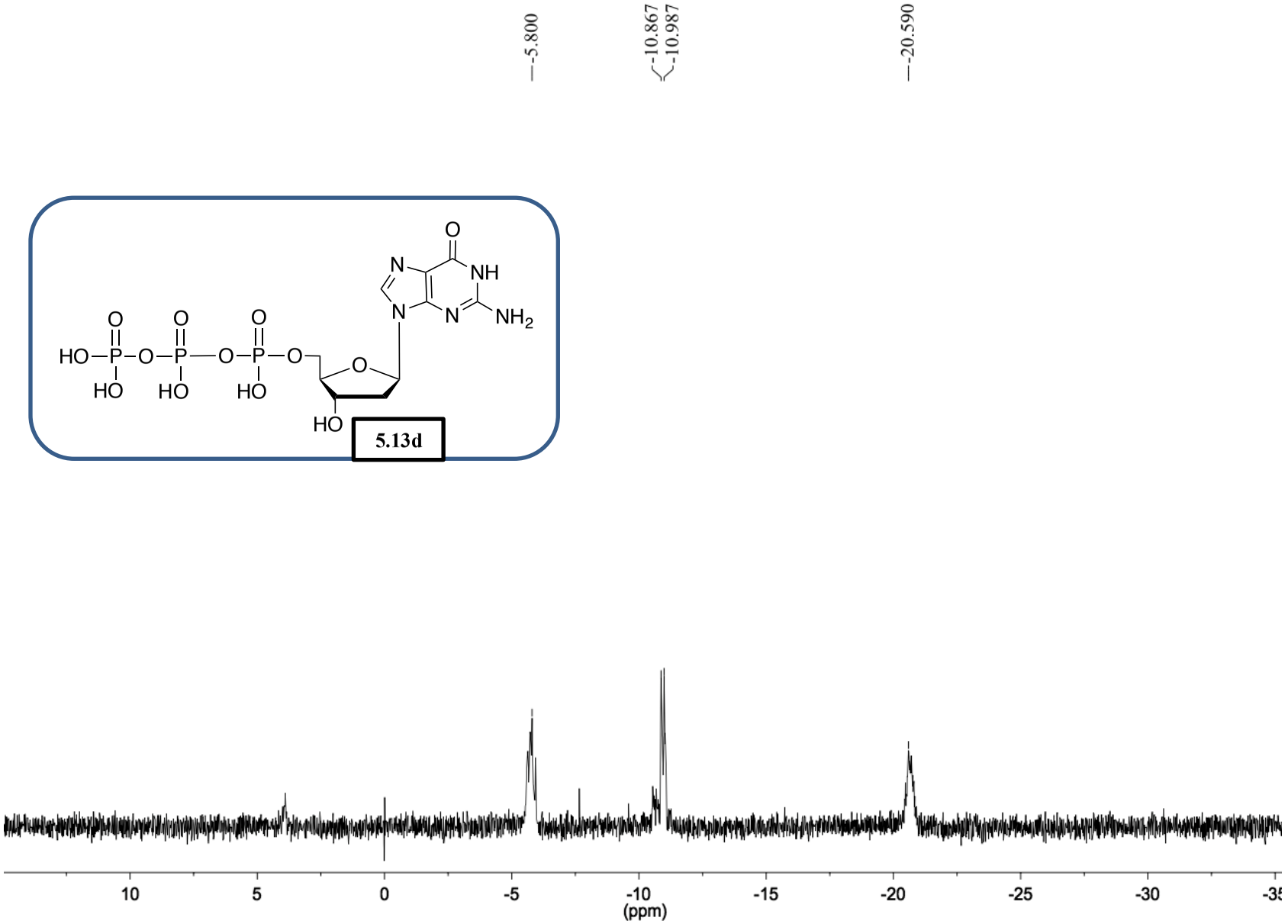
^{1}H NMR spectrum of compound **5.12d** (500 MHz,D_2O)



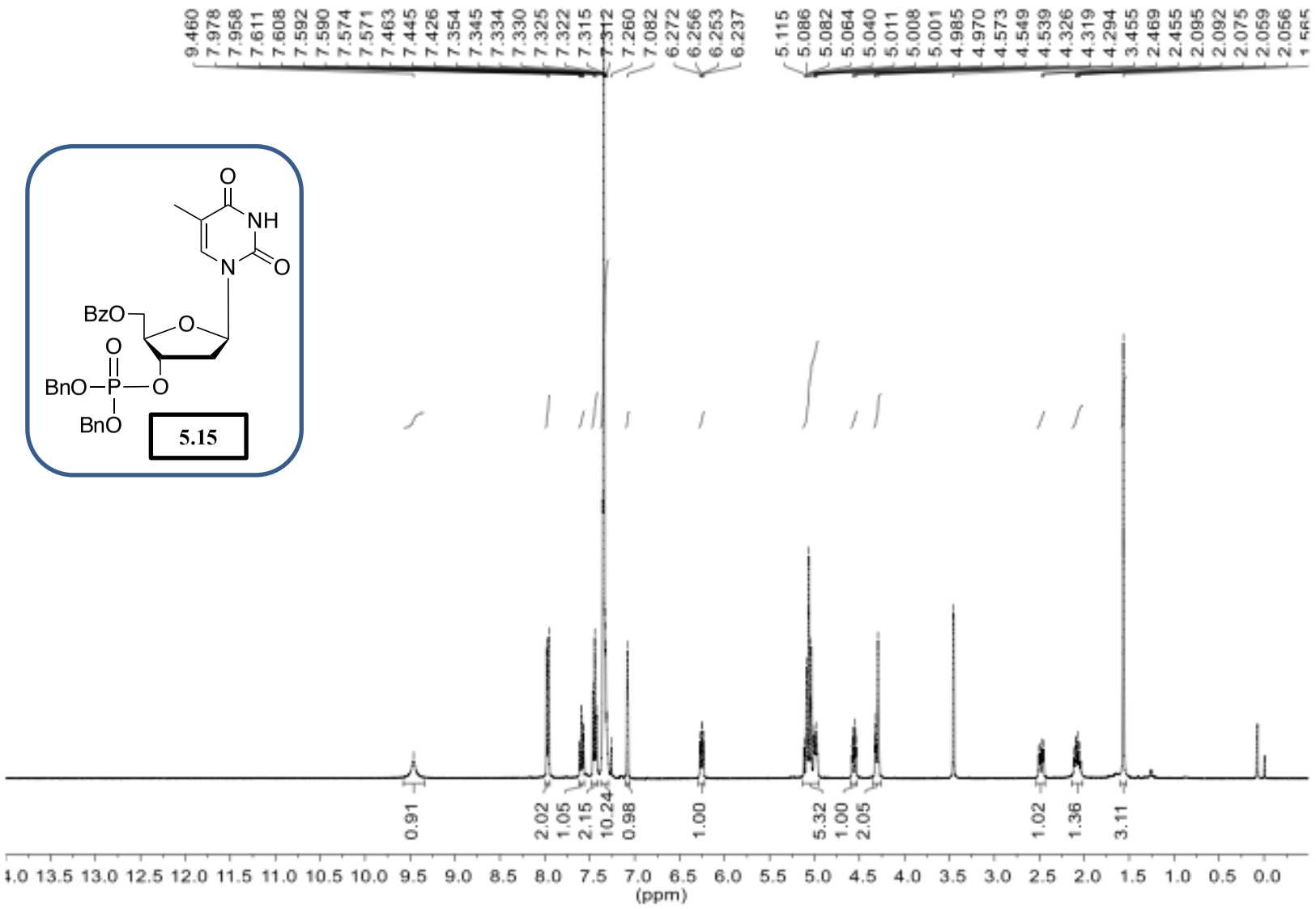
^{31}P NMR spectrum of compound **5.12d** (500 MHz,D_2O)



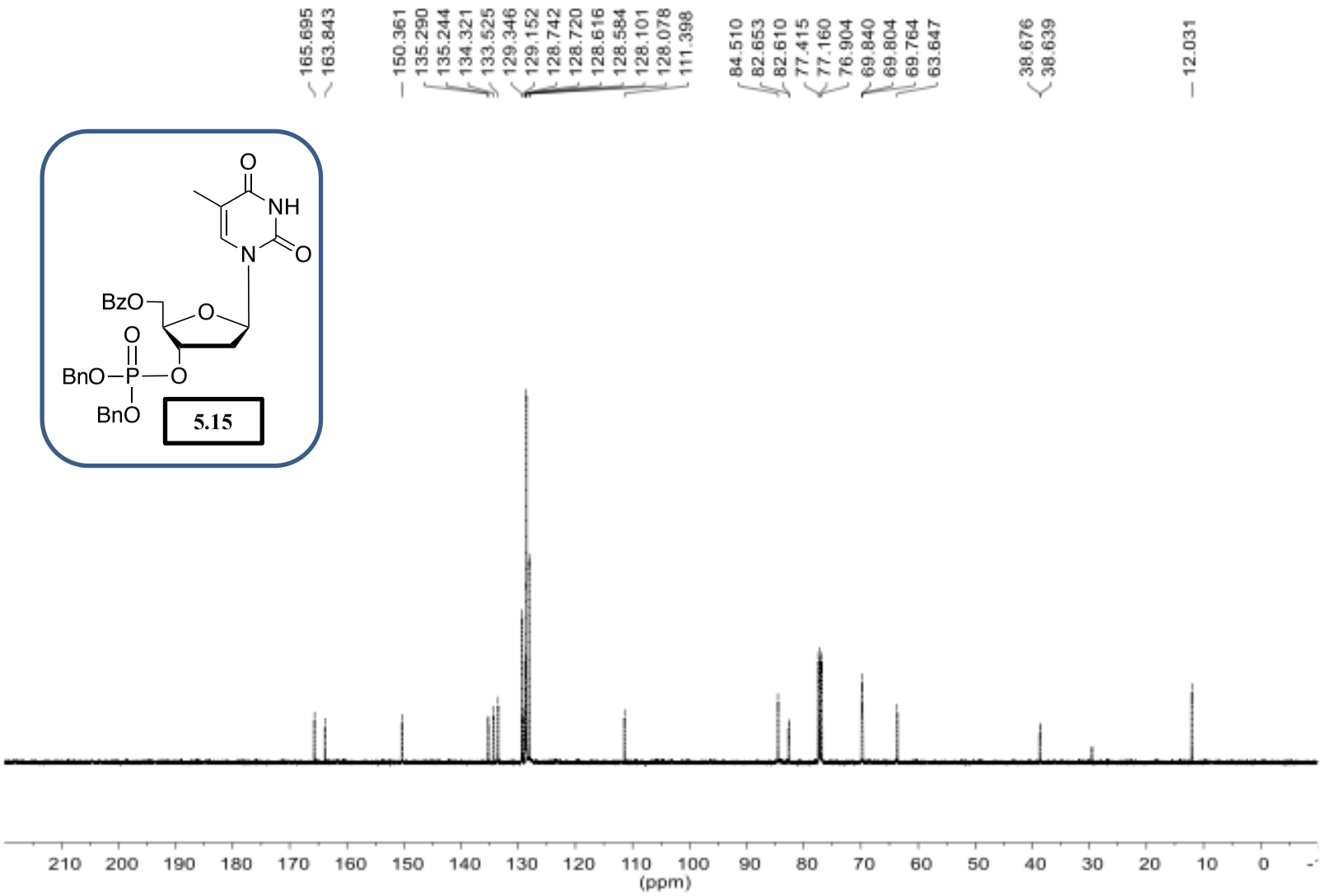
^{1}H NMR spectrum of compound **5.13d** (500 MHz,D_2O)



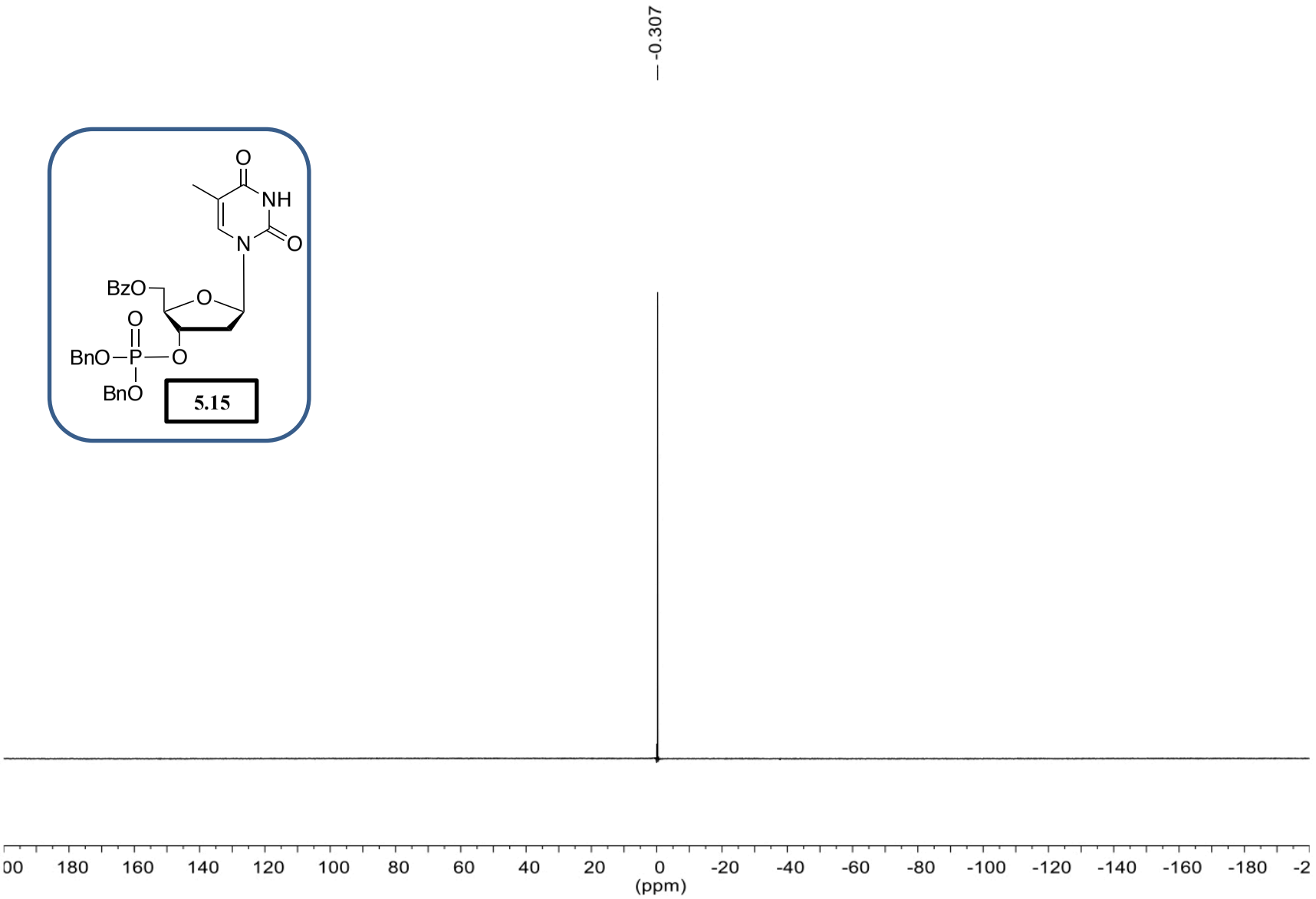
^{31}P NMR spectrum of compound **5.13d** (500 MHz,D_2O)



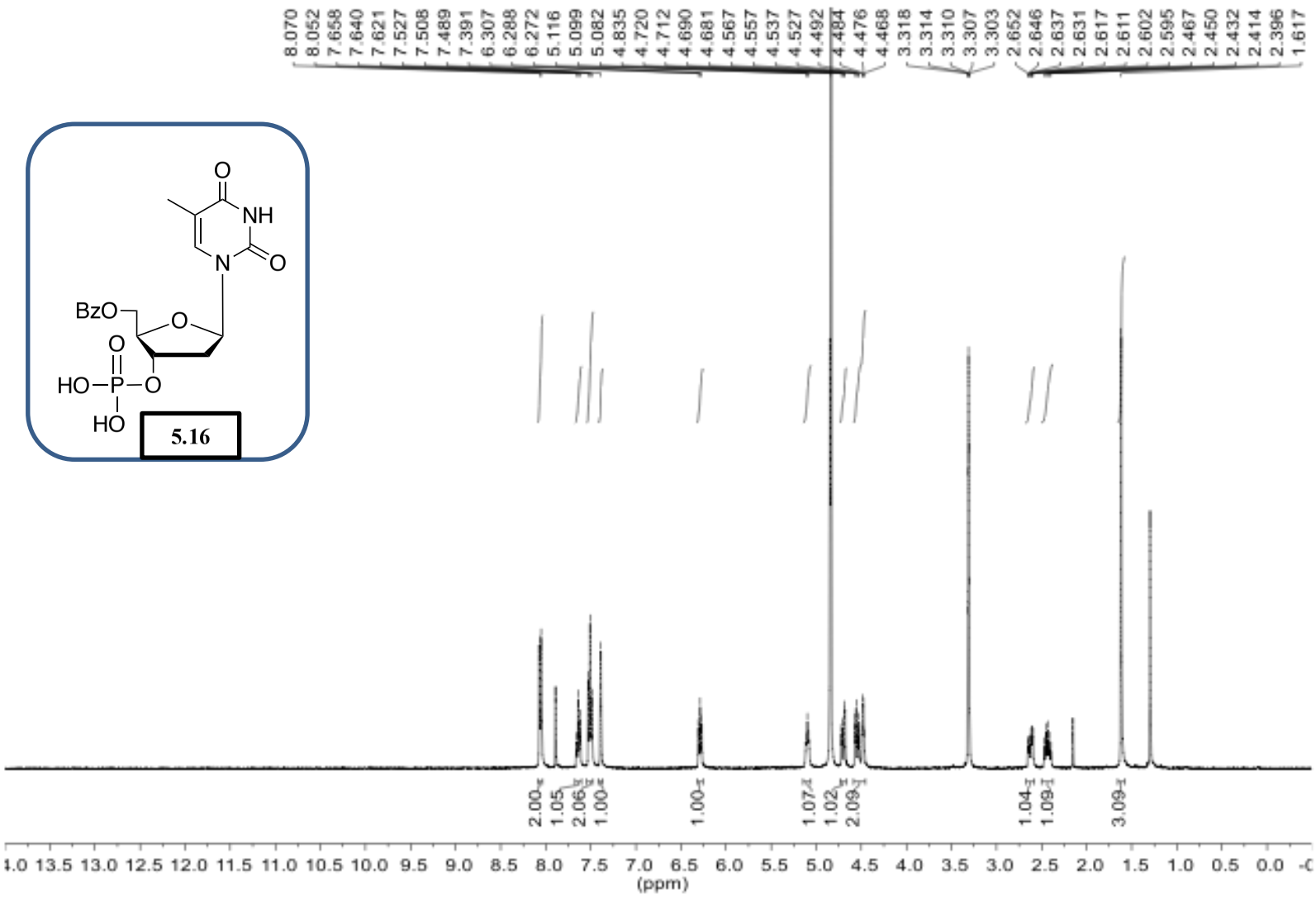
^{1}H NMR spectrum of compound **5.15** (500 MHz,$CDCl_3$)



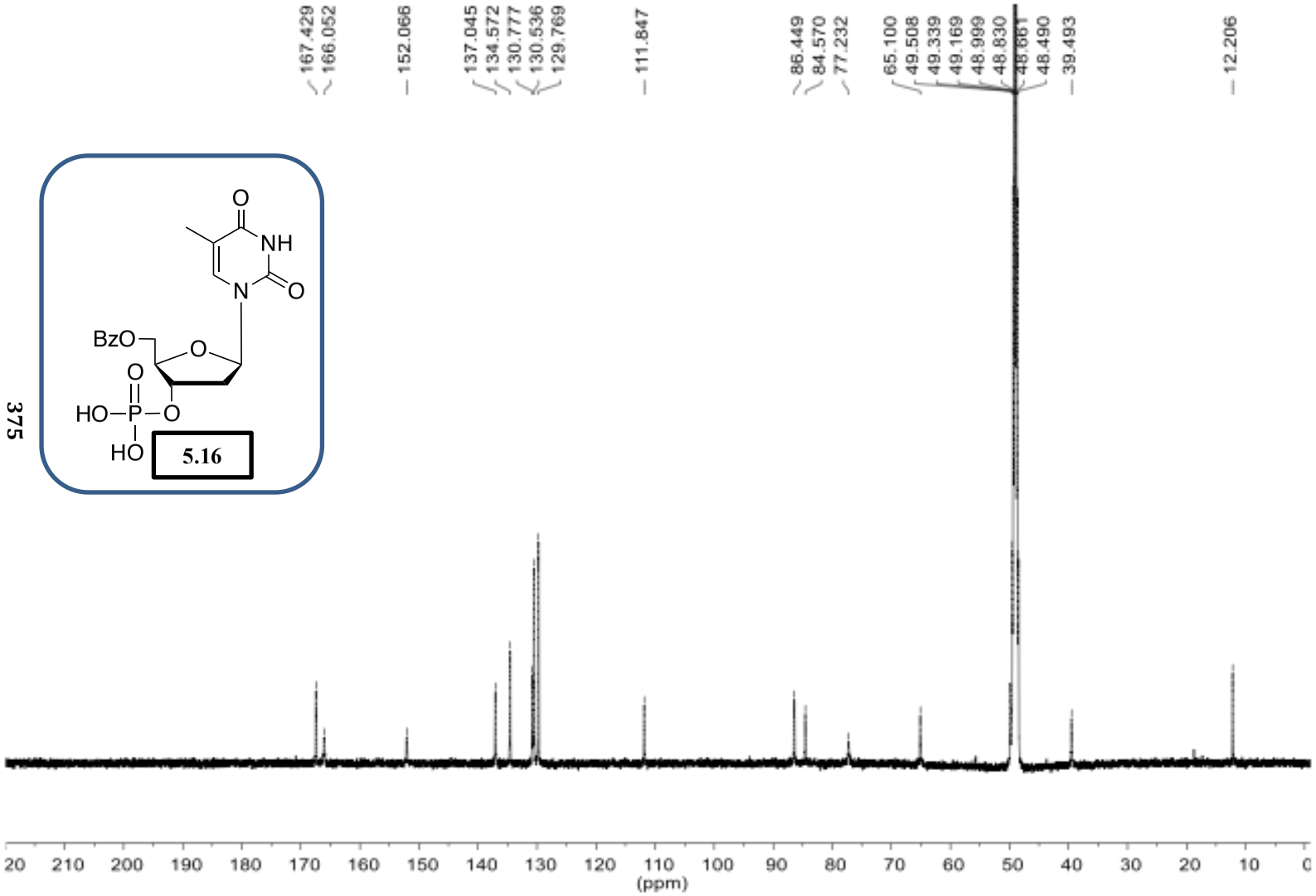
^{13}C NMR spectrum of compound **5.15** (500 MHz,$CDCl_3$)



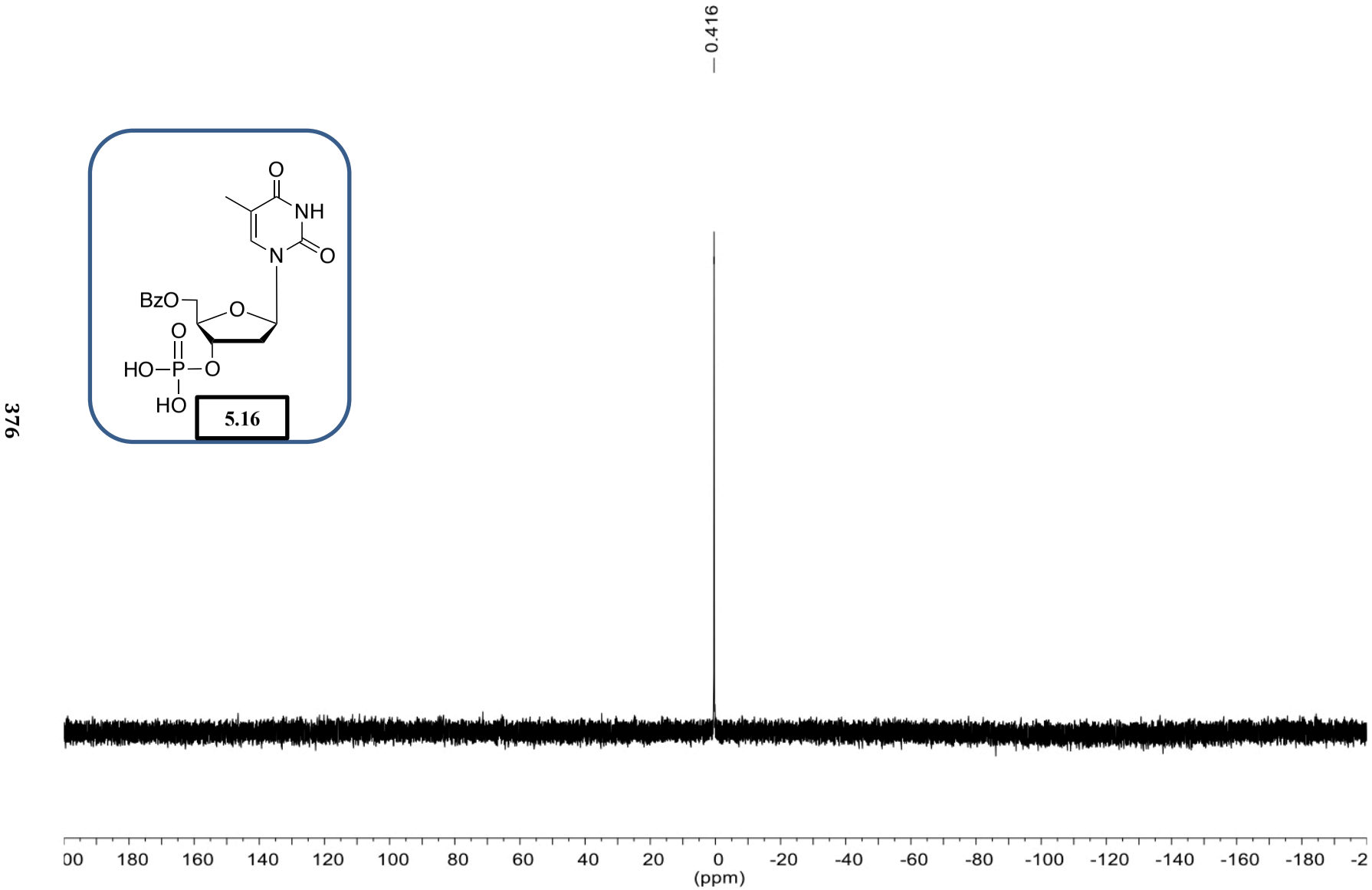
^{31}P NMR spectrum of compound **5.15** (500 MHz,$CDCl_3$)



