SYNTHESES AND BIOCHEMICAL EVALUATION OF *CAGED*-siRNA, *TROPOLONYL*-DNA AND *AMINOPYRAZOLONYL*-PEPTIDES

By
AMARNATH BOLLU
CHEM11201204001

National Institute of Science Education and Research
Bhubaneswar

A thesis submitted to the
Board of Studies in Chemical Sciences
In partial fulfillment of requirements
for the Degree of
DOCTOR OF PHILOSOPHY
of
HOMI BHABHA NATIONAL INSTITUTE

December, 2019

SYNTHESES AND BIOCHEMICAL EVALUATION OF *CAGED*-siRNA, *TROPOLONYL*-DNA AND *AMINOPYRAZOLONYL*-PEPTIDES

By
AMARNATH BOLLU
CHEM11201204001

National Institute of Science Education and Research
Bhubaneswar

A thesis submitted to the
Board of Studies in Chemical Sciences
In partial fulfillment of requirements
for the Degree of
DOCTOR OF PHILOSOPHY
of
HOMI BHABHA NATIONAL INSTITUTE

December, 2019

Homi Bhabha National Institute

Recommendations of the Viva Voce Committee

As members of the Viva Voce Committee, we certify that we have read the dissertation prepared by **Amarnath Bollu** entitled "**Syntheses and Biochemical Evaluation of *Caged*-siRNA, *Tropolonyl*-DNA and *Aminopyrazolonyl*-Peptides**" and recommend that it may be accepted as fulfilling the thesis requirement for the award of Degree of Doctor of Philosophy.

Chairman – Prof. A. Srinivasan	27.12.19
Guide / Convener – Dr. Nagendra K Sharma	27.12.19
Examiner - Prof. Surajit Sinha	S. Sinha 27.12.19
Member 1- Dr. Chandra Shekhar Purohit	27-12-19
Member 2- Dr. Sanjib Kar	27.12.19
Member 3- Dr. Manjusha Dixit	27.12.19

Final approval and acceptance of this thesis is contingent upon the candidate's submission of the final copies of the thesis to HBNI.

I hereby certify that I have read this thesis prepared under my direction and recommend that it may be accepted as fulfilling the thesis requirement.

Date: 27.12.19

Place: NISER-Bhubaneswar

Signature

Guide

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at Homi Bhabha National Institute (HBNI) and is deposited in the Library to be made available to borrowers under rules of the HBNI.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the Competent Authority of HBNI when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

AMARNATH BOLLU

DECLARATION

I, hereby declare that the investigation presented in the thesis has been carried out by me. The work is original and has not been submitted earlier as a whole or in part for a degree / diploma at this or any other Institution / University.

B. Amarnath

AMARNATH BOLLU

List of Publications arising from the thesis

Journal

1. "Tropolone conjugated DNA: pH Induced Fluorescence Enhancement of Tropolone in DNA Duplex", Amarnath Bollu and