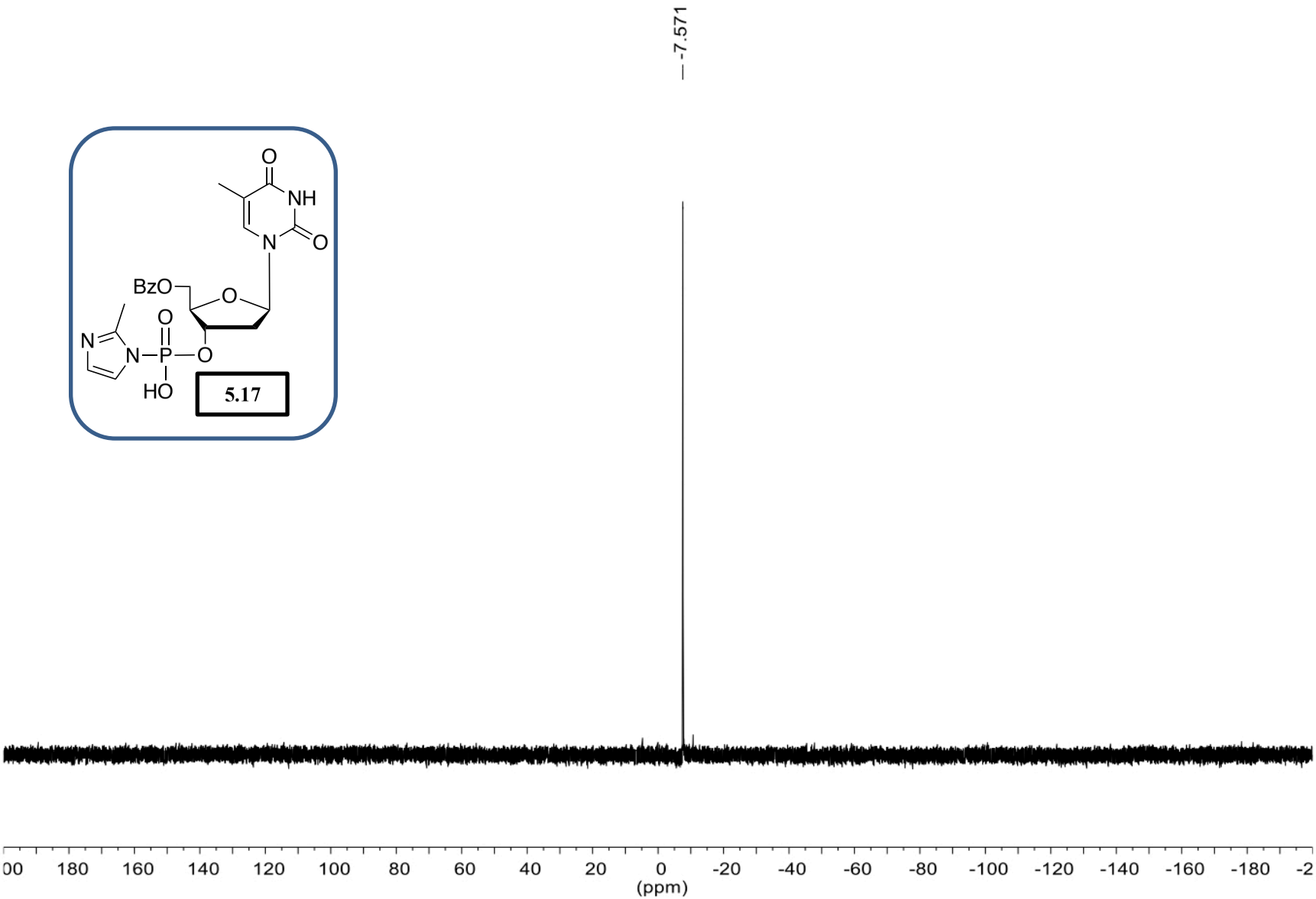
^{1}H NMR spectrum of compound **5.16** (500 MHz,CD_3OD)



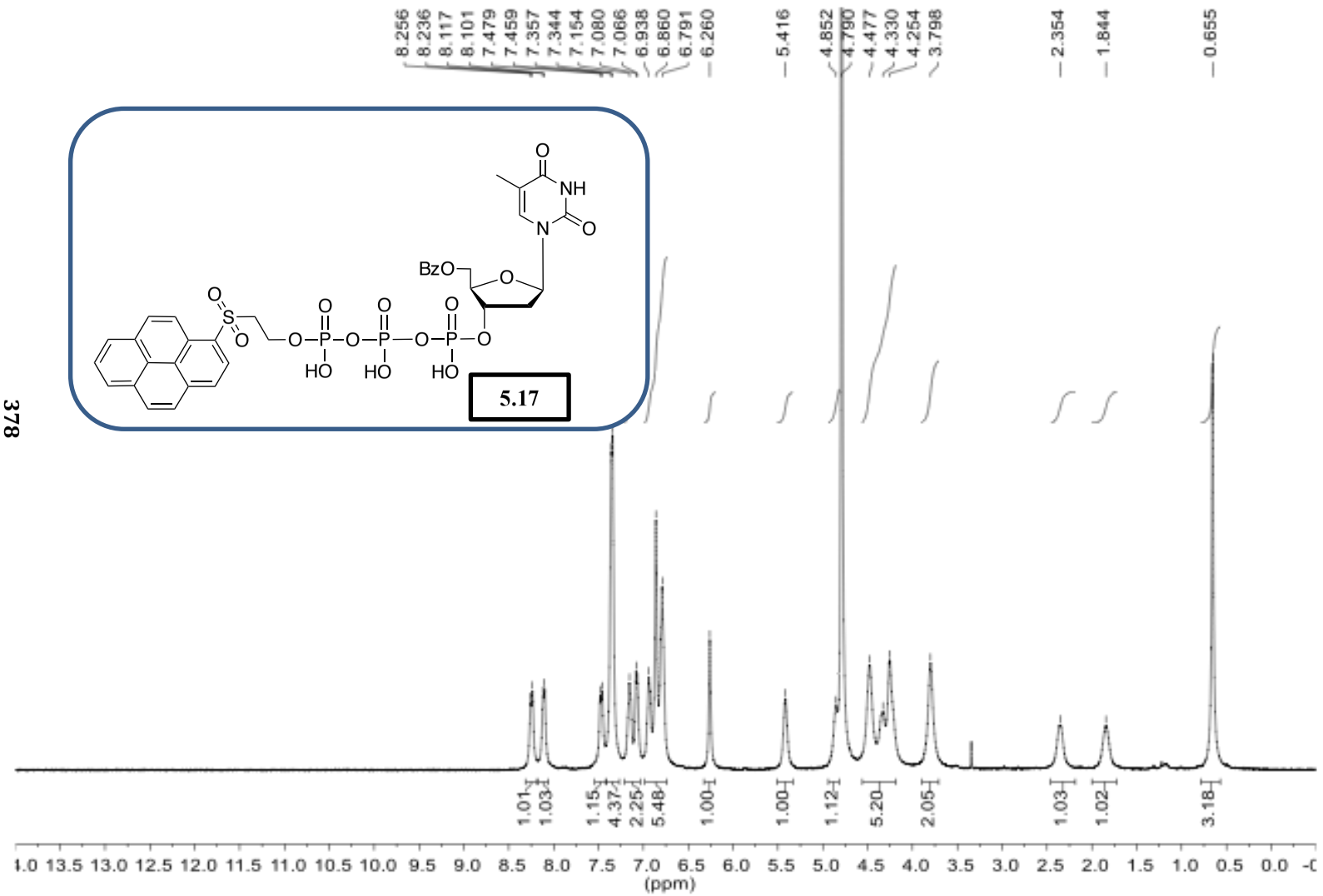
^{13}C NMR spectrum of compound **5.16** (500 MHz,CD_3OD)



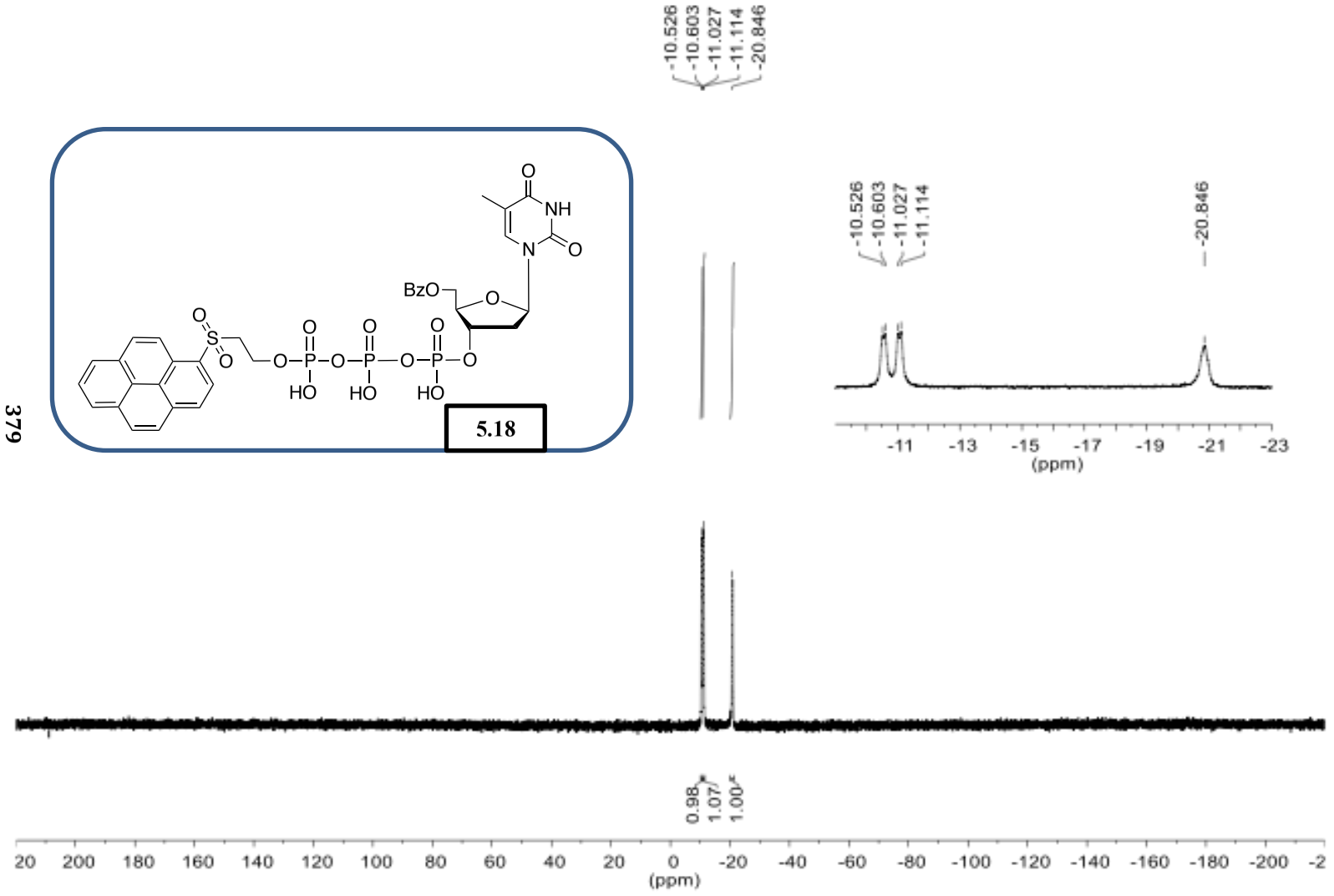
^{31}P NMR spectrum of compound **5.16** (500 MHz,CD_3OD)



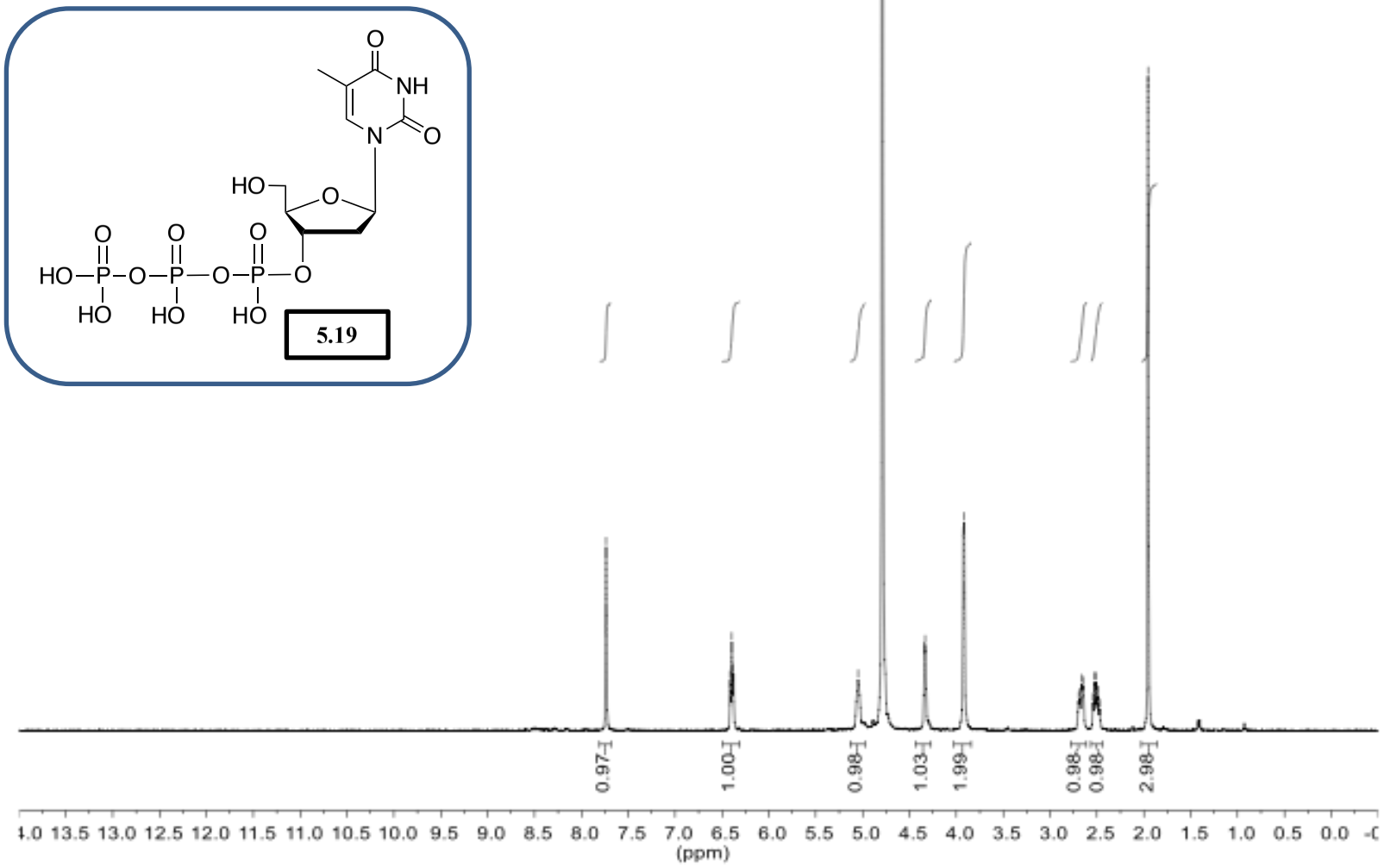
^{31}P NMR spectrum of compound **5.17** (500 MHz,D_2O)



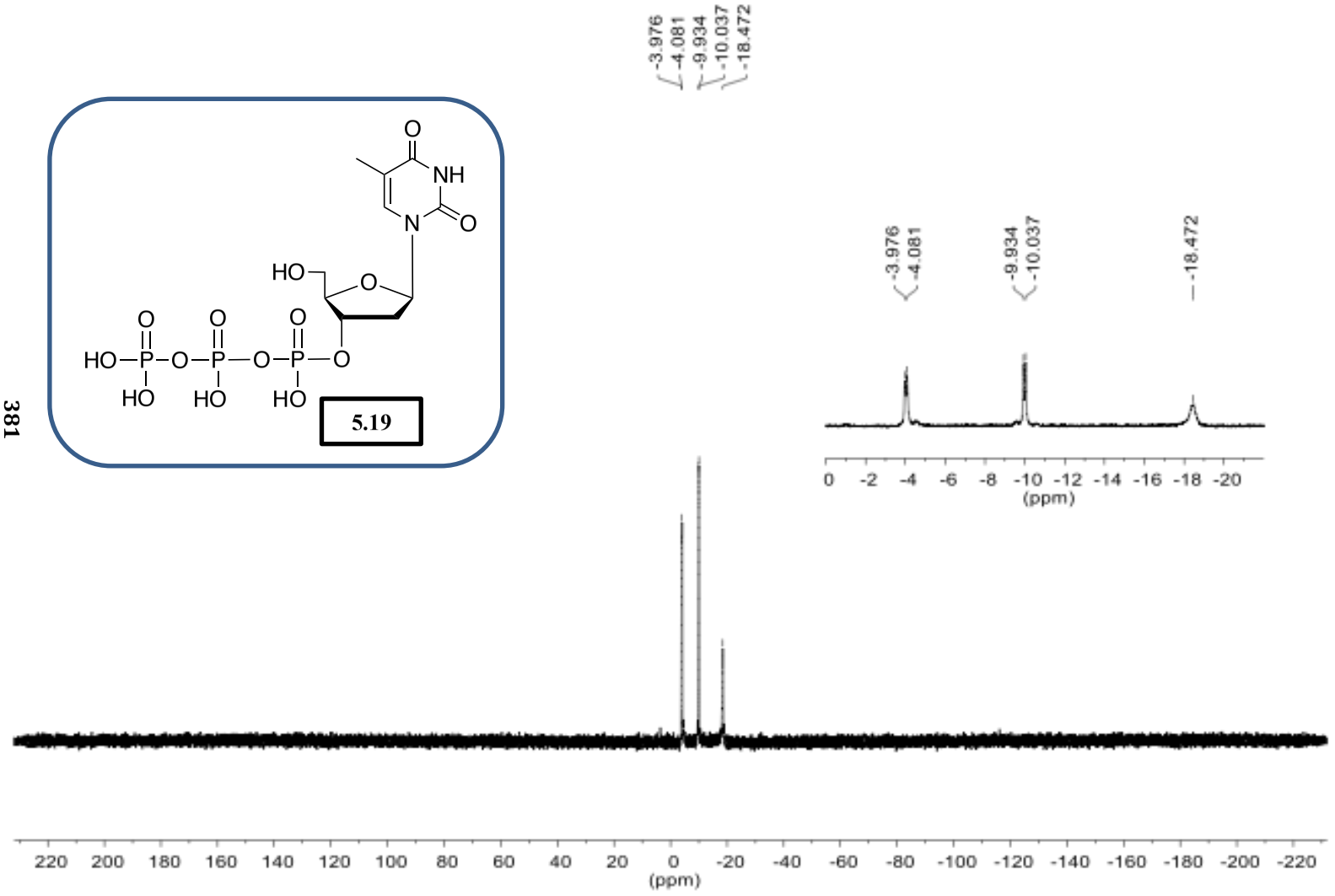
^{1}H NMR spectrum of compound **5.18** (500 MHz,D_2O)



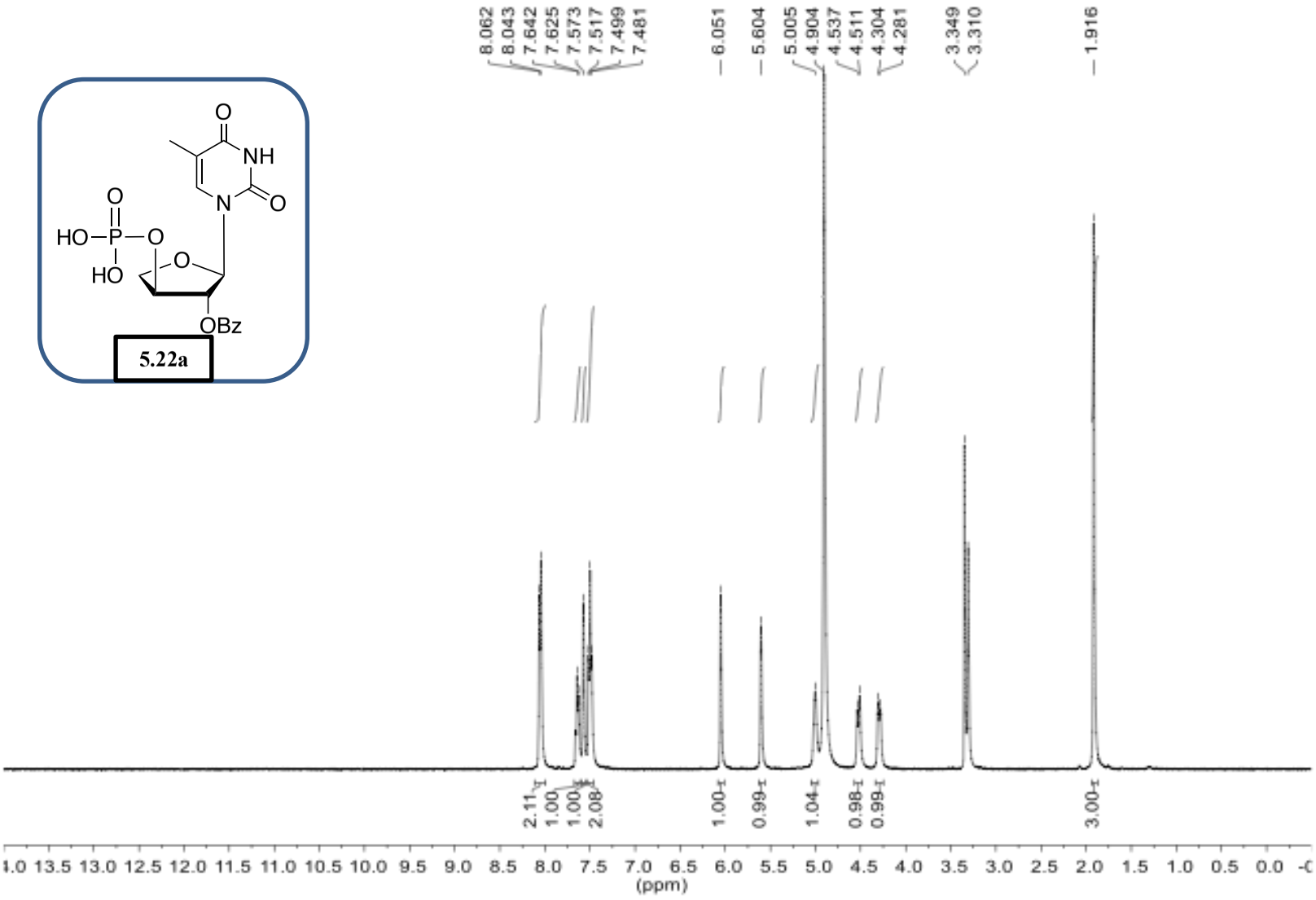
^{31}P NMR spectrum of compound **5.18** (500 MHz,D_2O)



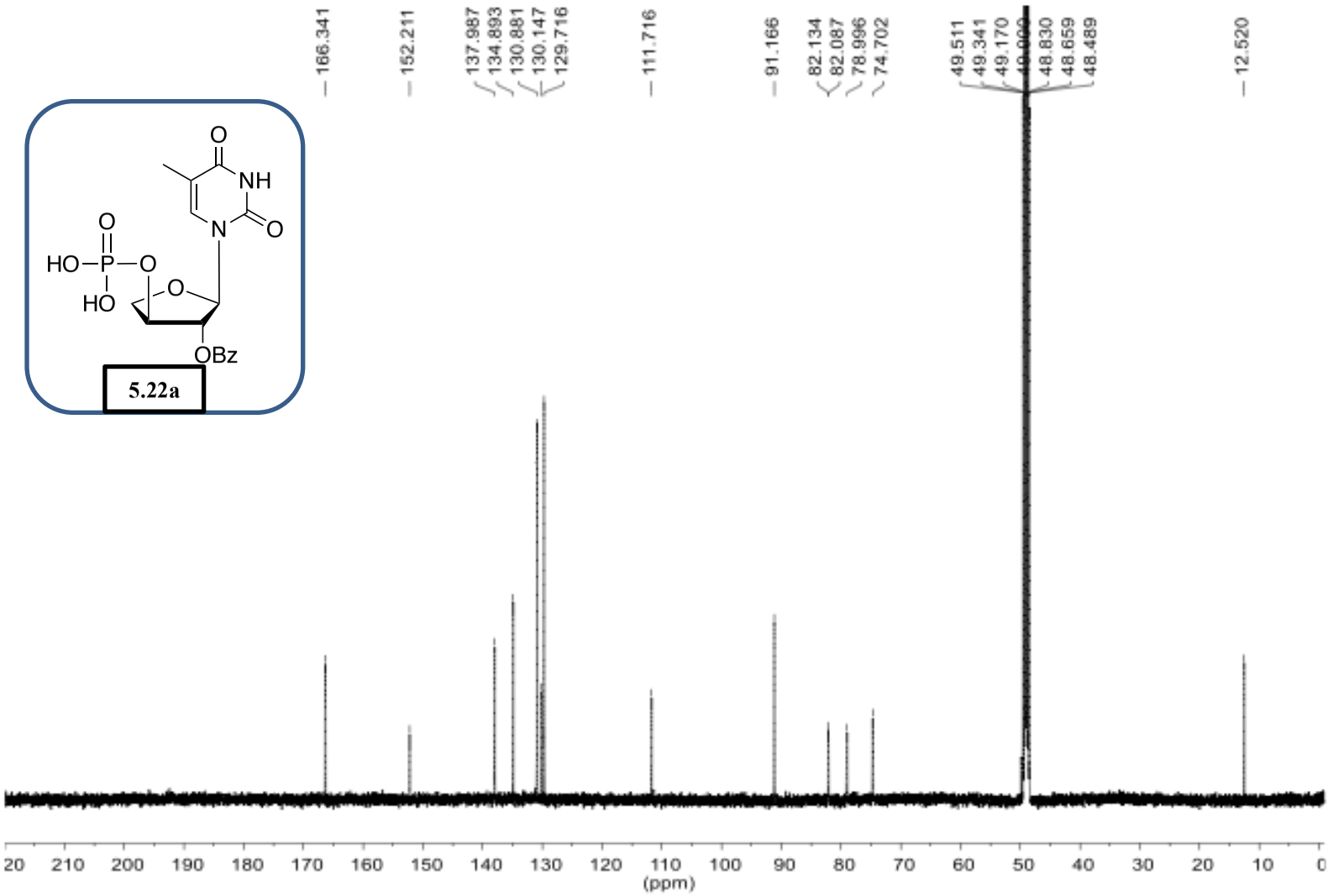
^{1}H NMR spectrum of compound **5.19** (500 MHz,D_2O)



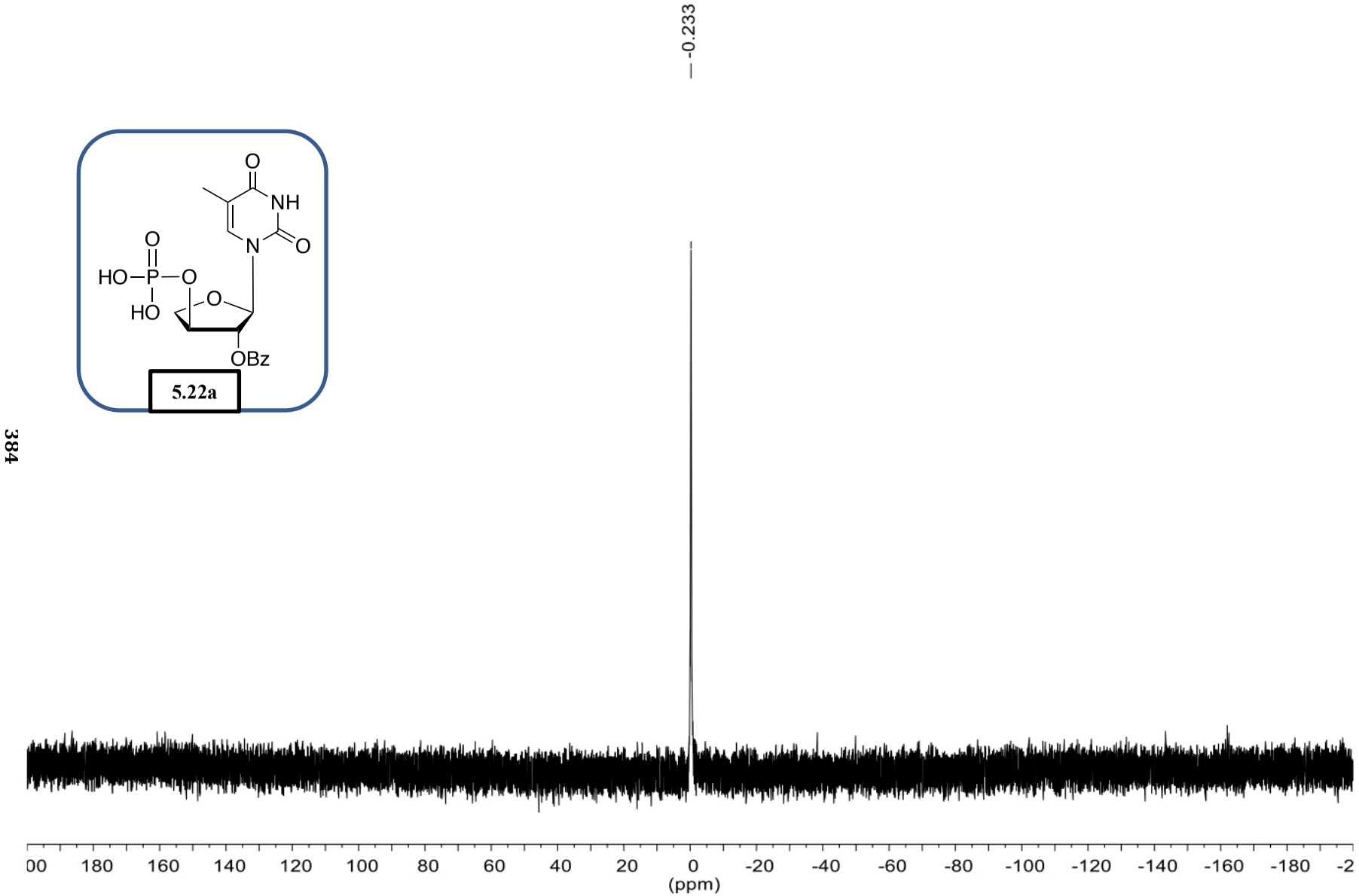
^{31}P NMR spectrum of compound **5.19** (500 MHz,D_2O)



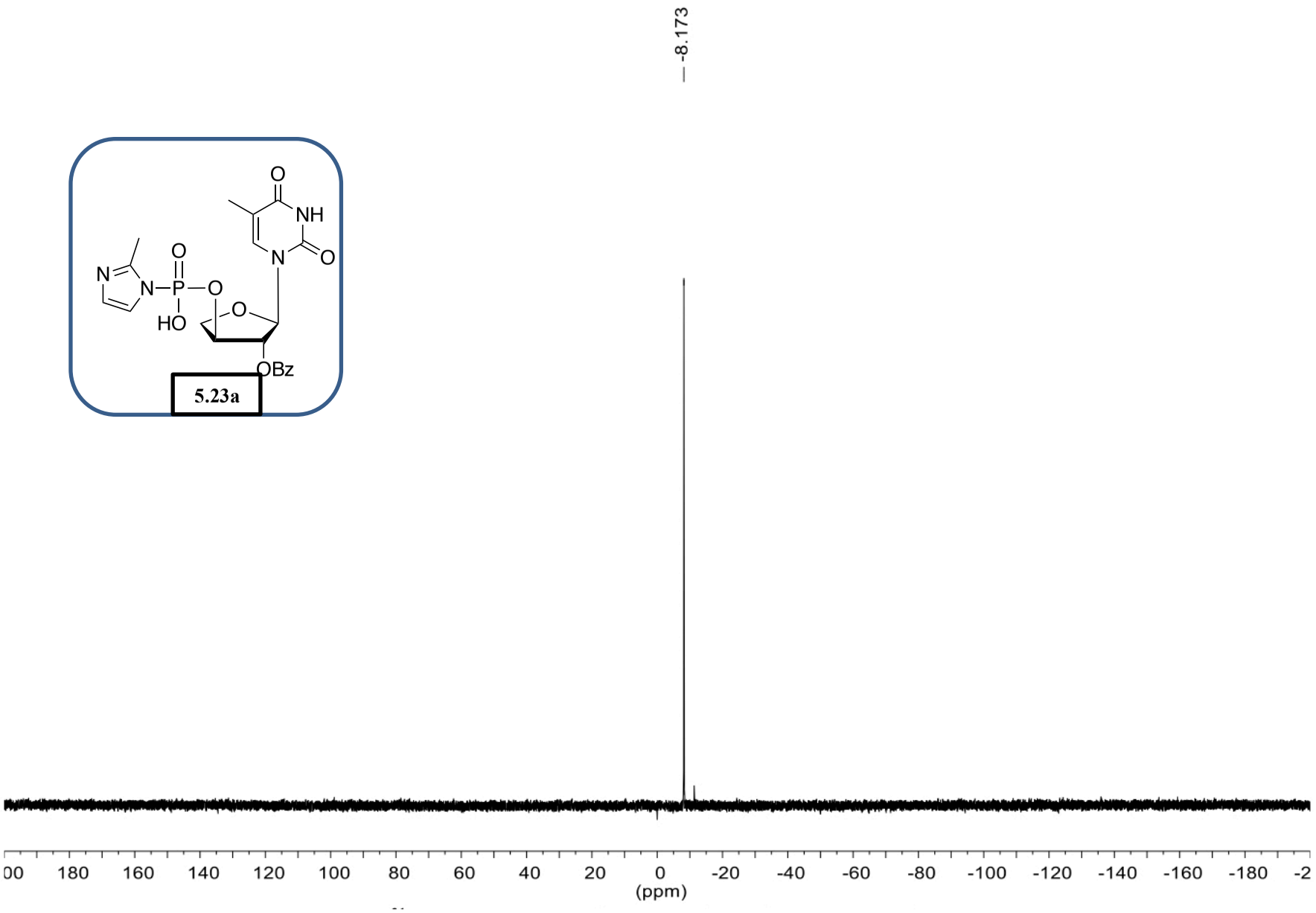
^{1}H NMR spectrum of compound **5.22a** (500 MHz,CD_3OD)



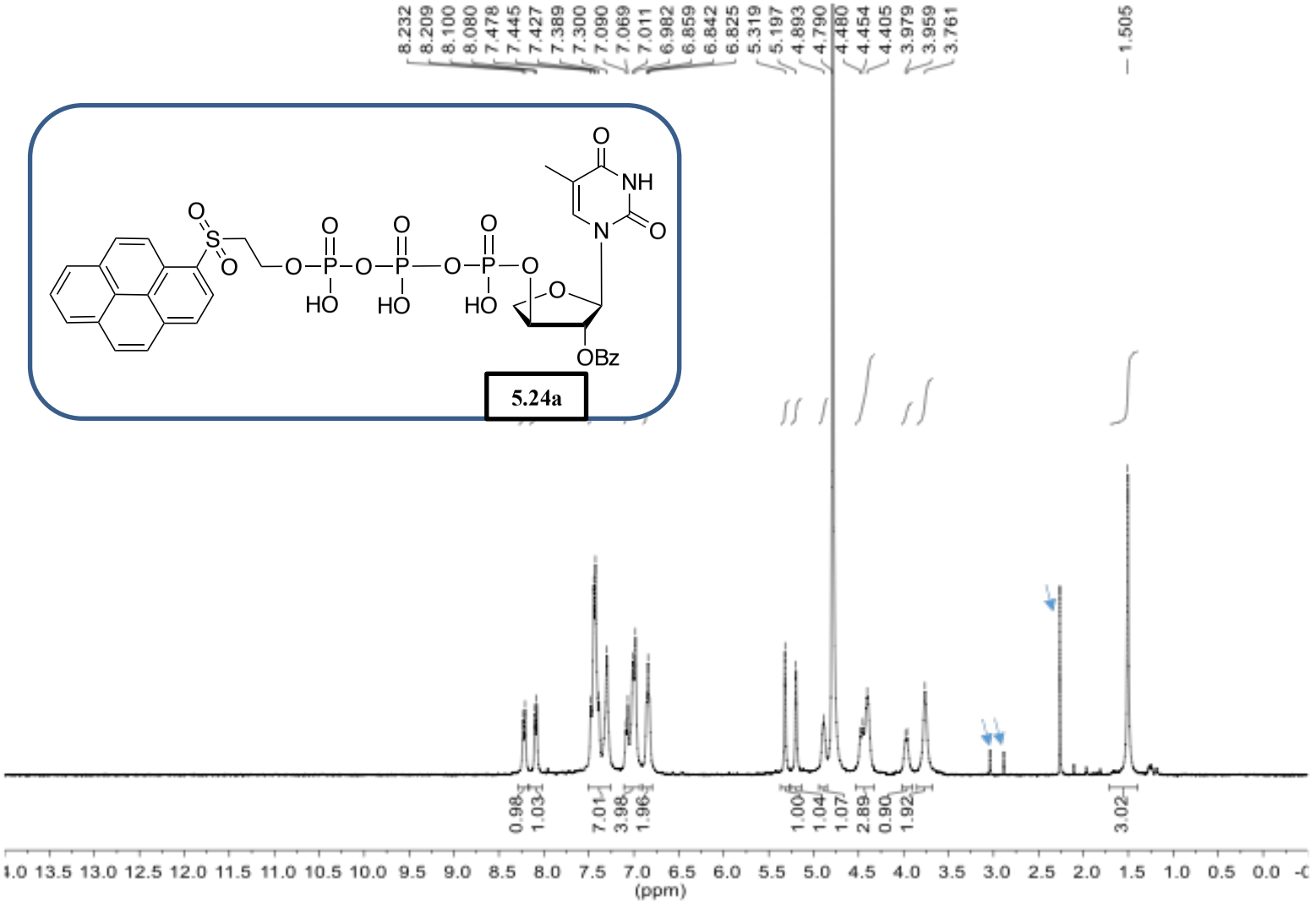
^{13}C NMR spectrum of compound **5.22a** (500 MHz,CD_3OD)



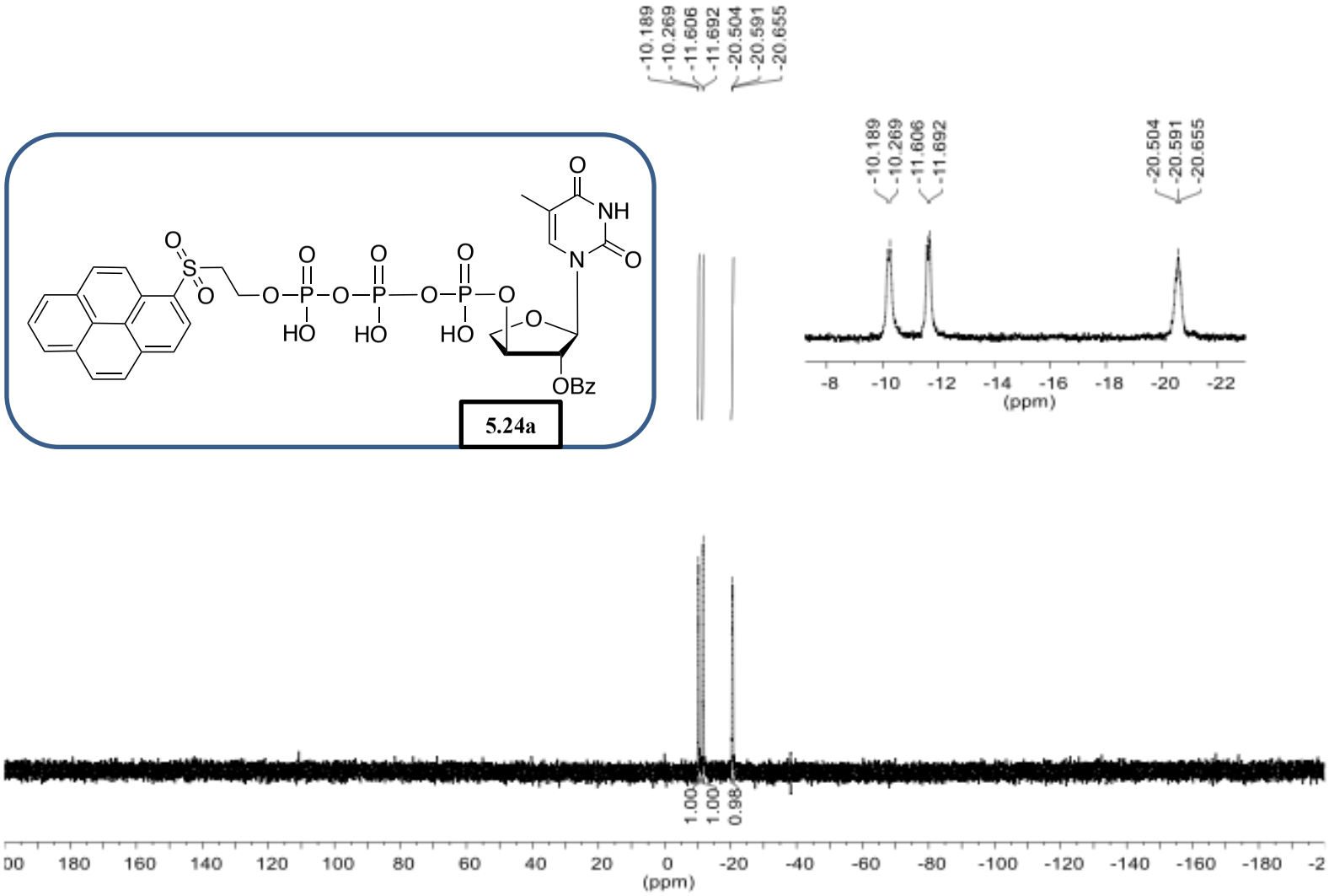
^{31}P NMR spectrum of compound **5.22a** (500 MHz,CD_3OD)



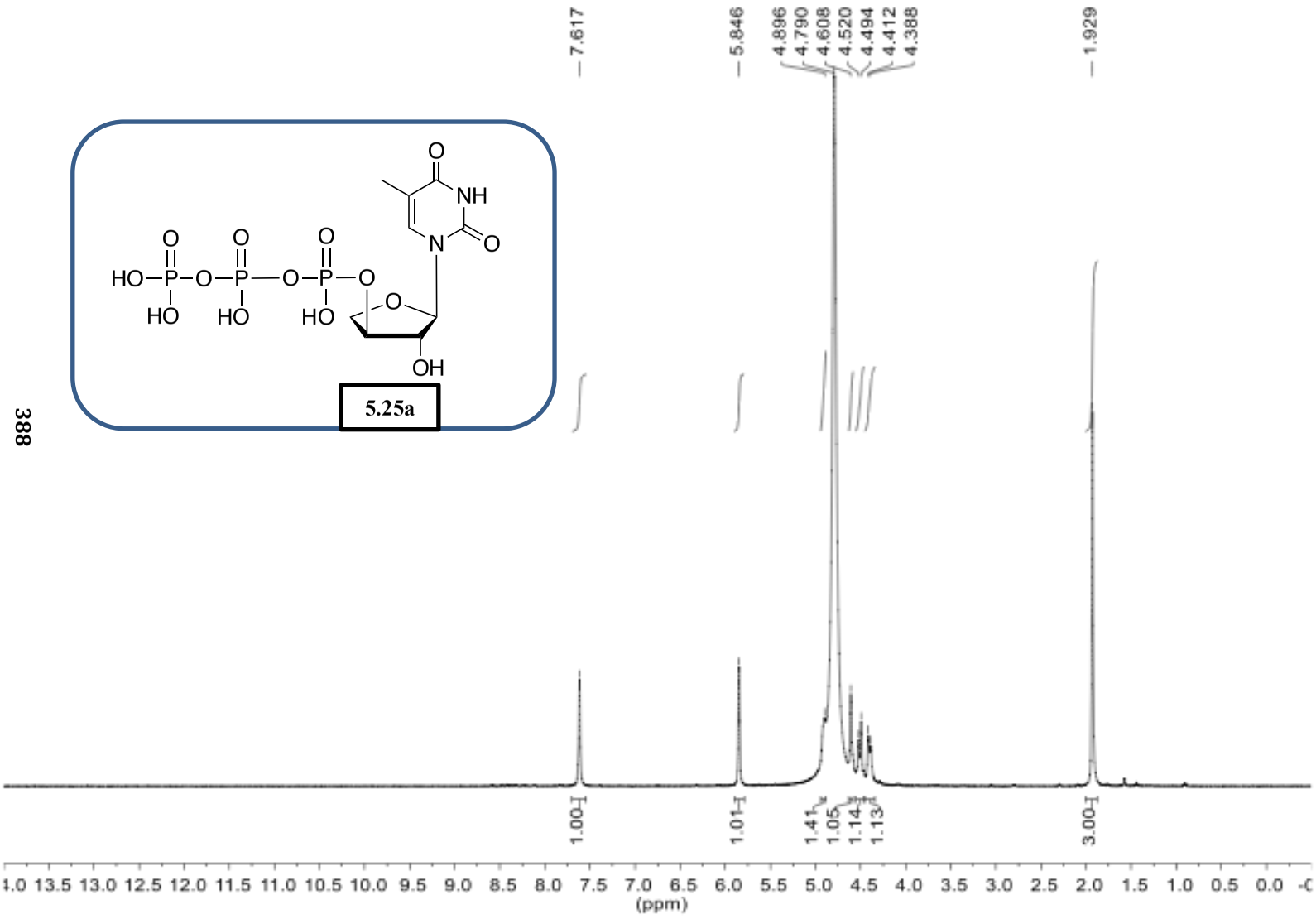
^{31}P NMR spectrum of compound **5.23a** (500 MHz,CD_3OD)



^{1}H NMR spectrum of compound **5.24a** (500 MHz,D_2O)

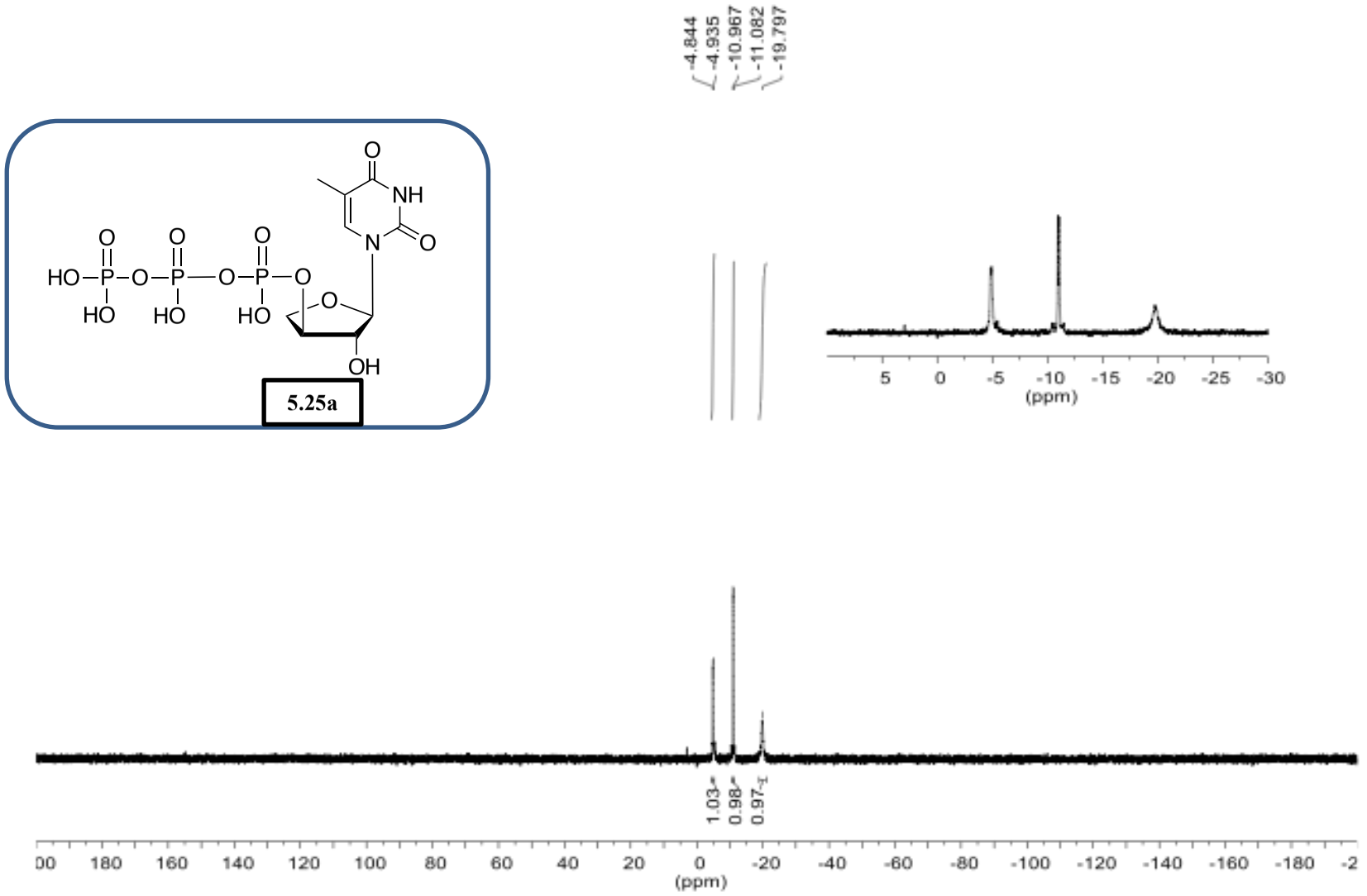


^{31}P NMR spectrum of compound **5.24a** (500 MHz,D_2O)

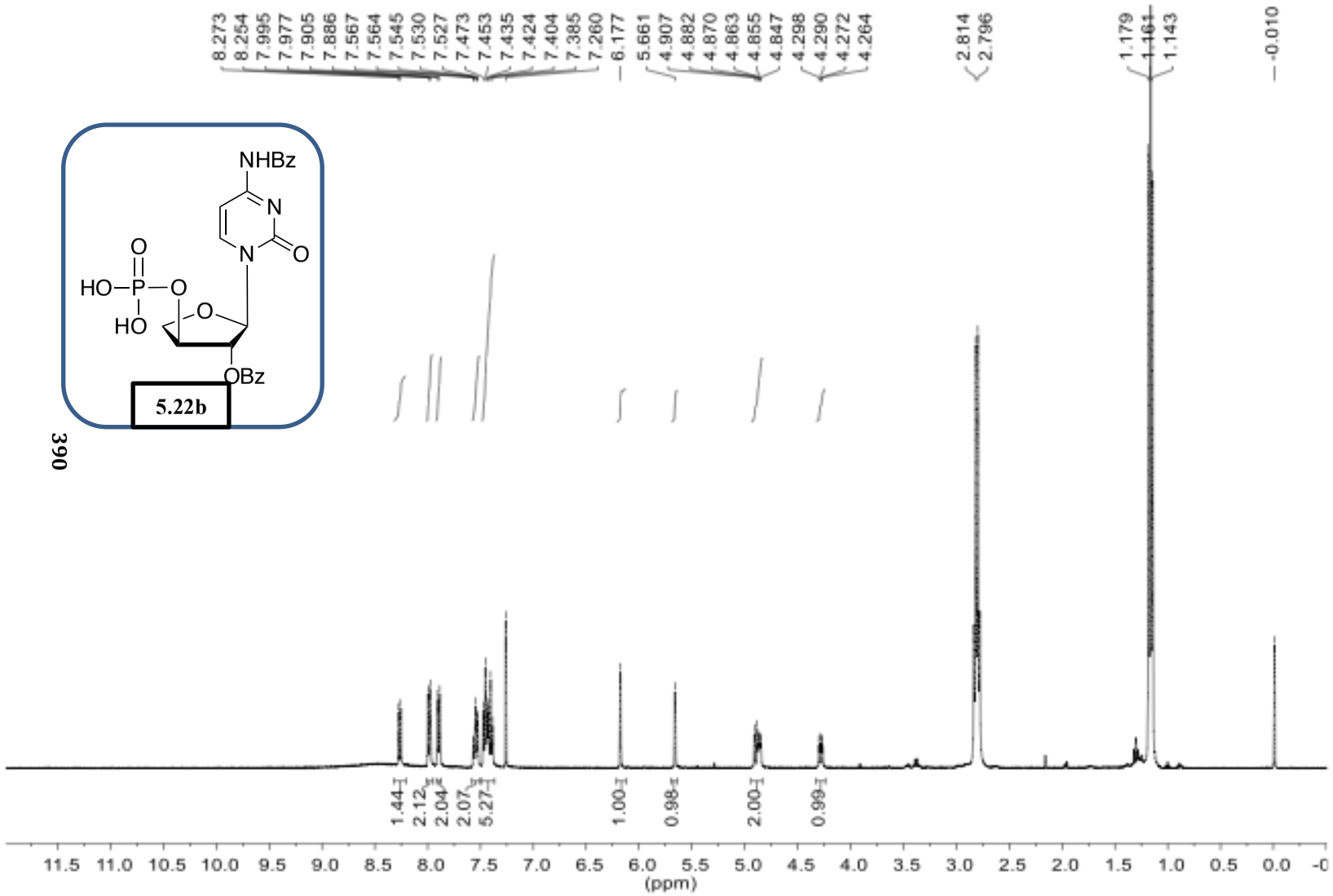


388

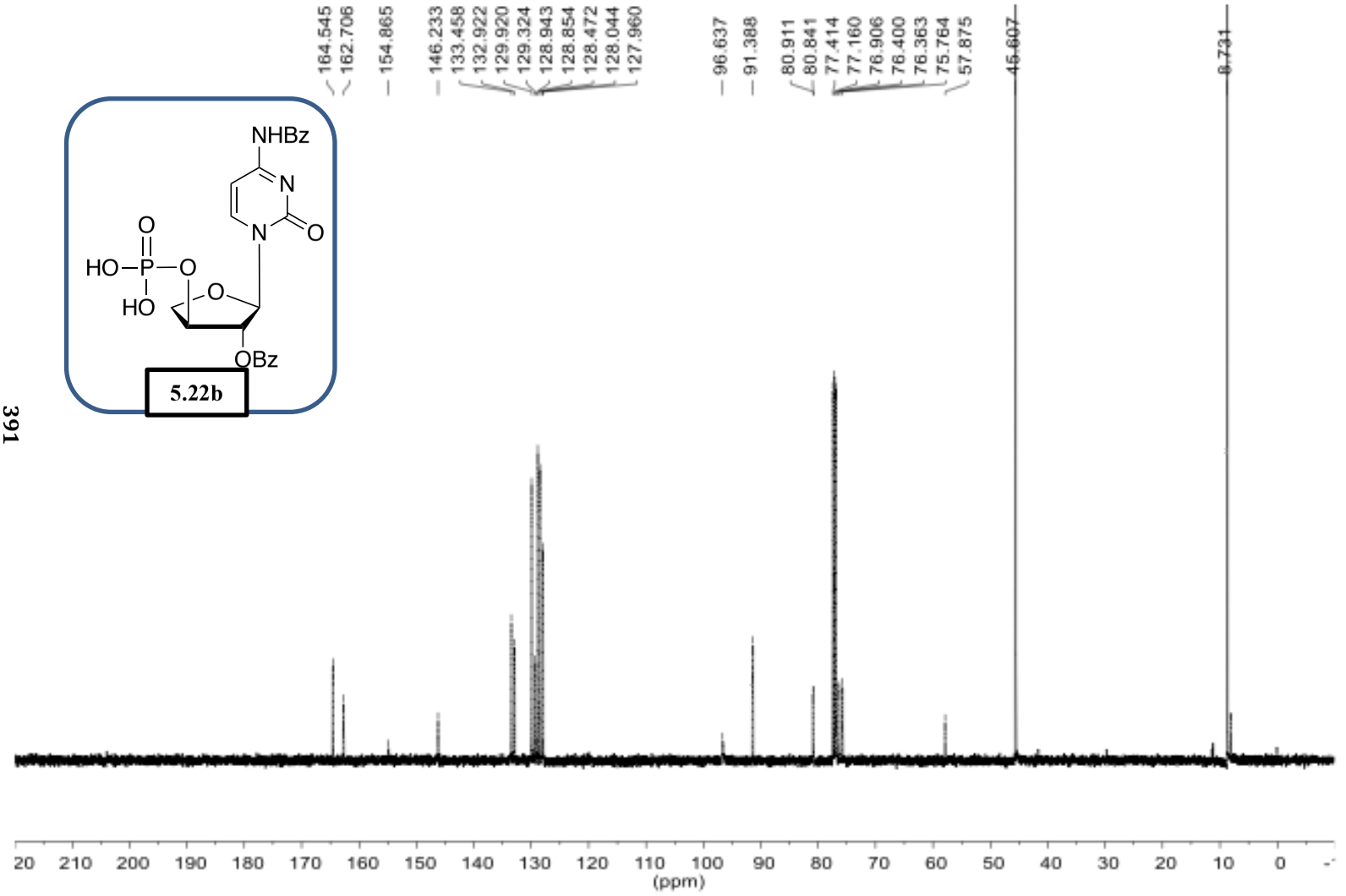
^{1}H NMR spectrum of compound **5.25a** (500 MHz,D_2O)



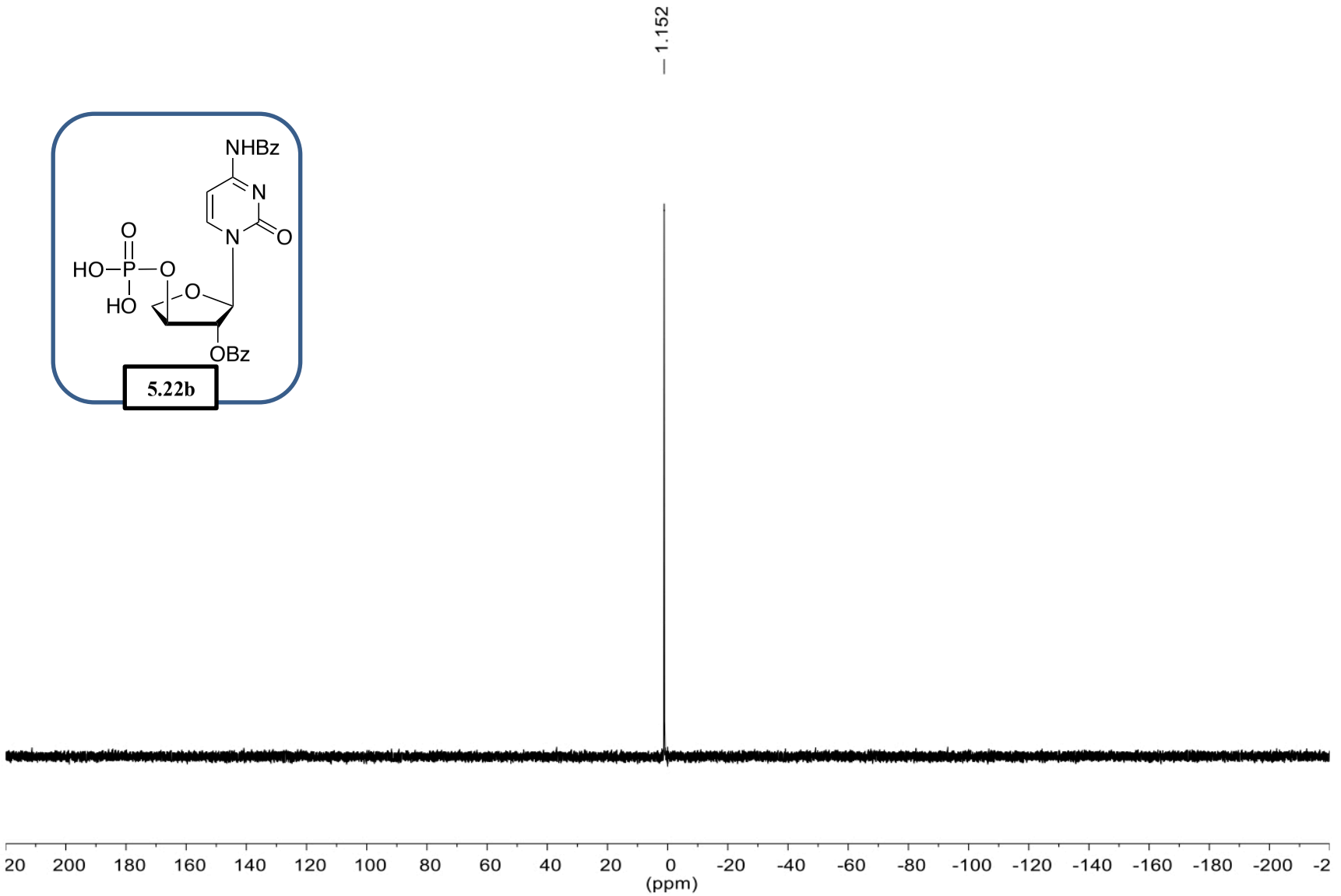
^{31}P NMR spectrum of compound **5.25a** (500 MHz,D_2O)



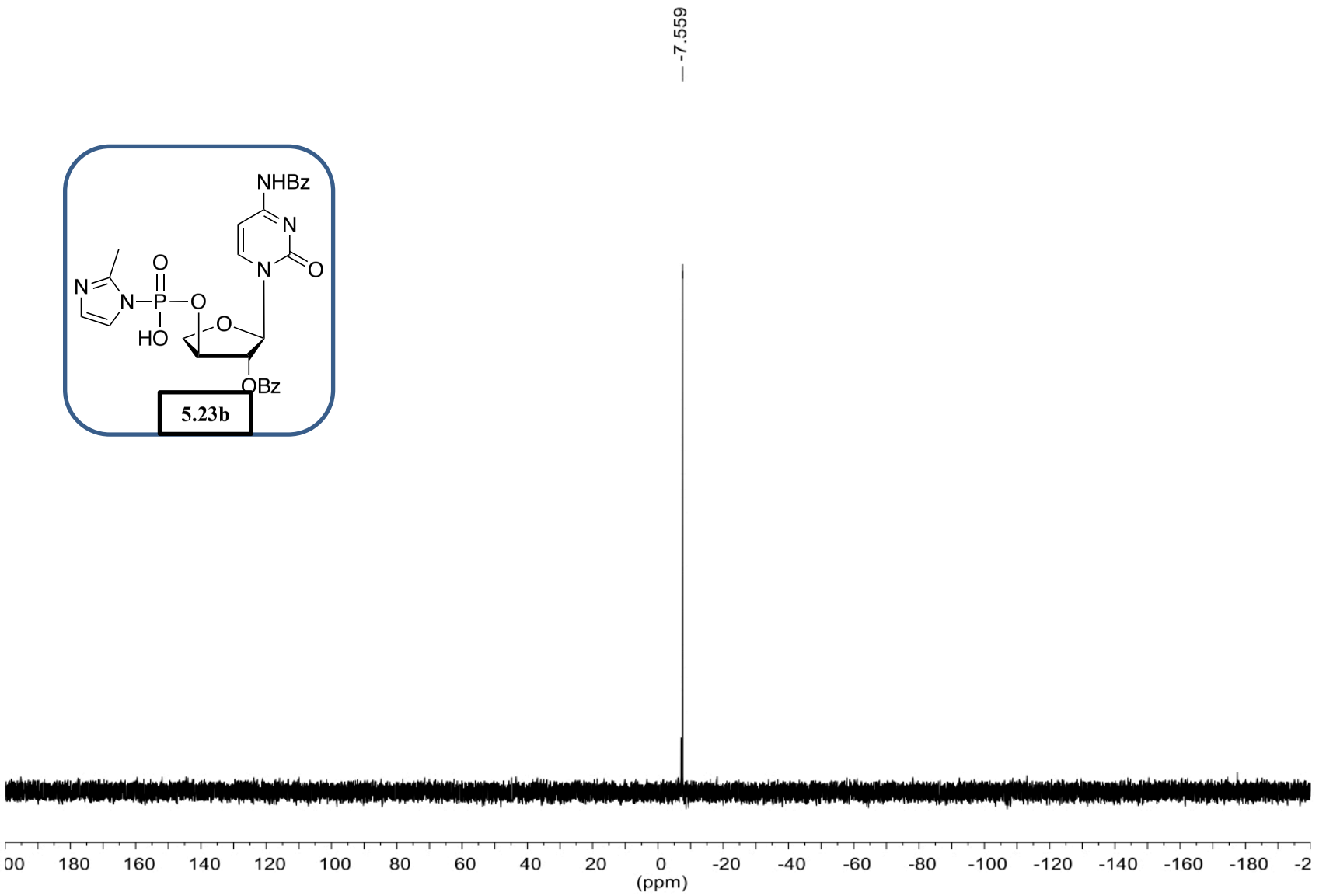
^{1}H NMR spectrum of compound **5.22b** (500 MHz,D_2O)



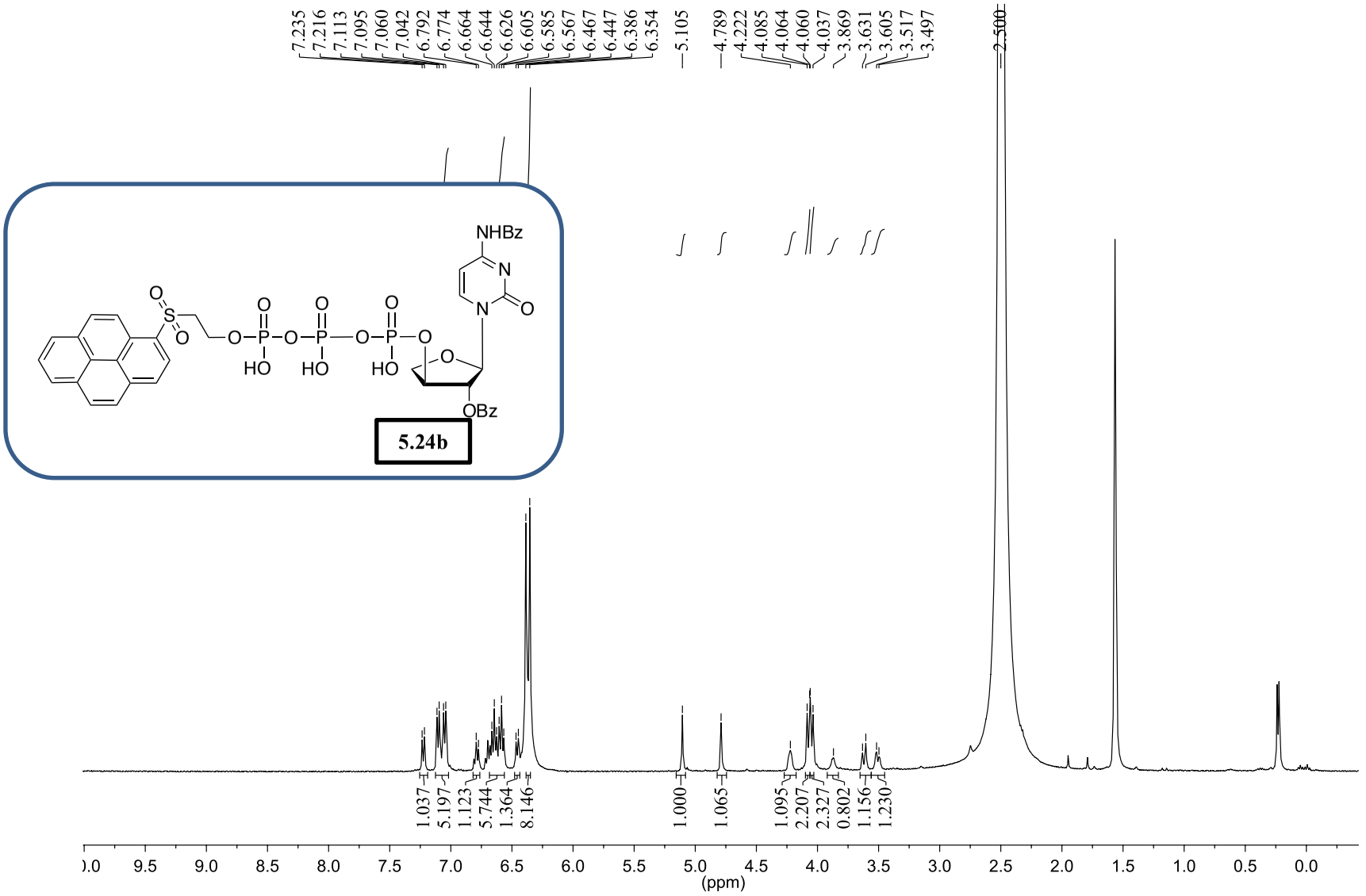
^{13}C NMR spectrum of compound **5.22b** (500 MHz,D_2O)



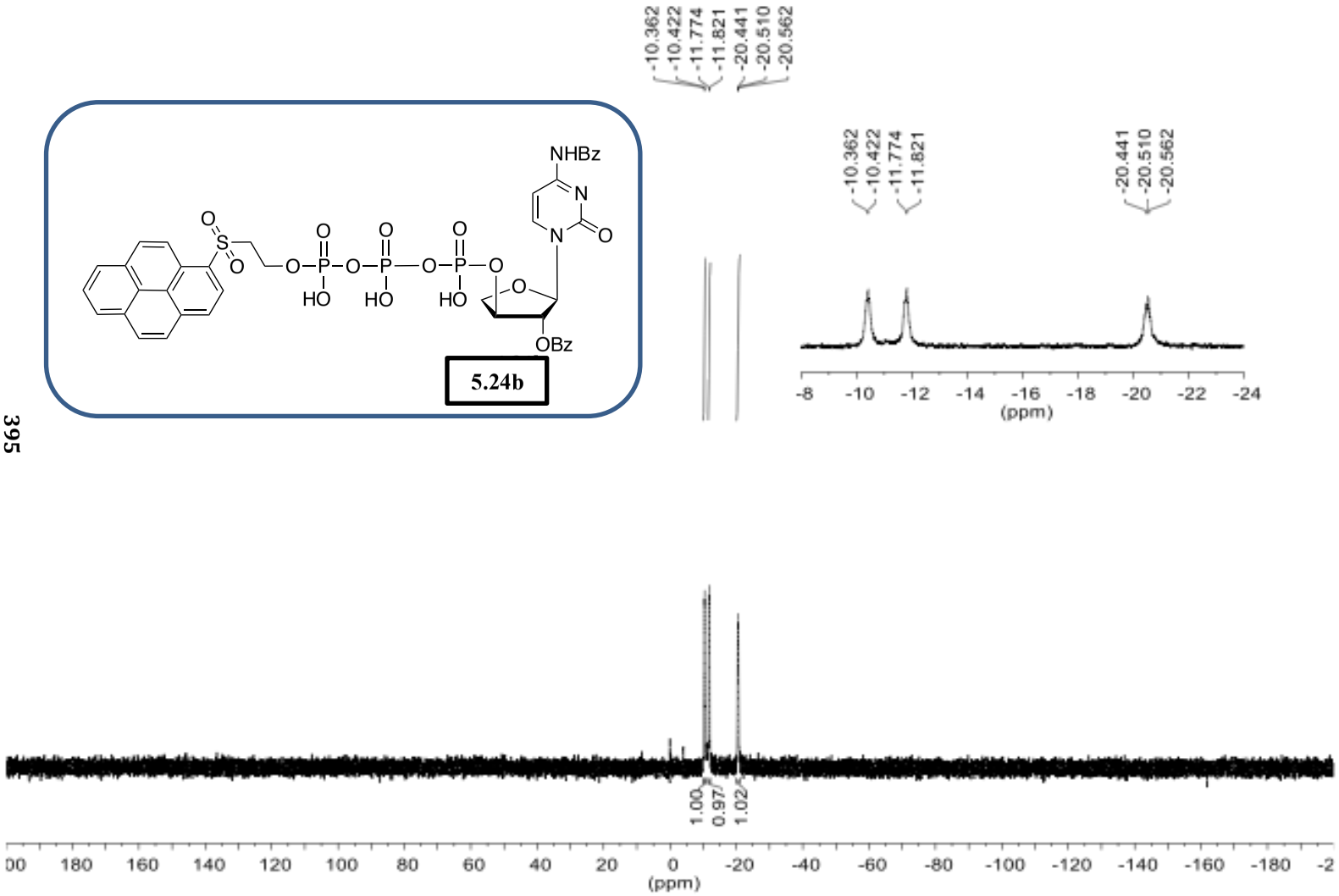
^{31}P NMR spectrum of compound **4.22b** (500 MHz,D_2O)



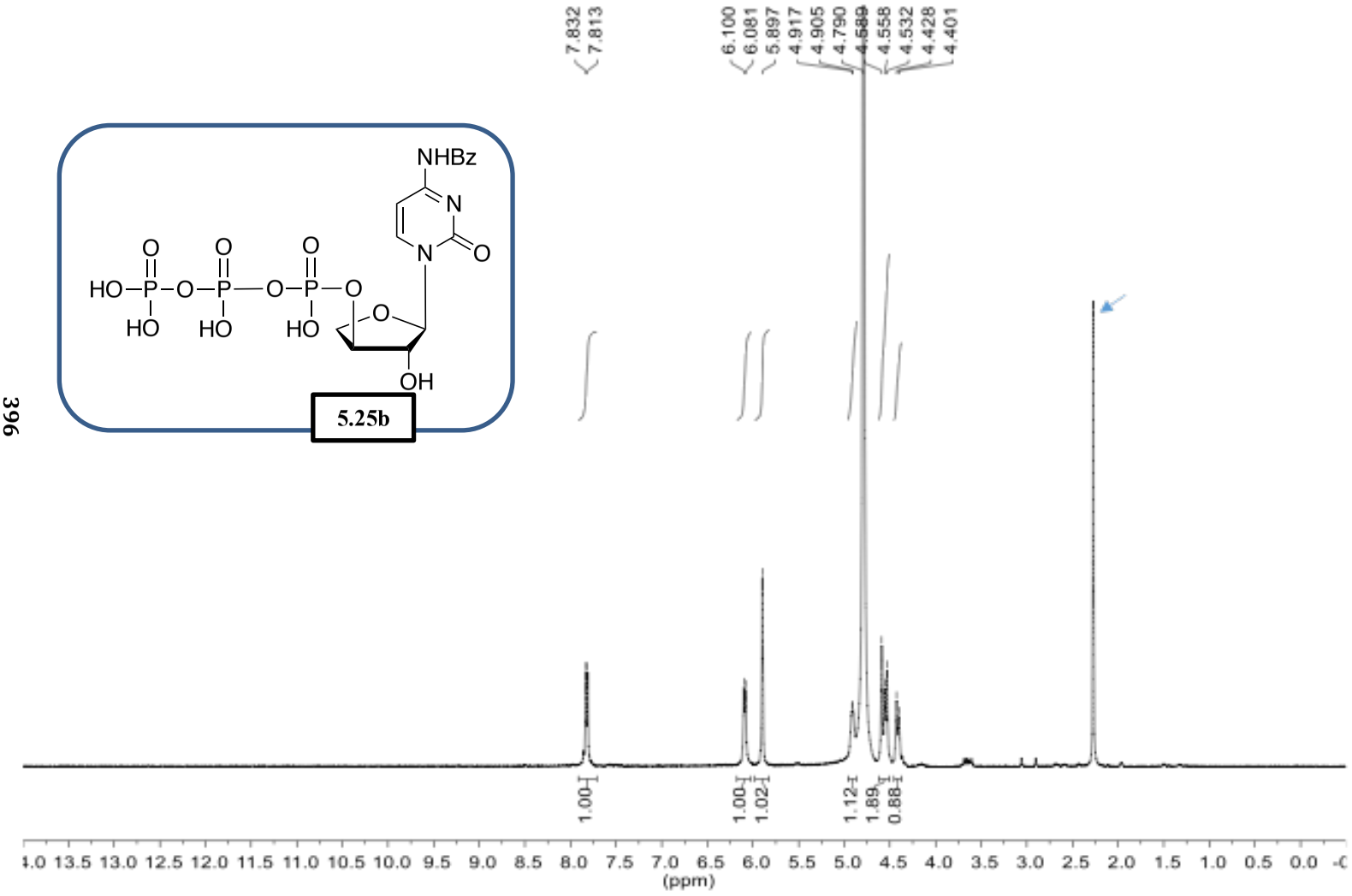
^{31}P NMR spectrum of compound **5.23b** (500 MHz,D_2O)



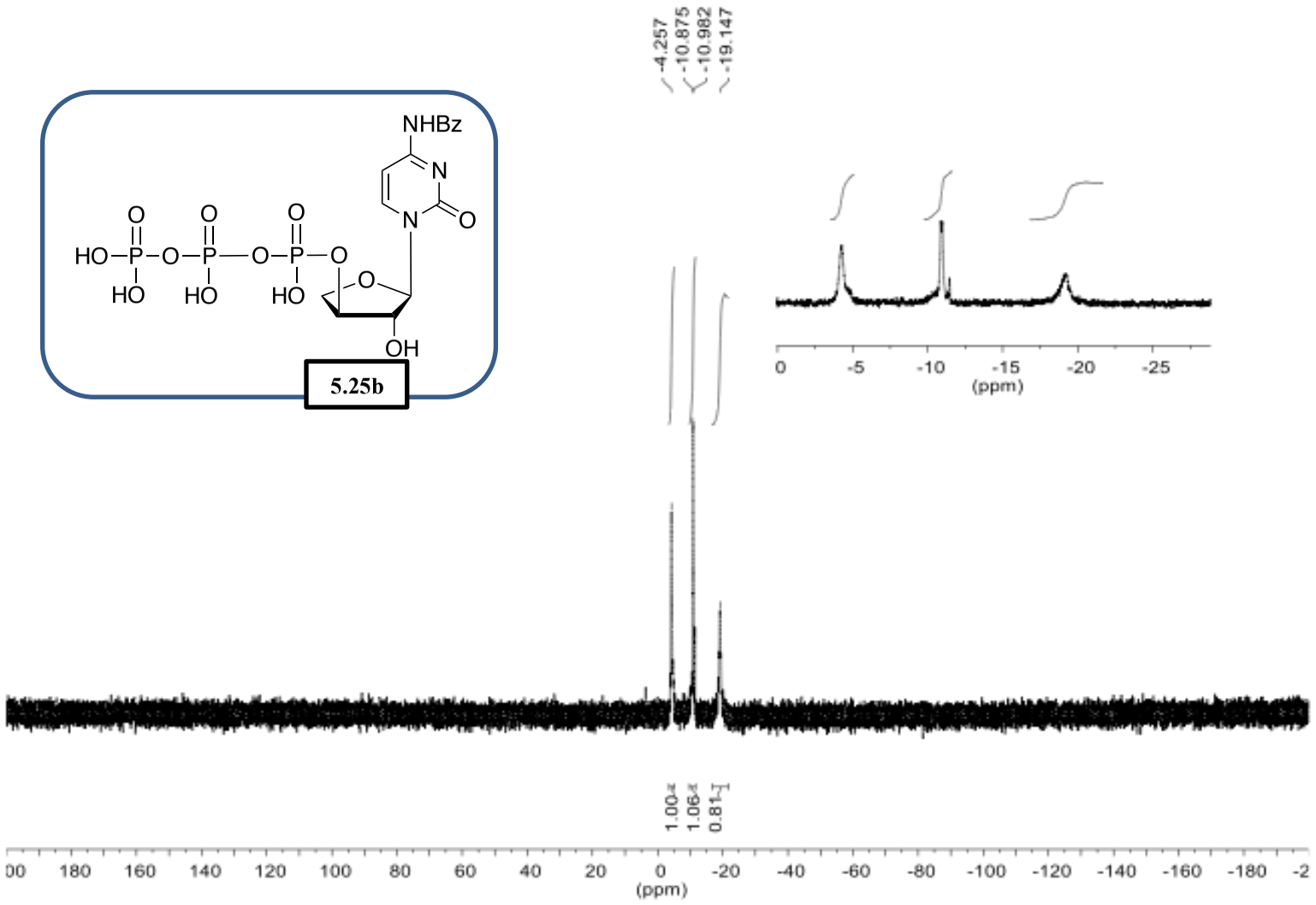
^{1}H NMR spectrum of compound **5.24b** (500 MHz,D_2O)



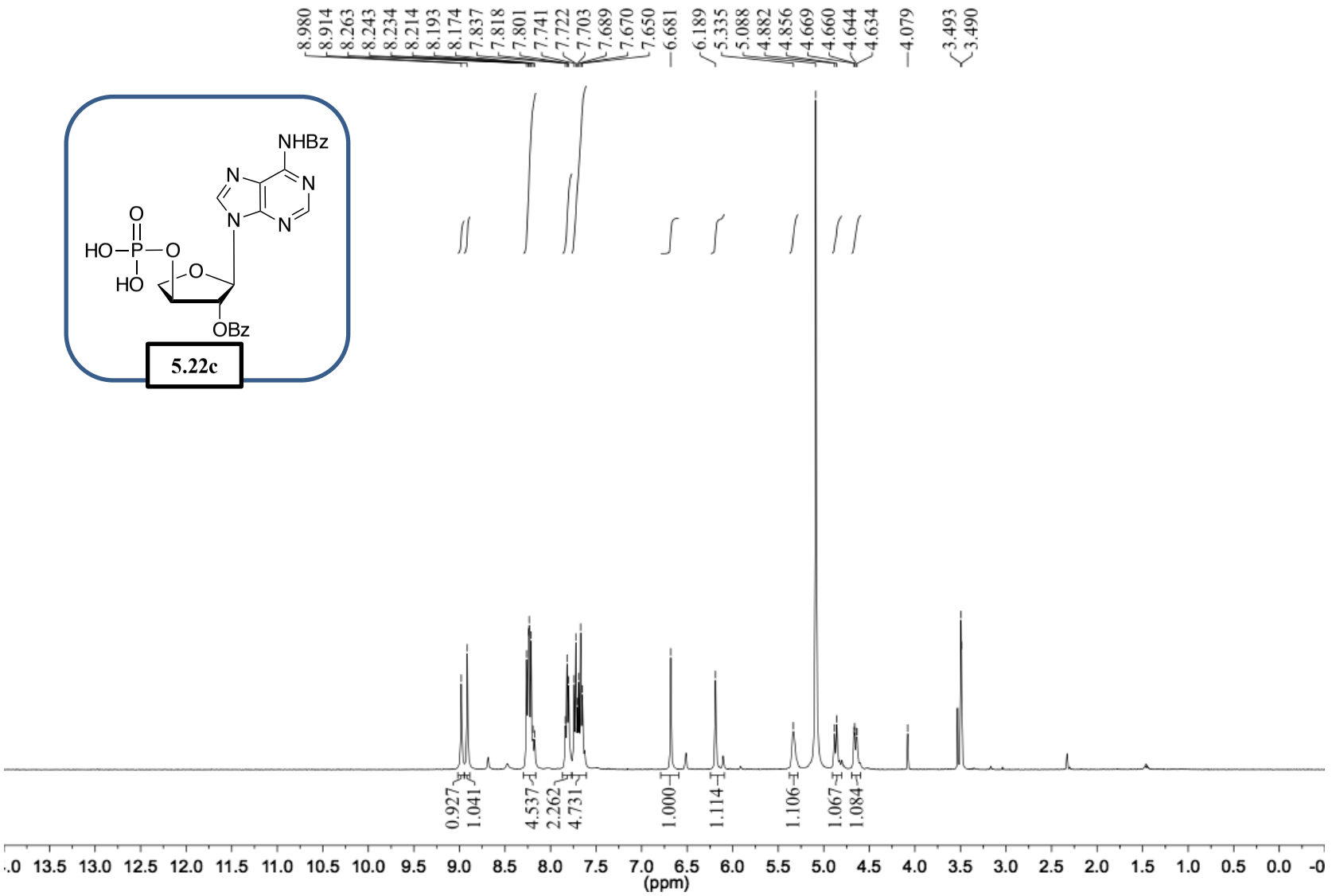
^{31}P NMR spectrum of compound **5.24b** (500 MHz,D_2O)



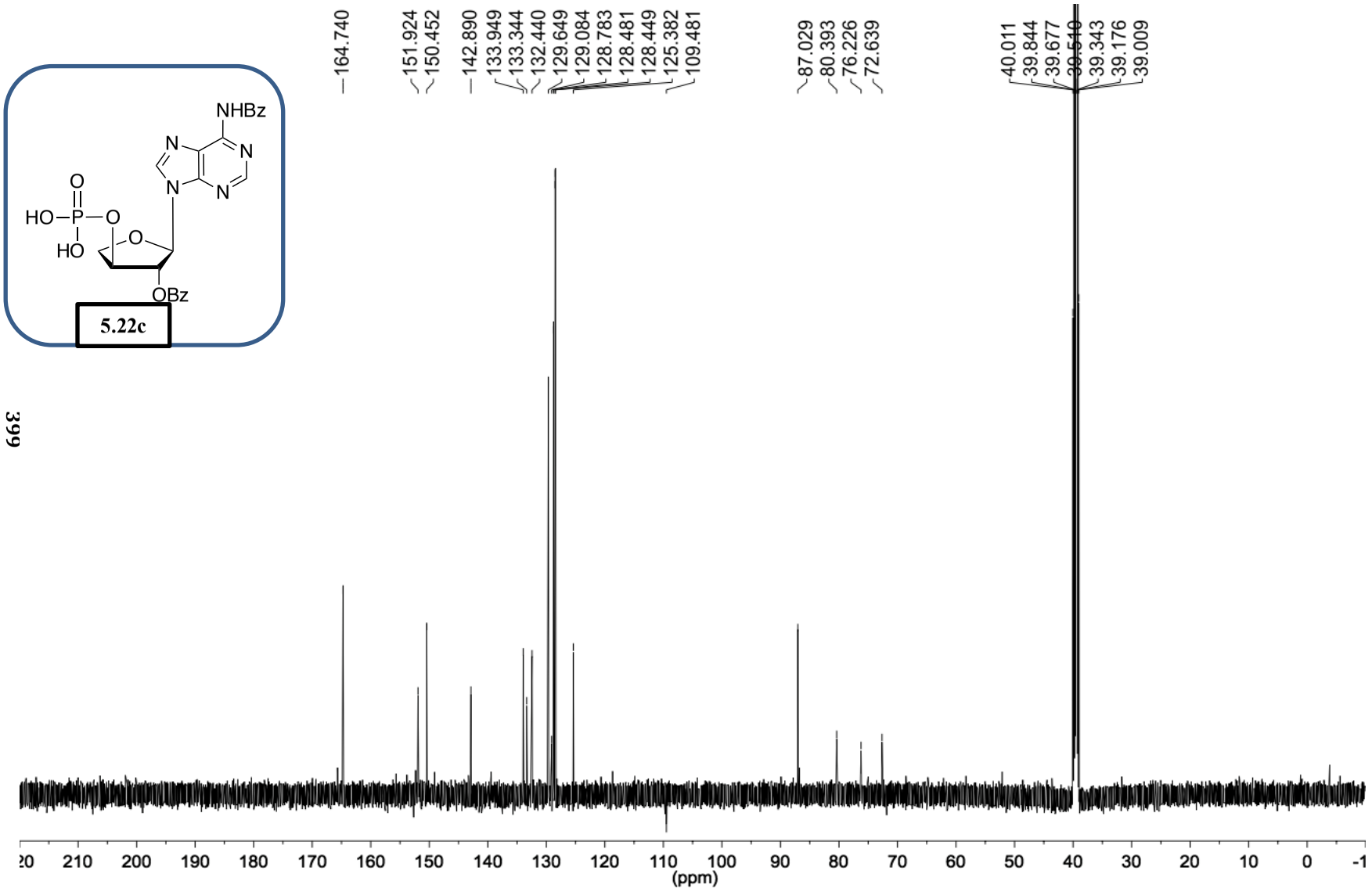
^{1}H NMR spectrum of compound **5.25b** (500 MHz,D_2O)



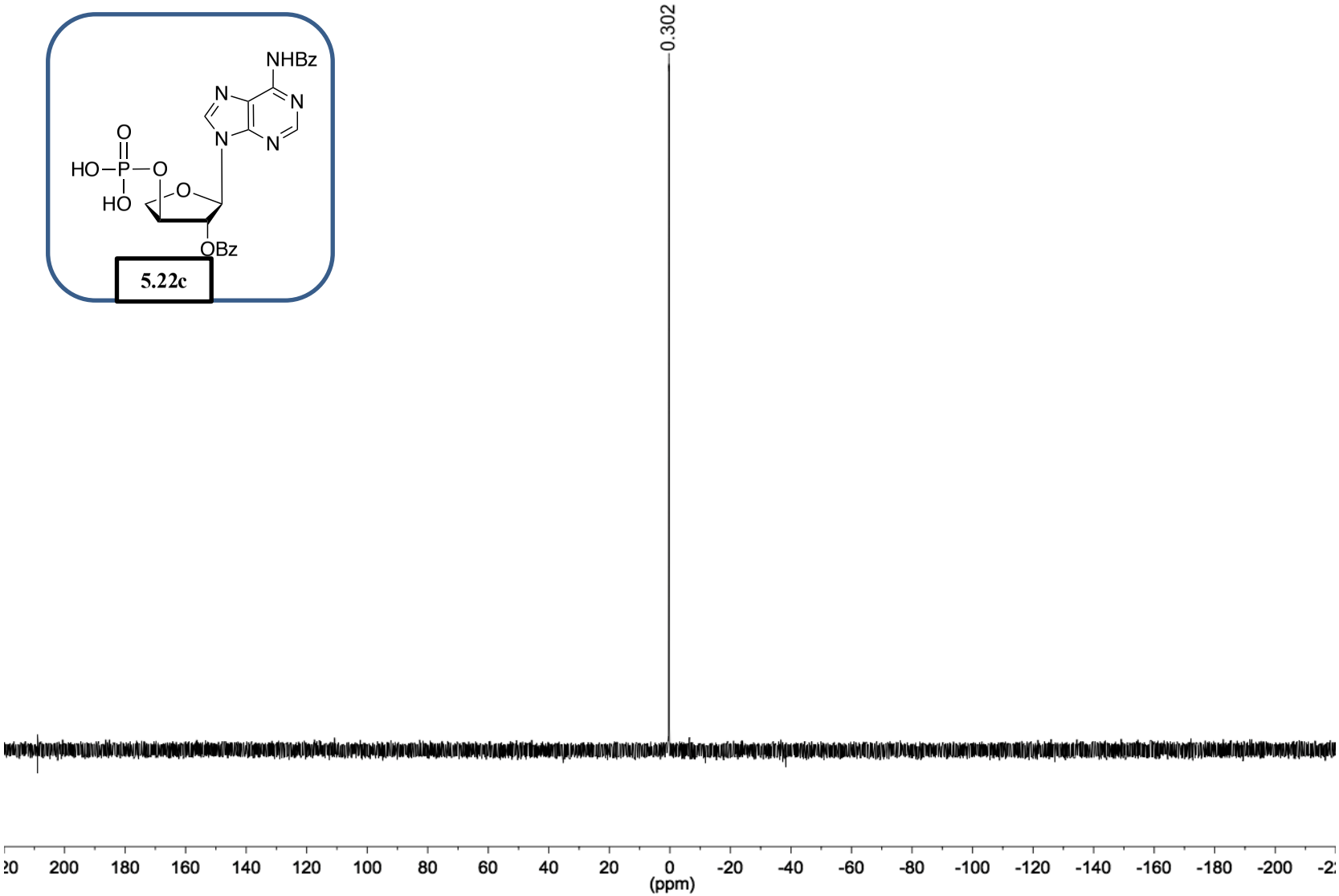
^{31}P NMR spectrum of compound **5.25b** (500 MHz,D_2O)



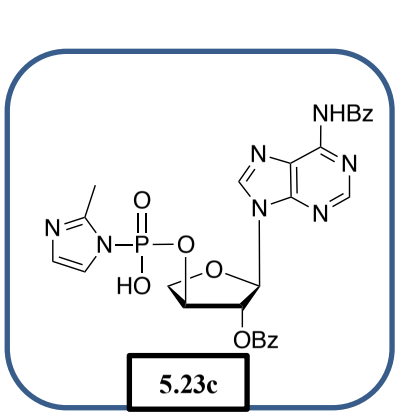
^{1}H NMR spectrum of compound **5.22c** (500 MHz,CD_3OD)



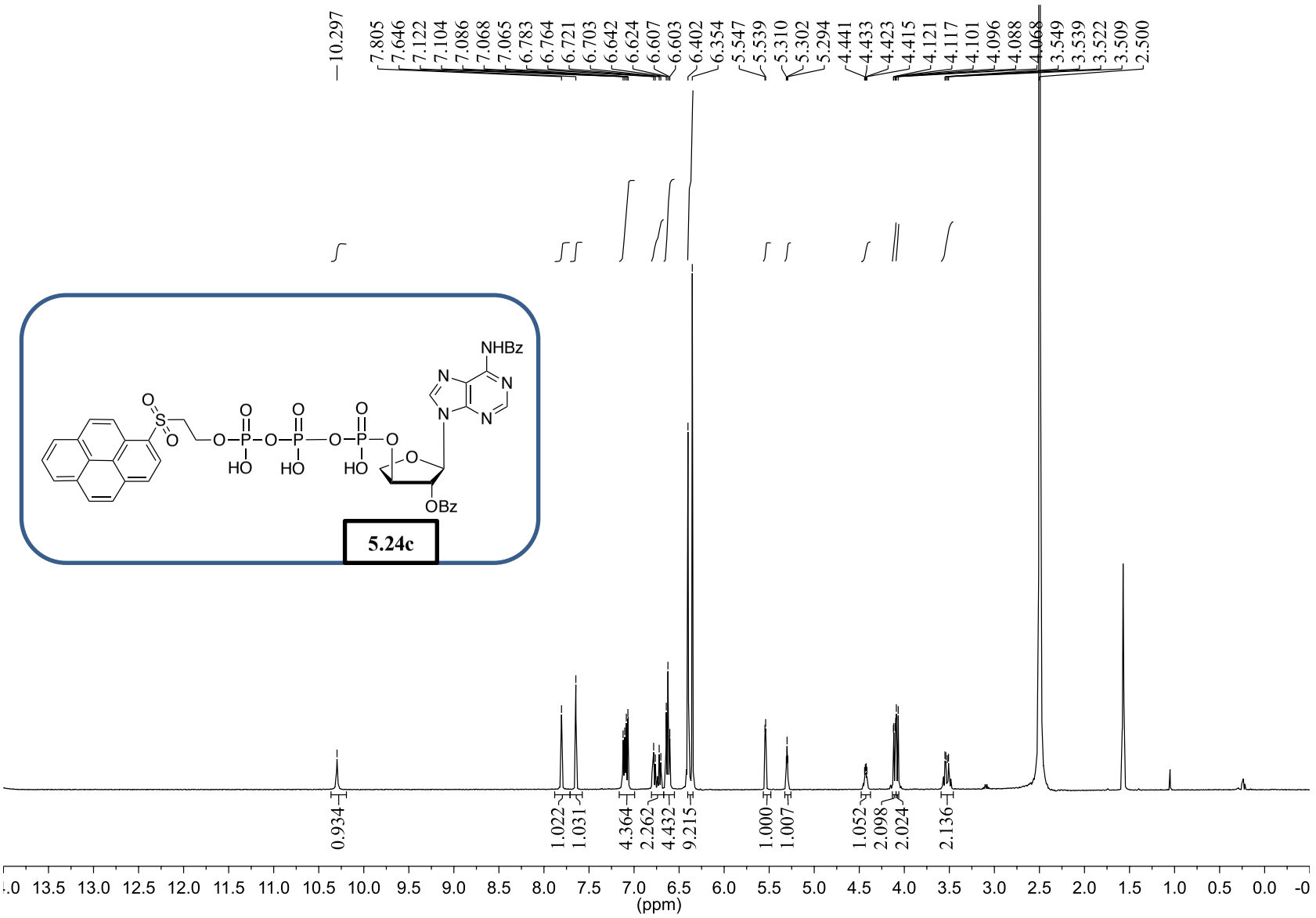
^{13}C NMR spectrum of compound **5.22c** (500 MHz,CD_3OD)



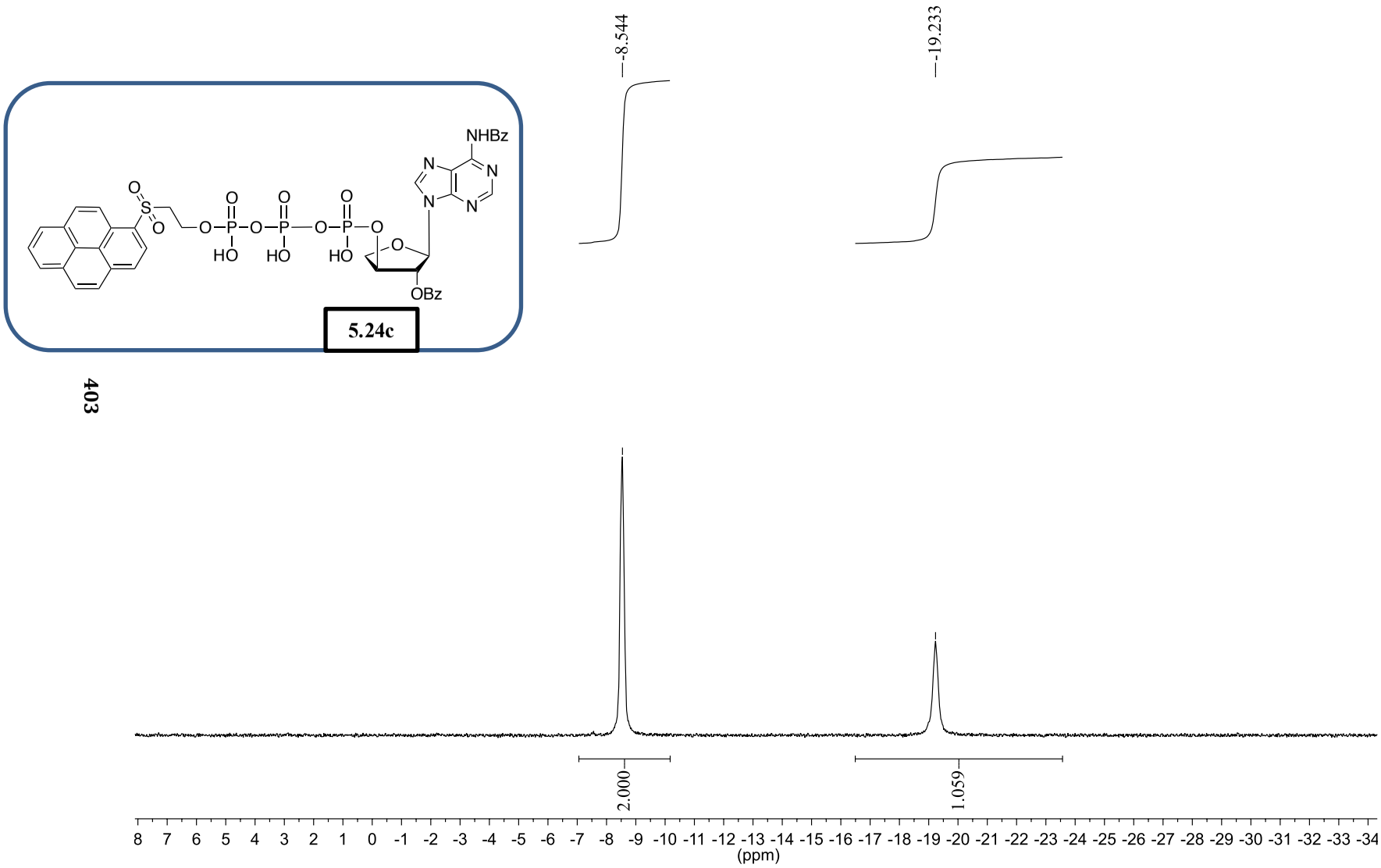
^{31}P NMR spectrum of compound **5.22c** (500 MHz,CD_3OD)



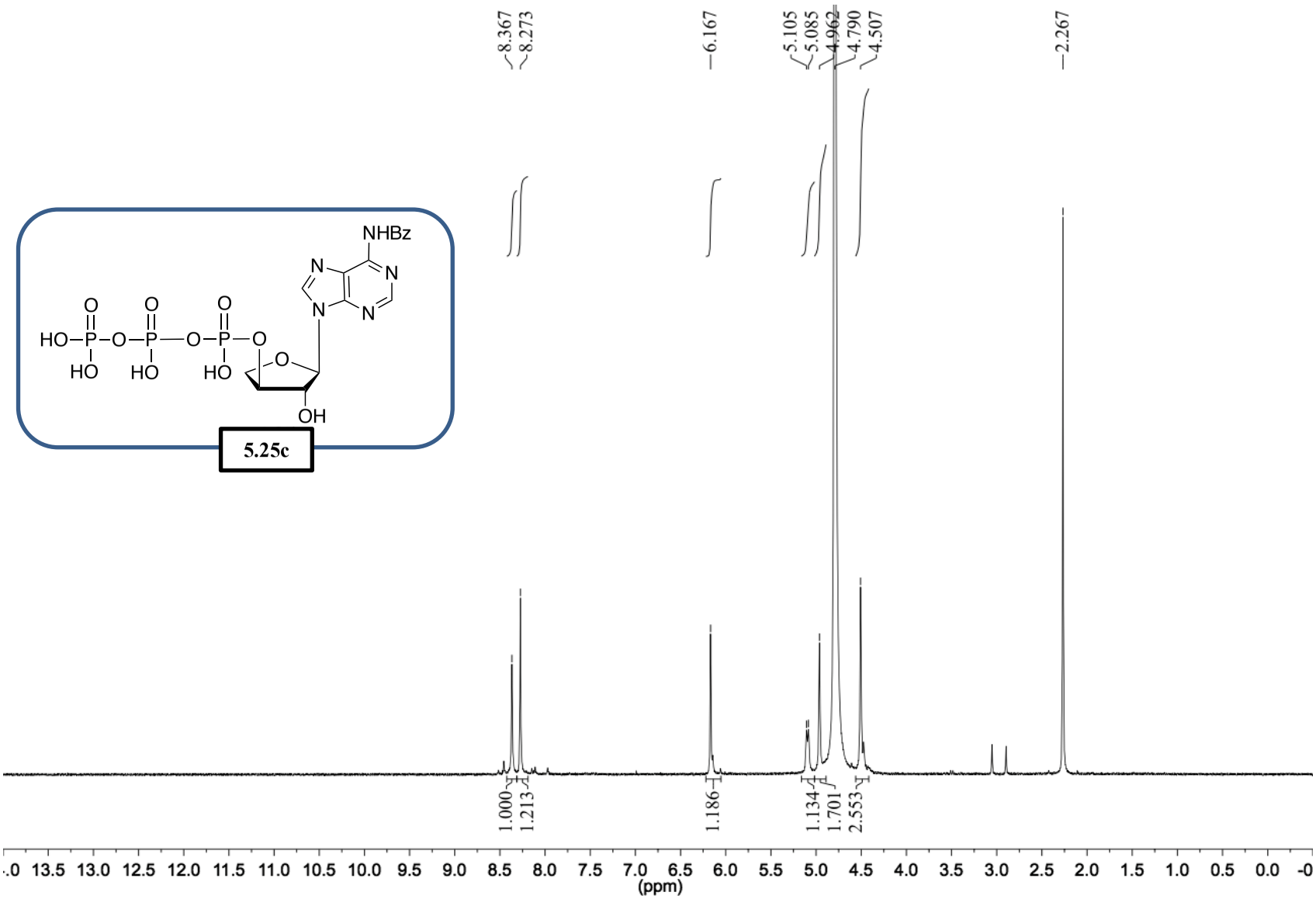
^{31}P NMR spectrum of compound **5.23c** (500 MHz,CD_3OD)



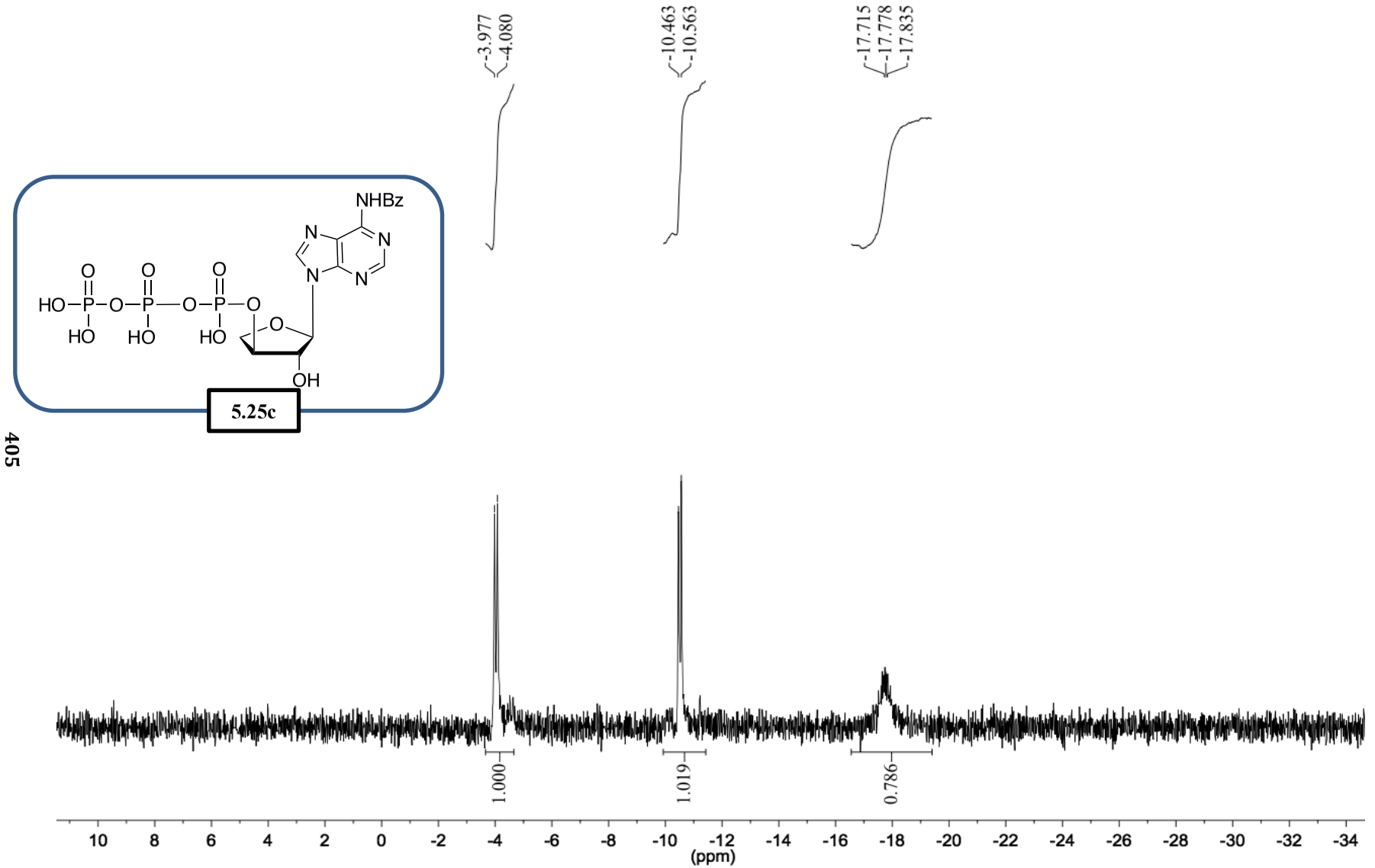
^{1}H NMR spectrum of compound **5.24c** (500 MHz,D_2O)



^{31}P NMR spectrum of compound **5.24c** (500 MHz,D_2O)

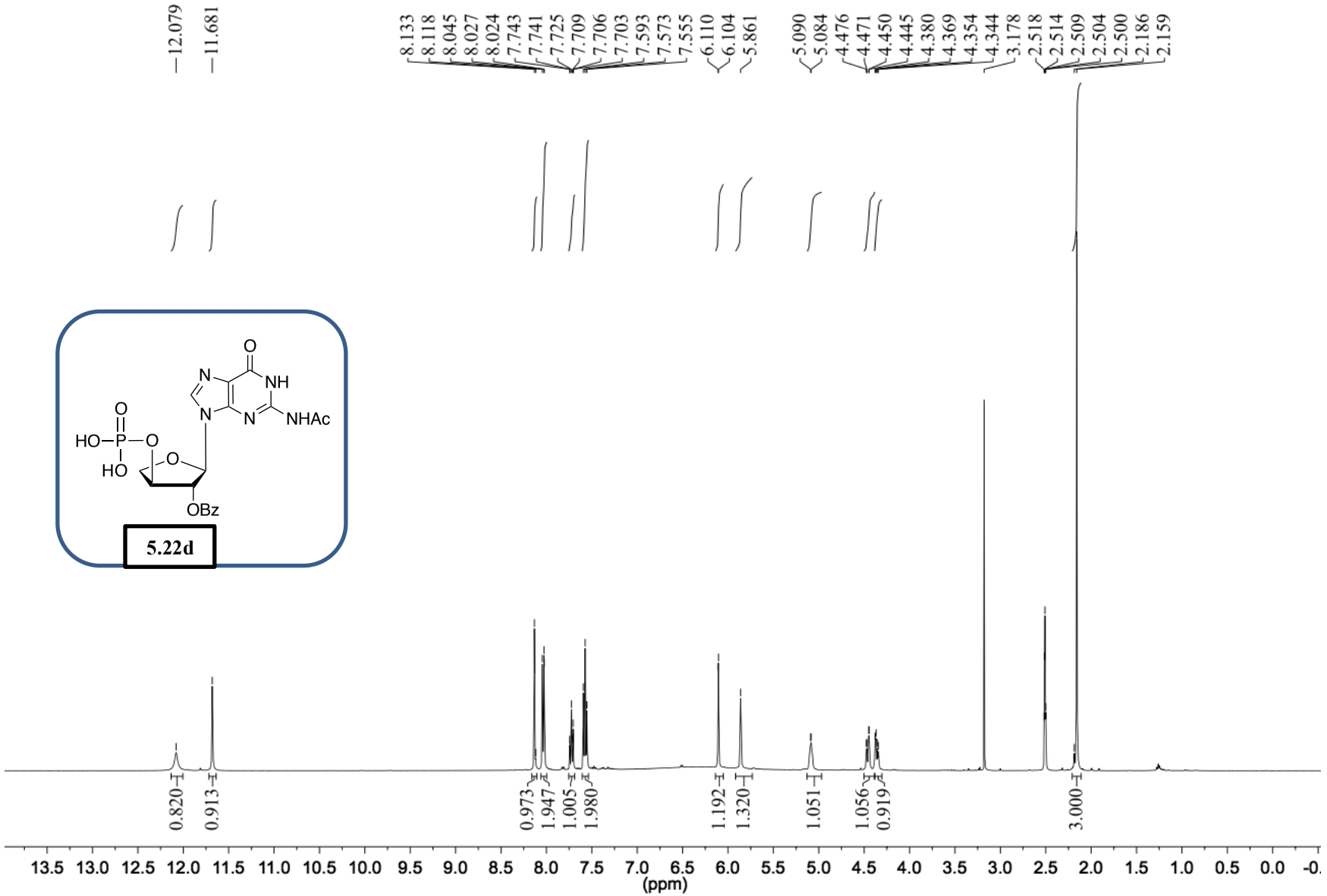


^{1}H NMR spectrum of compound **4.25c** (500 MHz,D_2O)

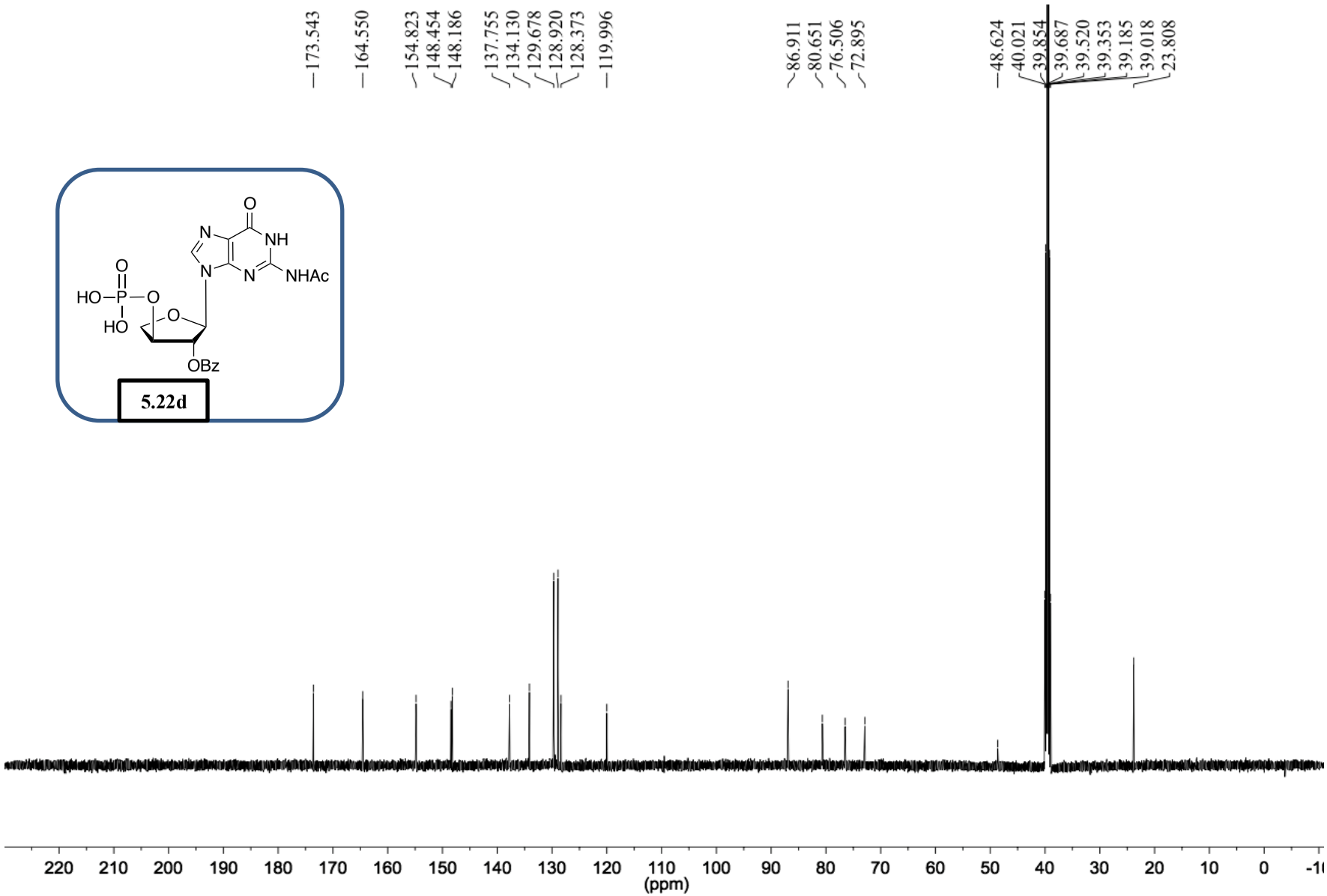


405

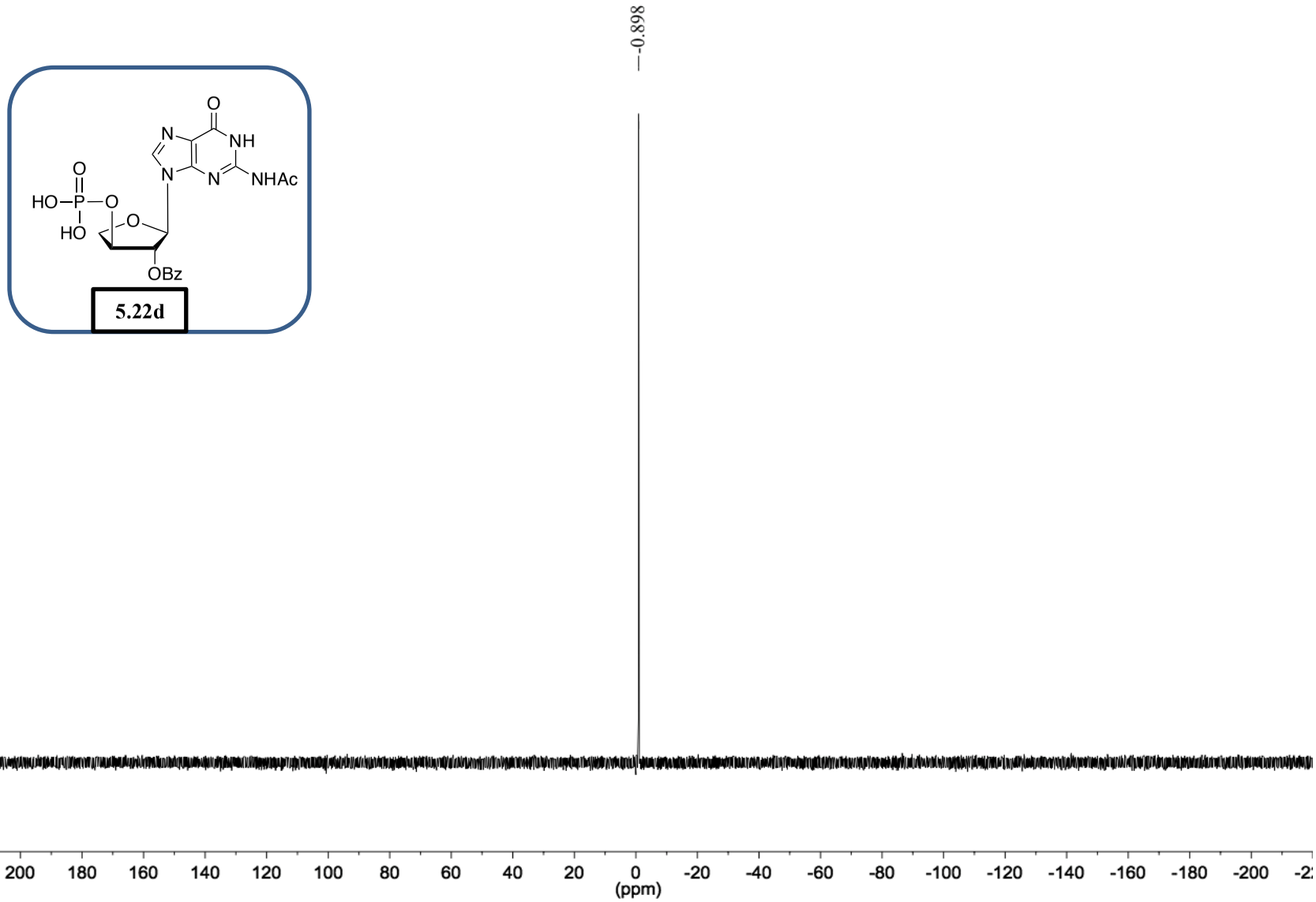
^{31}P NMR spectrum of compound **5.25c** (500 MHz,D_2O)



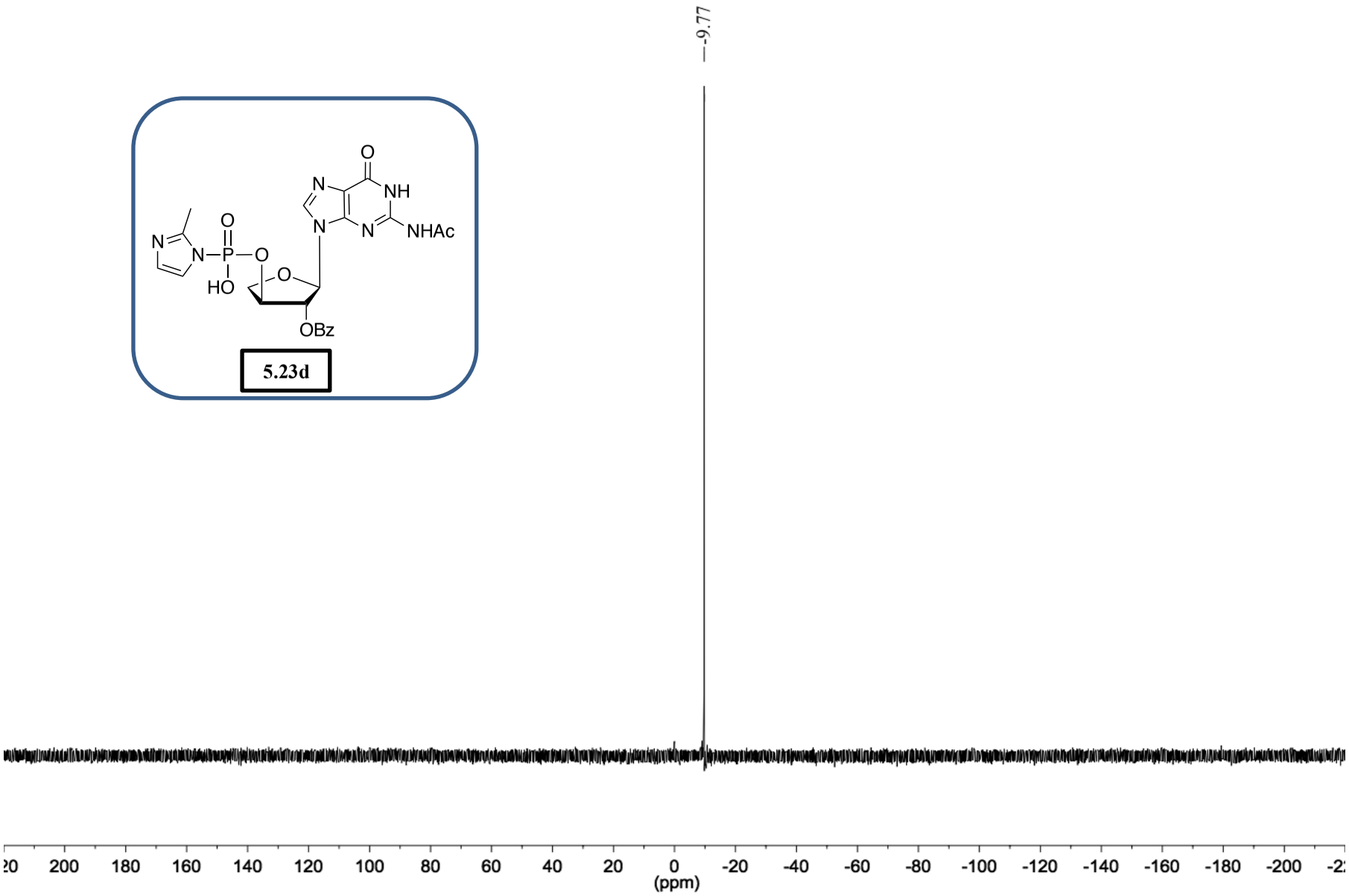
^{1}H NMR spectrum of compound **5.22d** (500 MHz,CD_3OD)



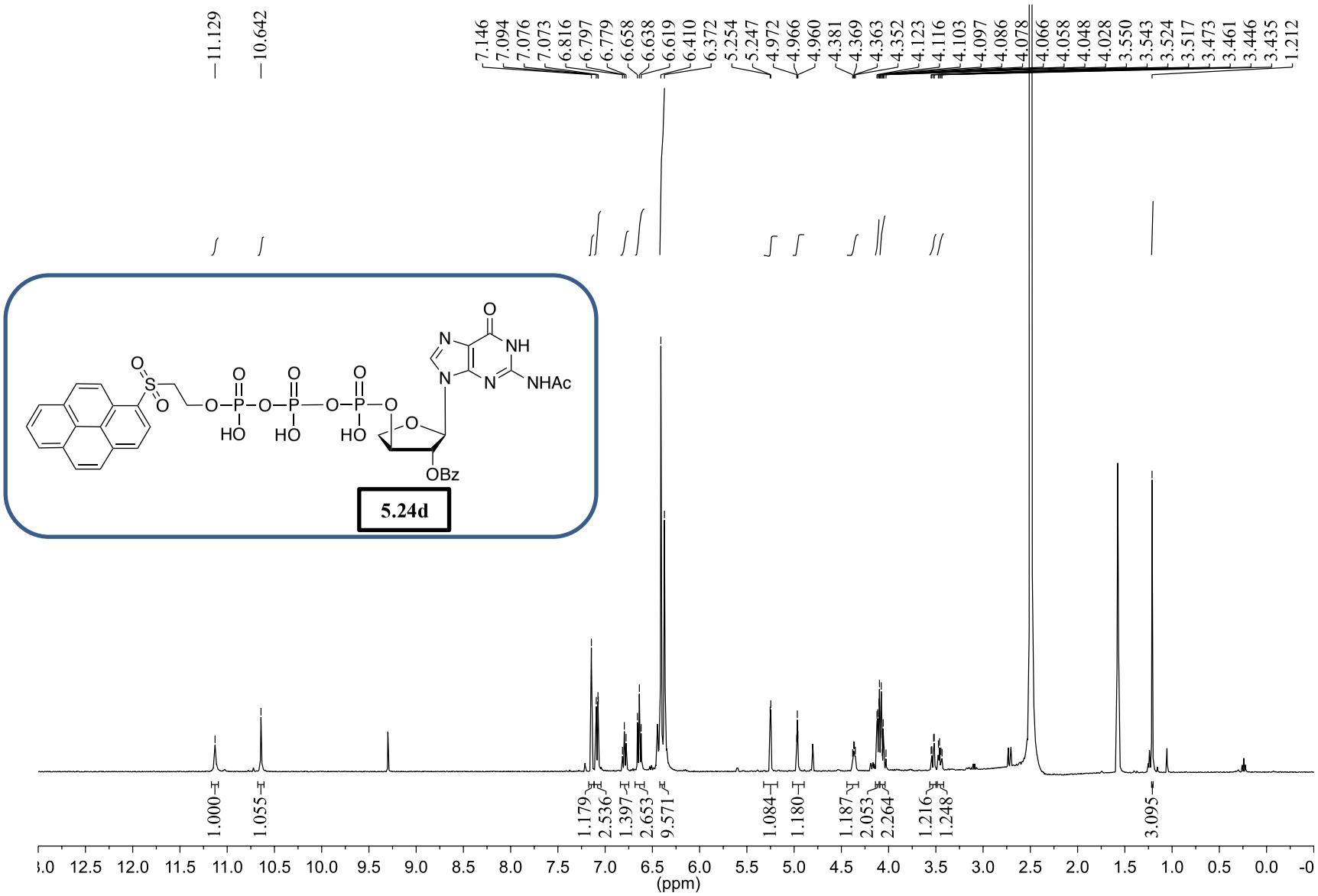
^{13}C NMR spectrum of compound **5.22d** (500 MHz,CD_3OD)



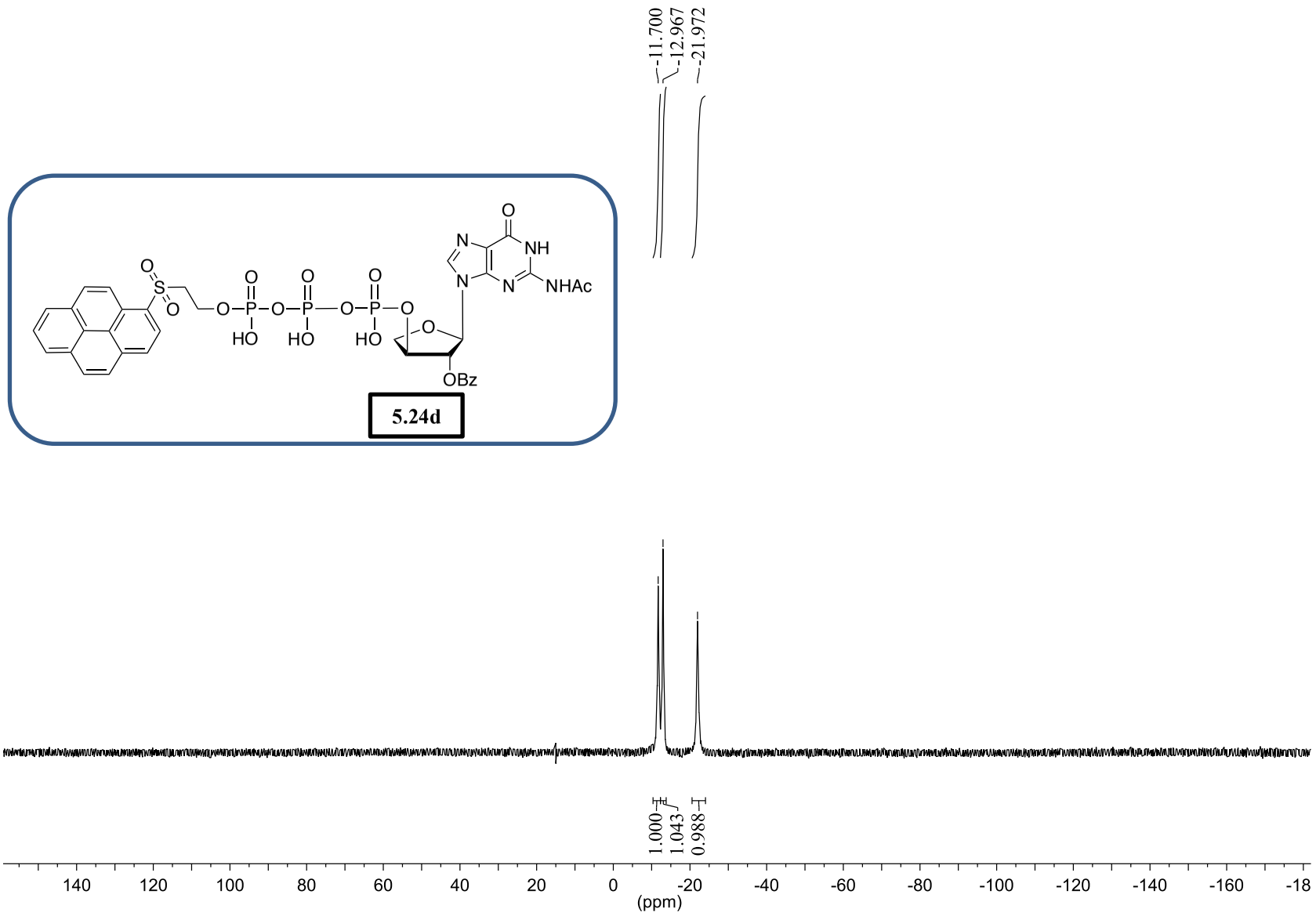
^{31}P NMR spectrum of compound **5.22d** (500 MHz,CD_3OD)



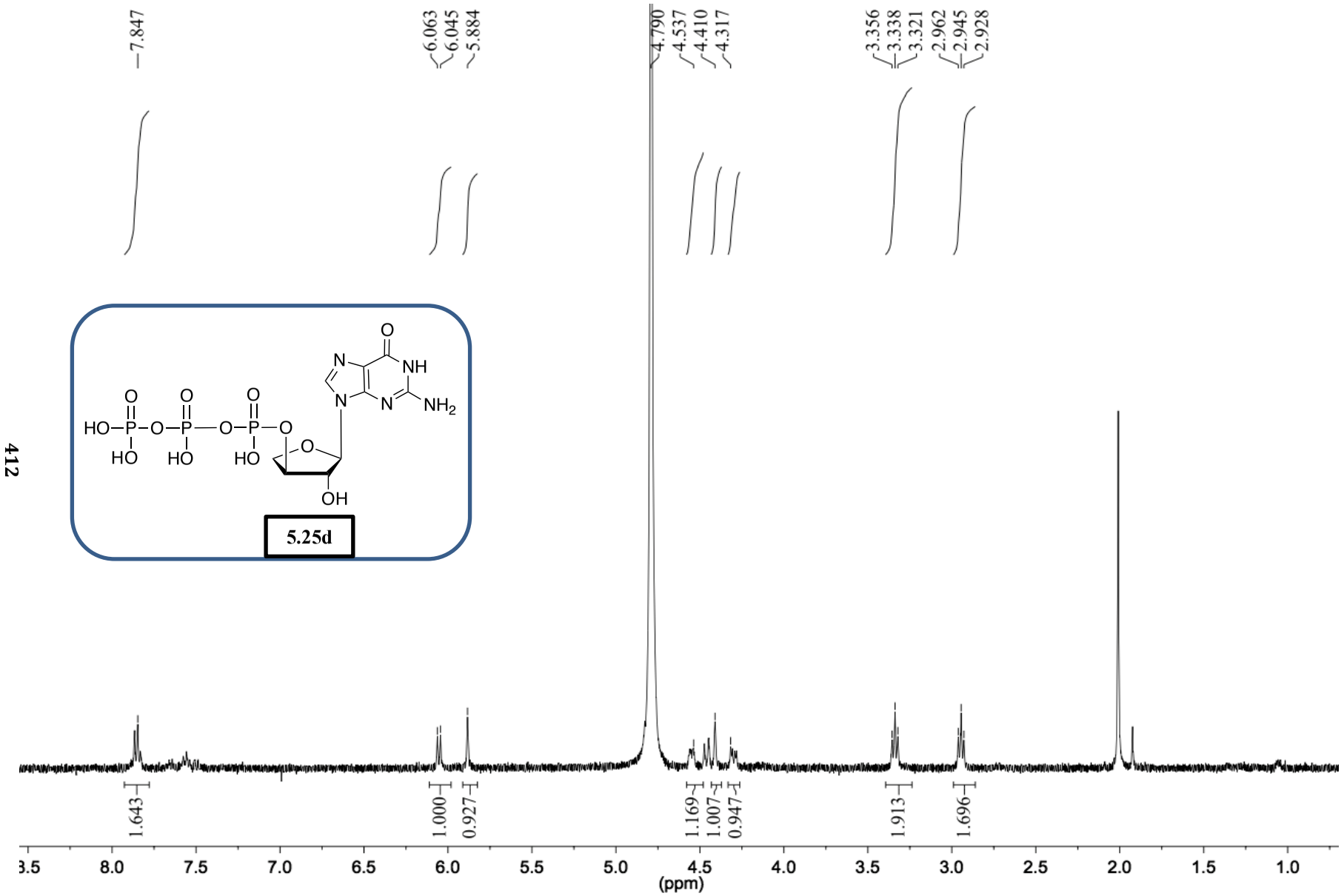
^{31}P NMR spectrum of compound **5.23d** (500 MHz,CD_3OD)



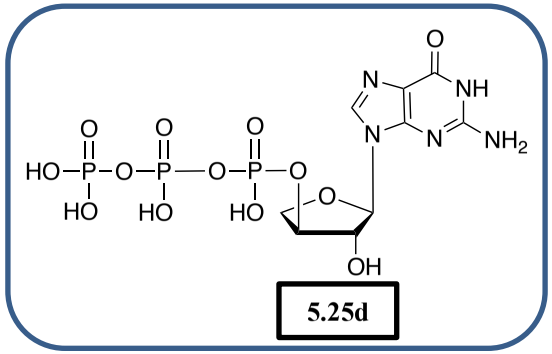
^{1}H NMR spectrum of compound **5.24d** (500 MHz,D_2O)



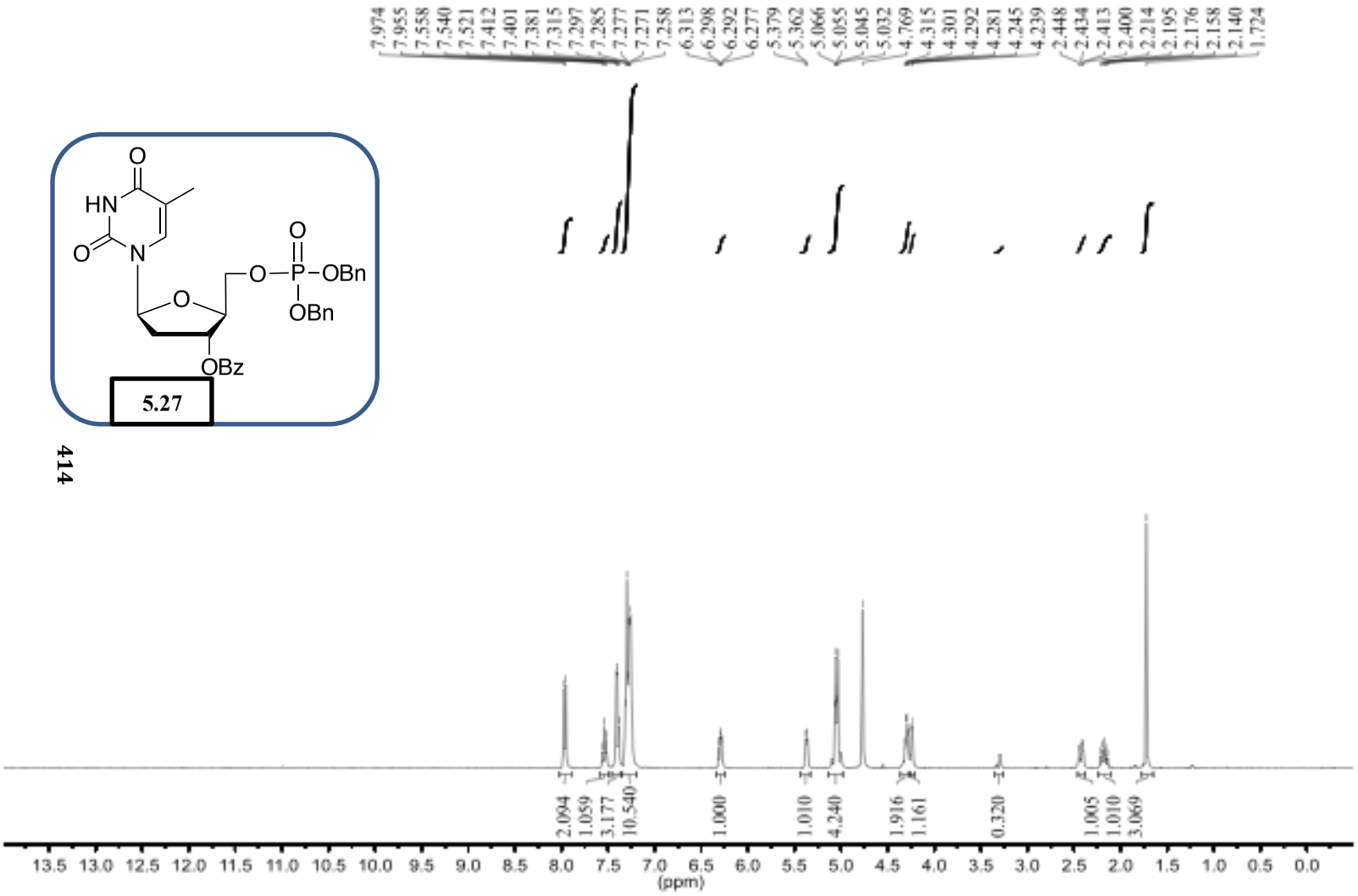
^{31}P NMR spectrum of compound **5.24d** (500 MHz,D_2O)



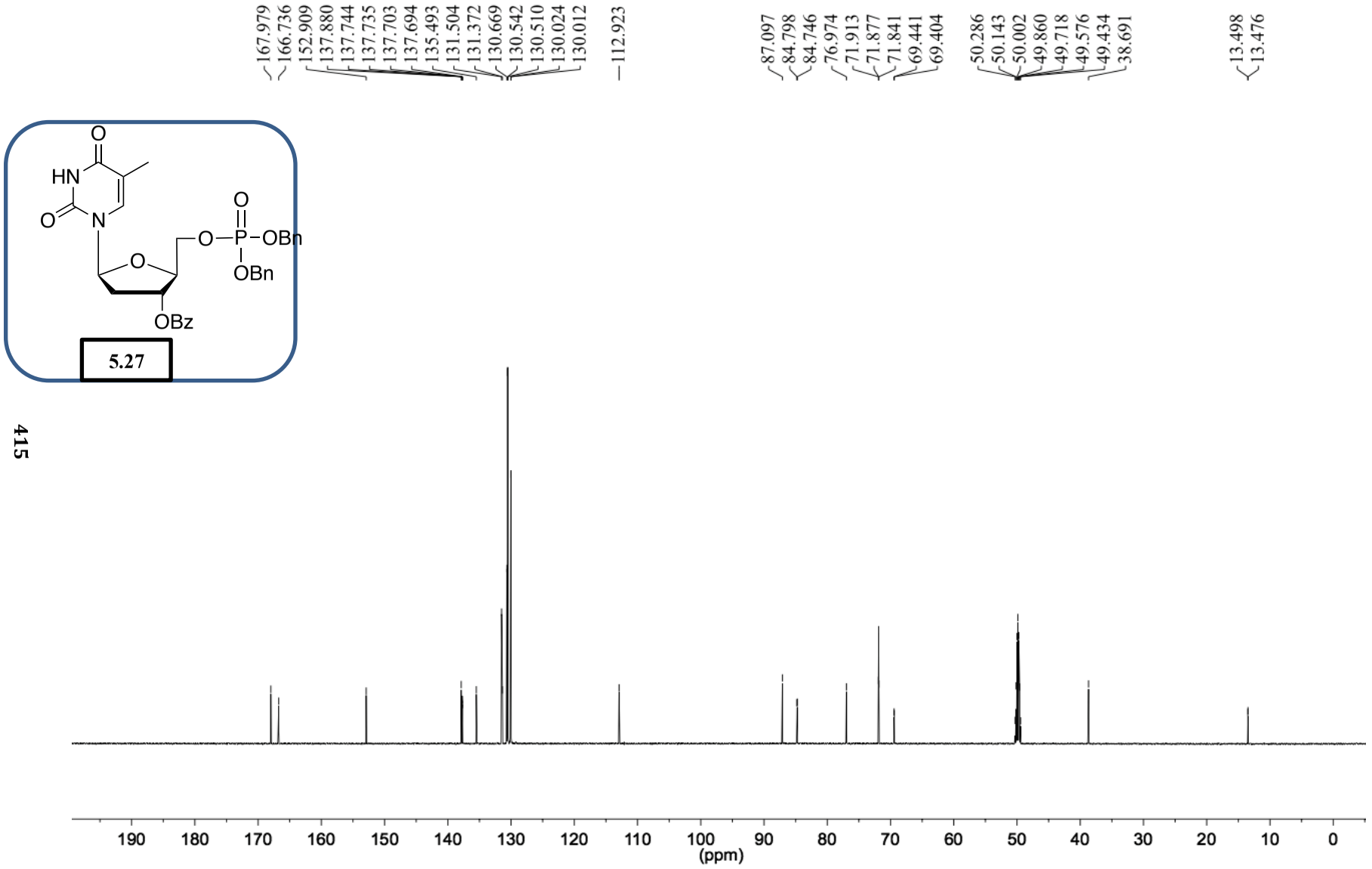
^{1}H NMR spectrum of compound **5.25d** (500 MHz,D_2O)



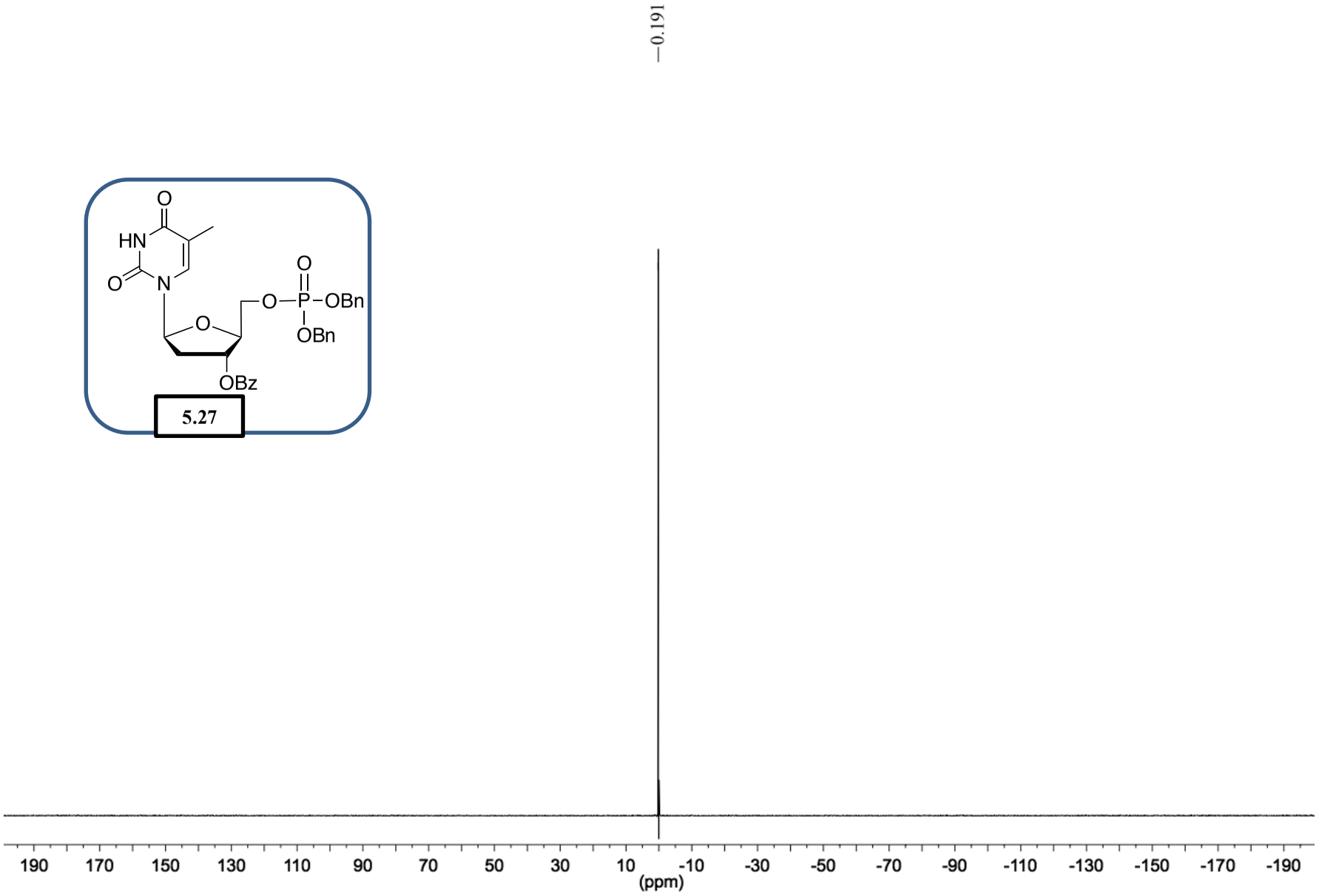
^{31}P NMR spectrum of compound **5.25d** (500 MHz,D_2O)



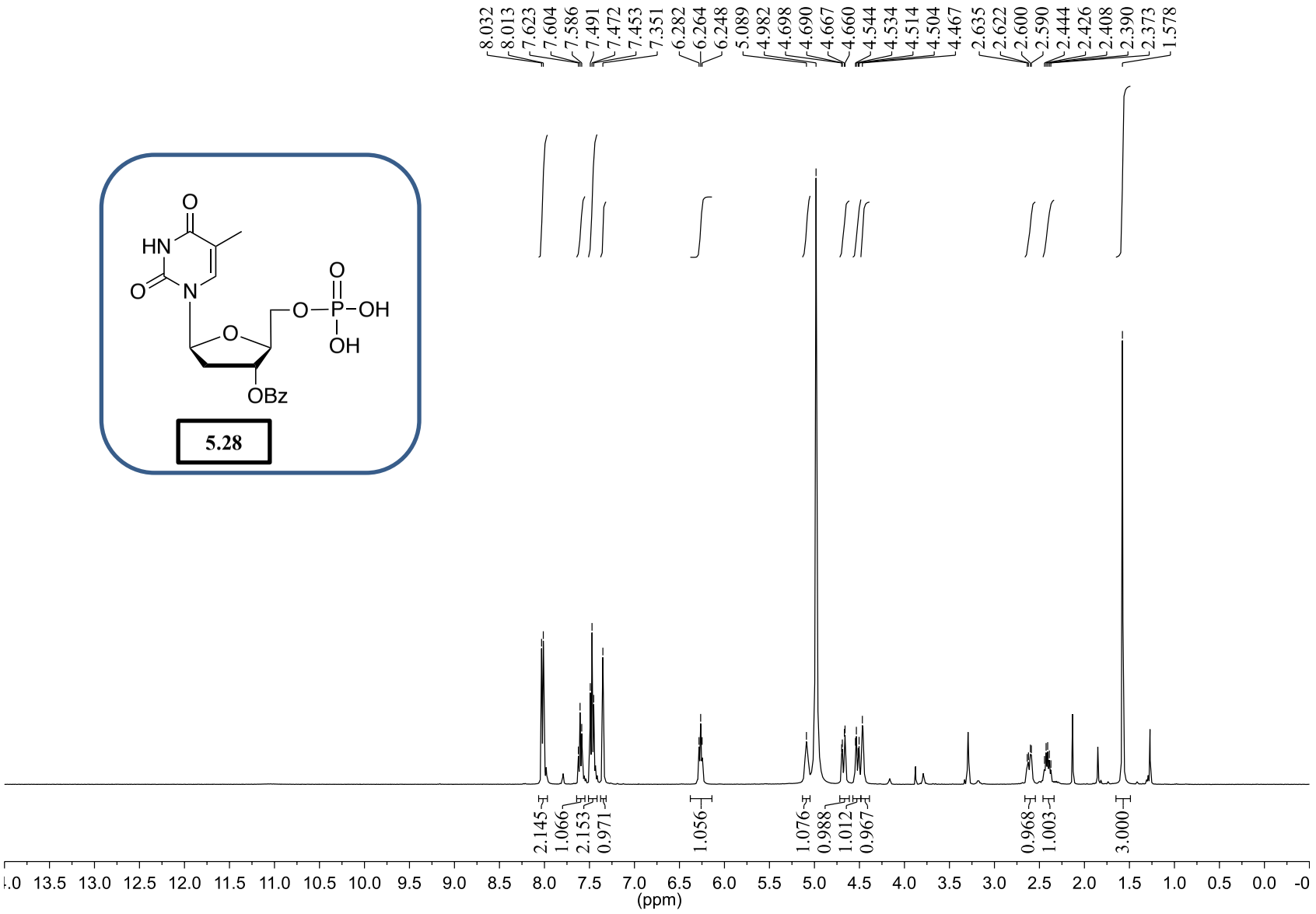
^{1}H NMR spectrum of compound **4.27** (500 MHz,$CDCl_3$)



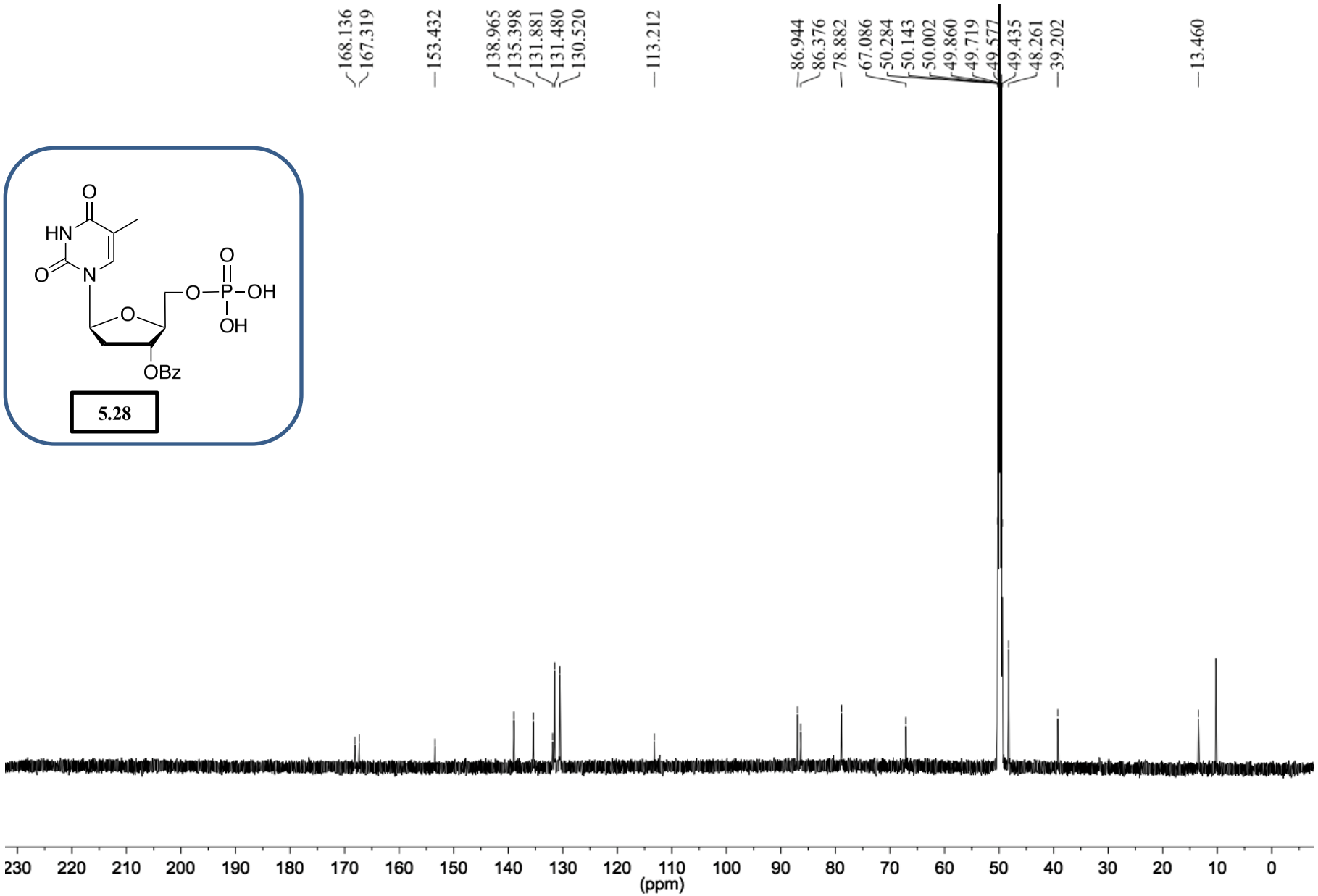
^{13}C NMR spectrum of compound **5.27** (500 MHz,$CDCl_3$)



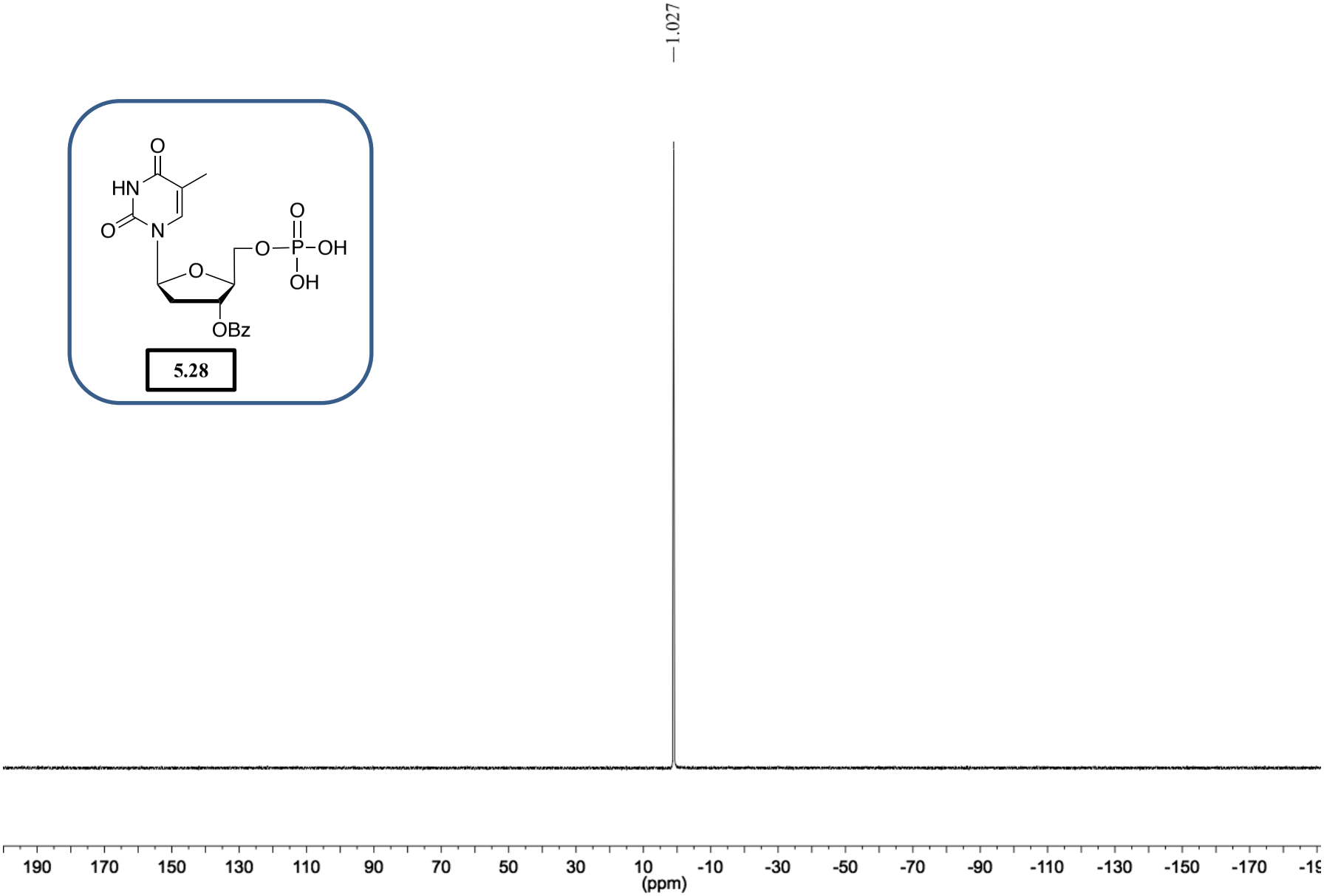
^{31}P NMR spectrum of compound **4.27** (500 MHz,$CDCl_3$)



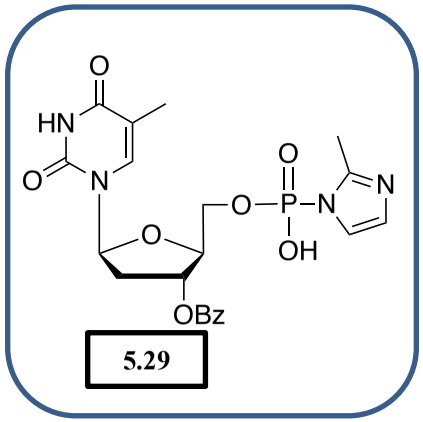
^{1}H NMR spectrum of compound **5.28** (500 MHz,CD_3OD)



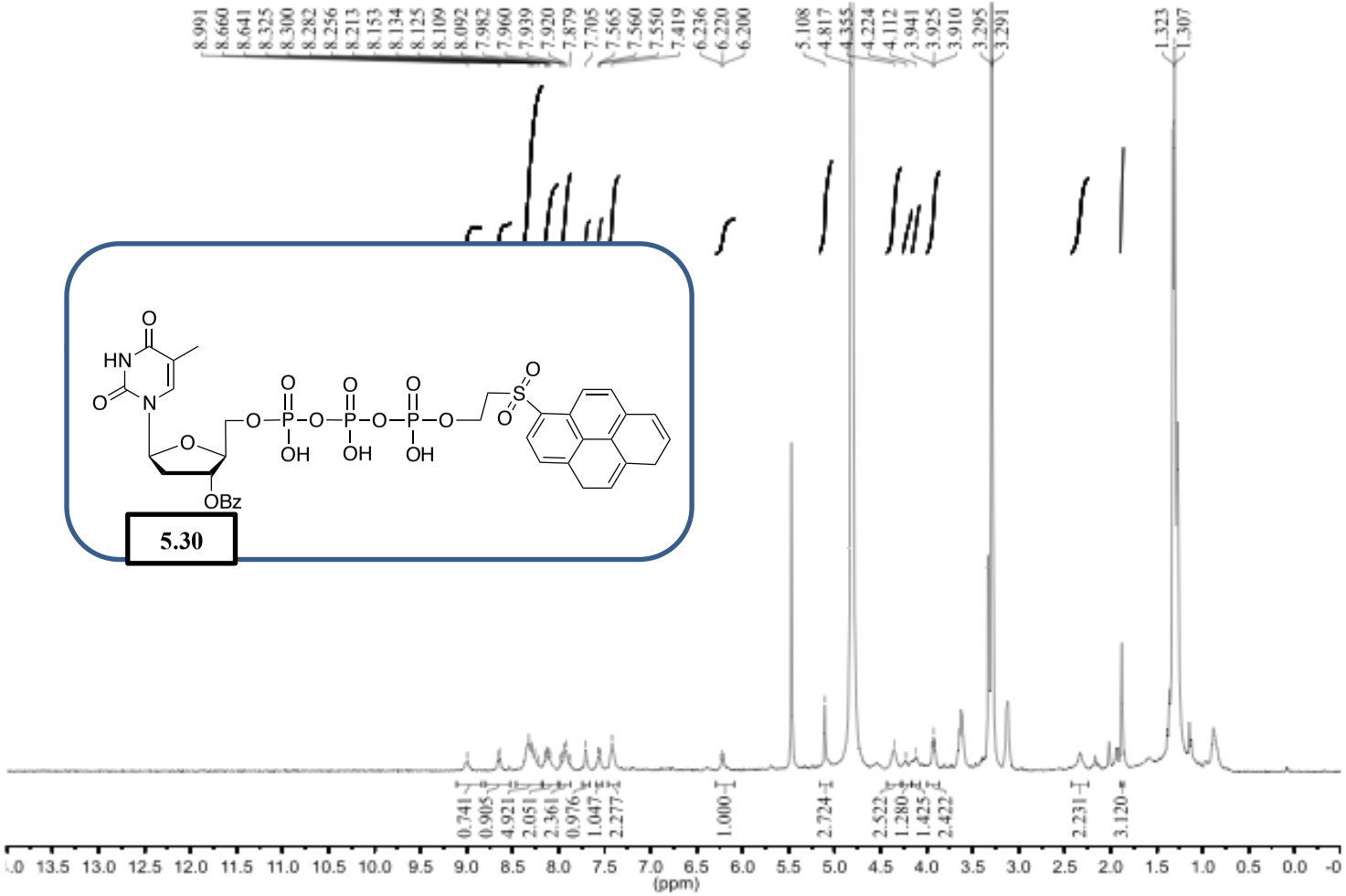
^{13}C NMR spectrum of compound **5.28** (500 MHz,CD_3OD)



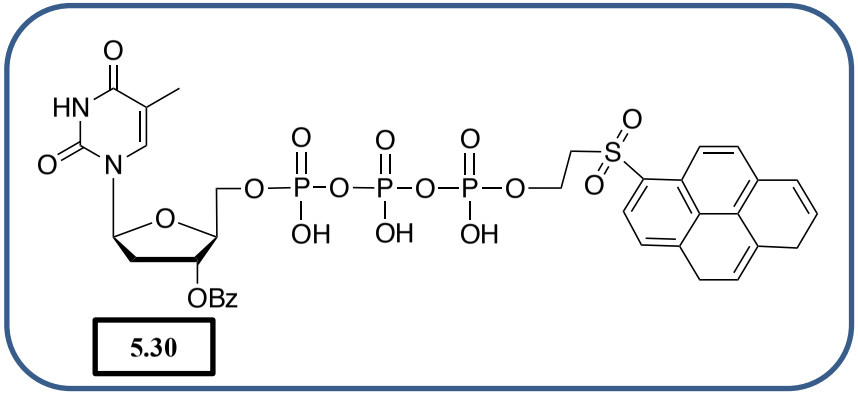
^{31}P NMR spectrum of compound **5.28** (500 MHz,CD_3OD)



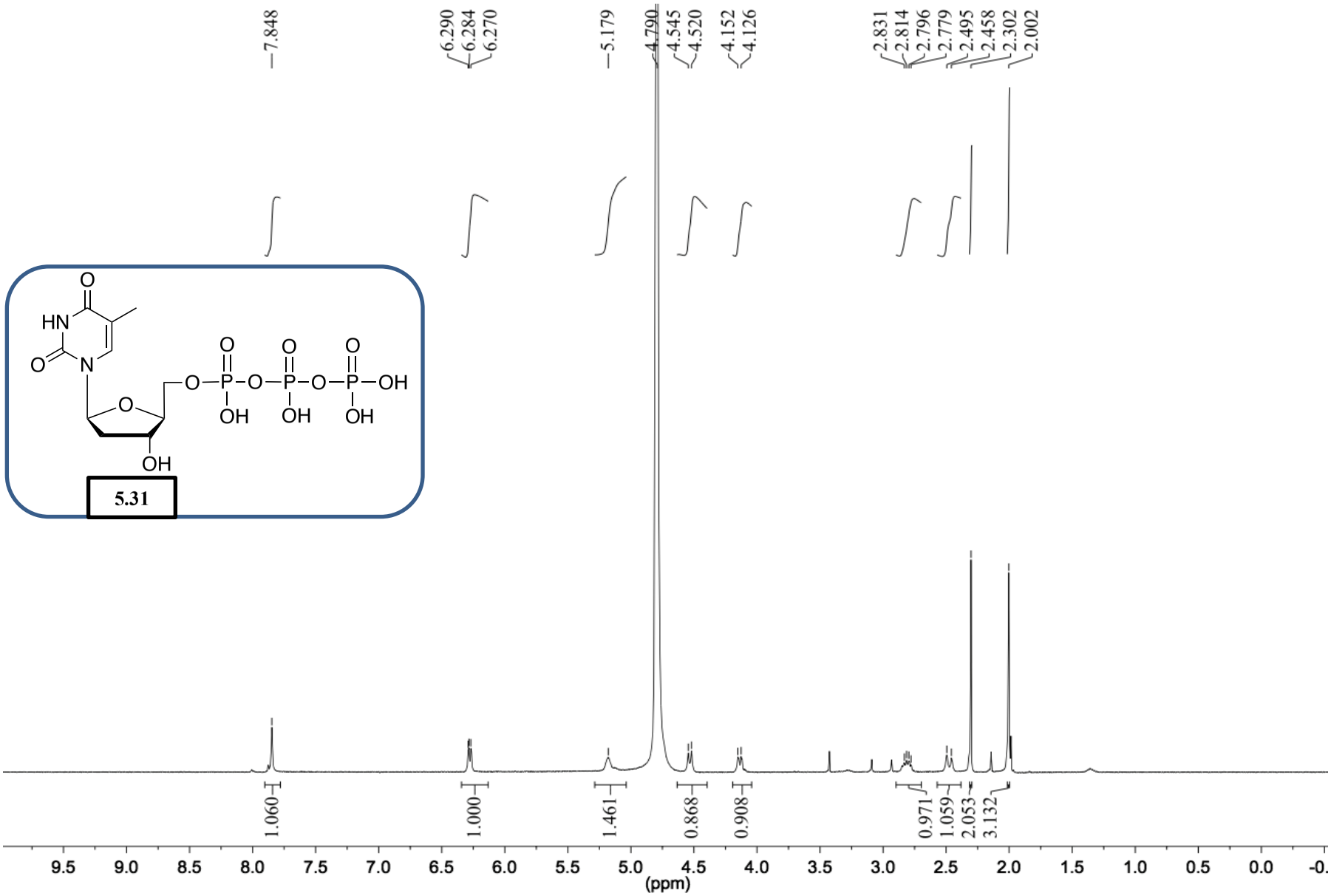
^{31}P NMR spectrum of compound **5.29** (500 MHz,D_2O)



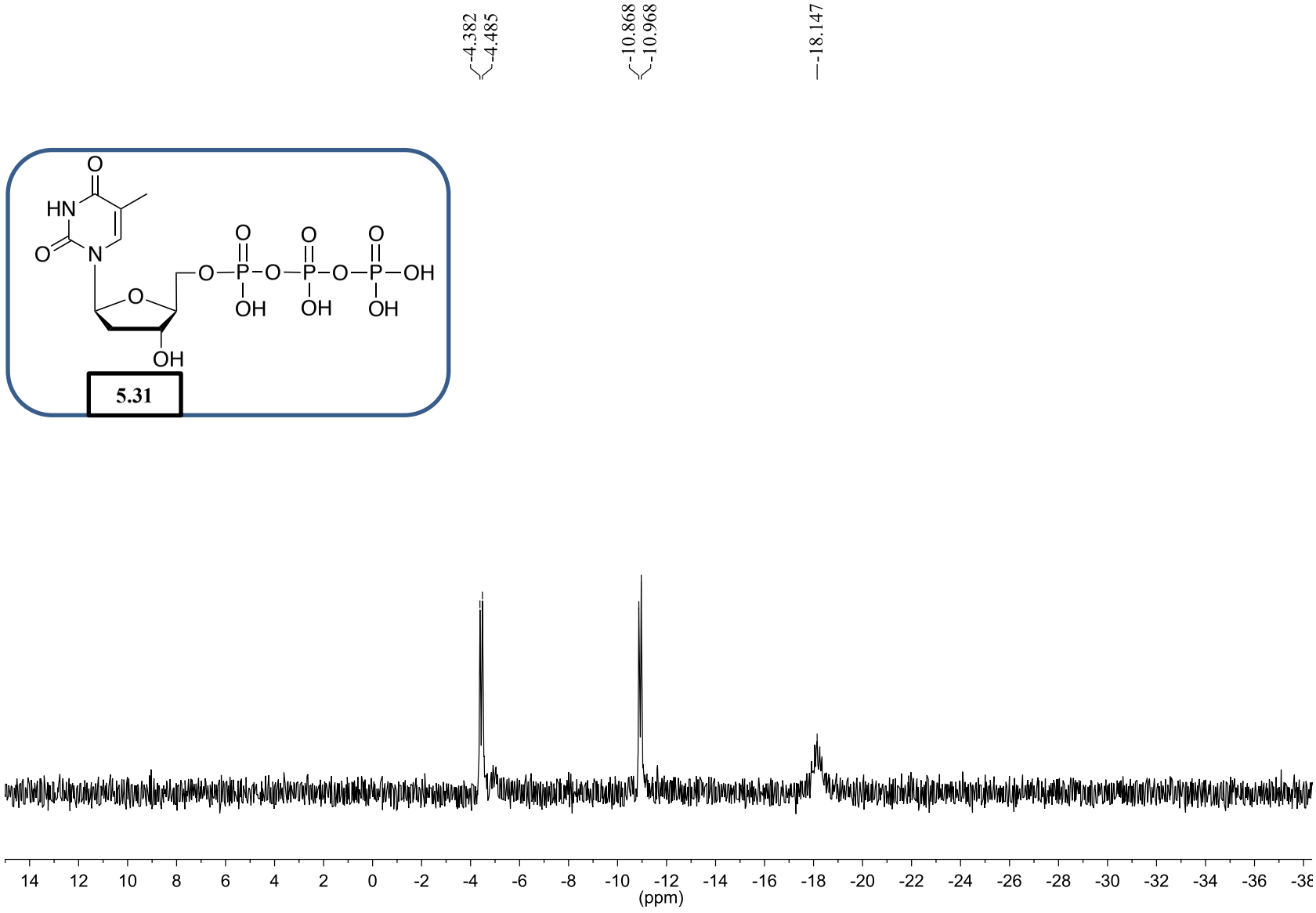
^{1}H NMR spectrum of compound **5.30** (500 MHz,D_2O)



^{31}P NMR spectrum of compound **5.30** (500 MHz,D_2O)



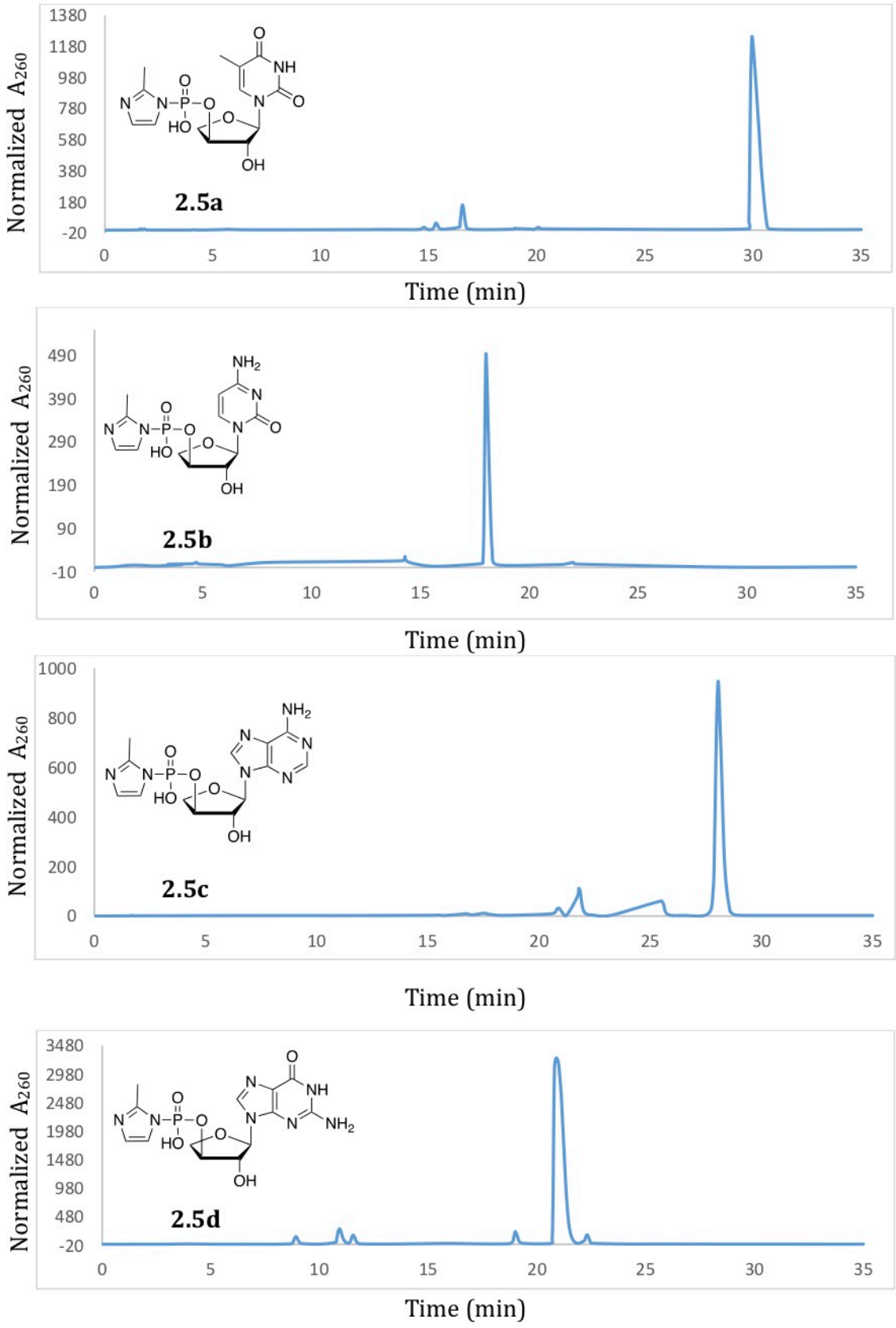
^{1}H NMR spectrum of compound **5.31** (500 MHz,D_2O)



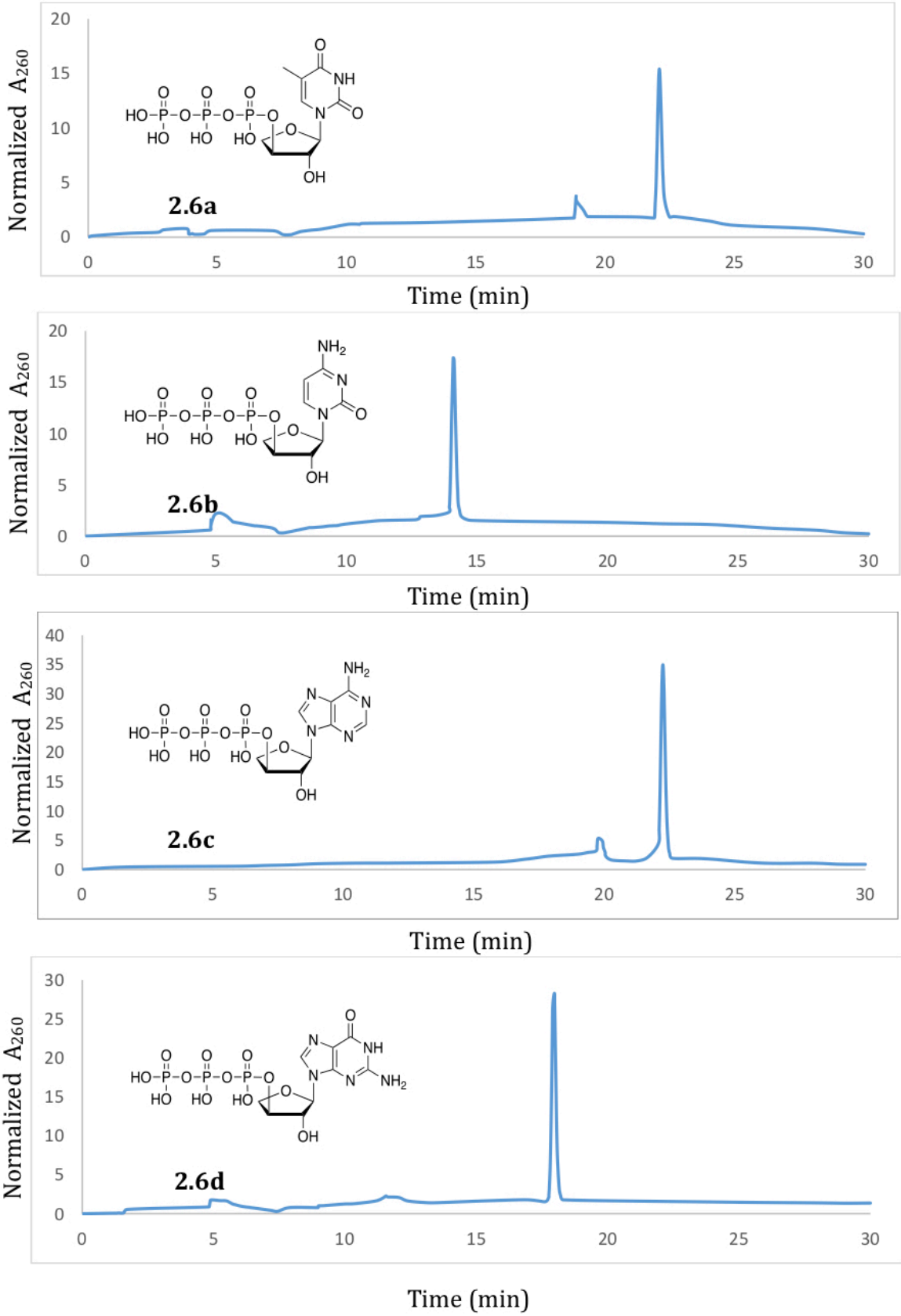
^{31}P NMR spectrum of compound **5.31** (500 MHz,D_2O)

APPENDIX B

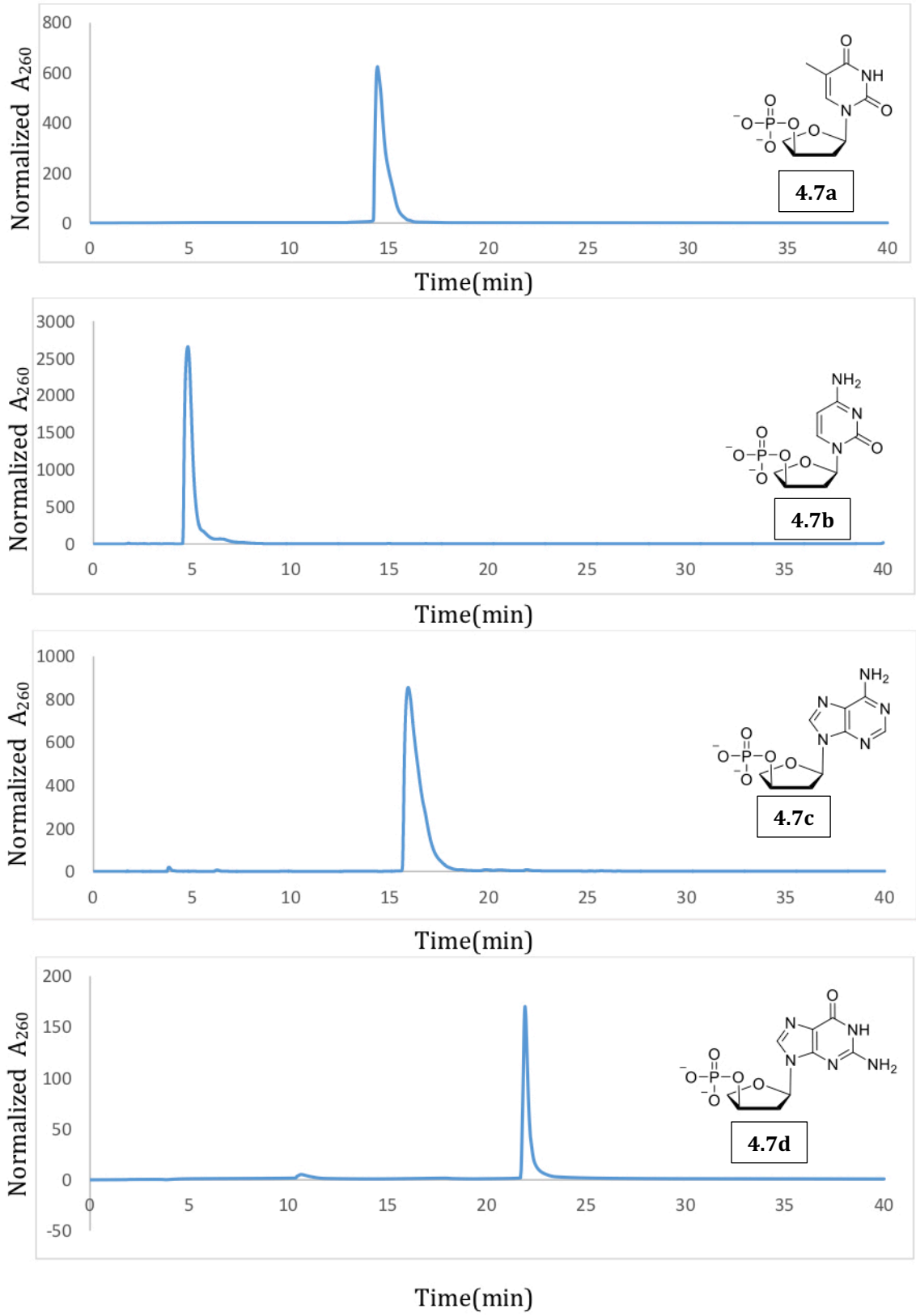
SUPPLEMENTAL HPLC CHROMATOGRAM



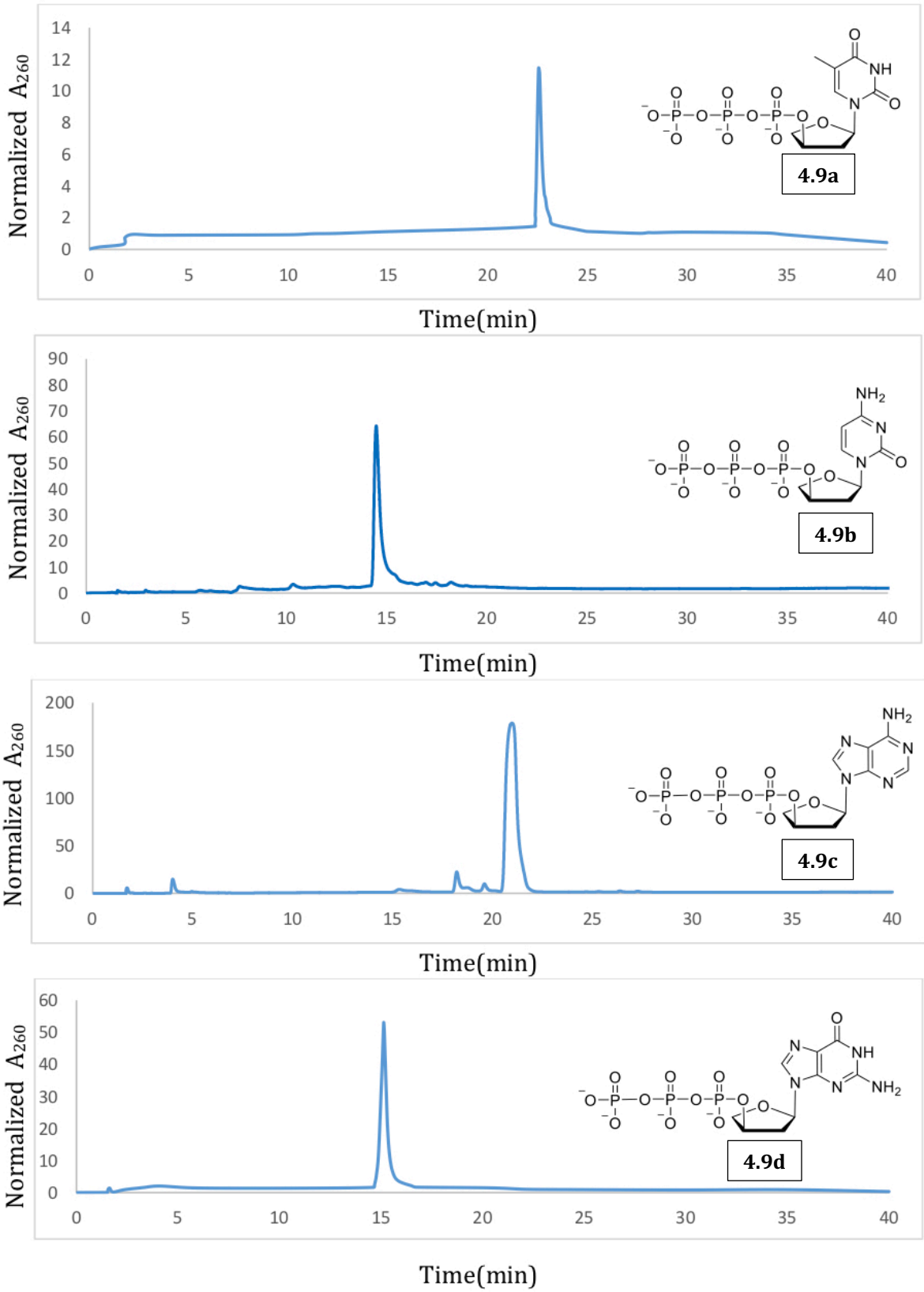
HPLC analysis of **2.5a-d** (MeCN/0.1M TEAA from 0 to 50%).



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



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



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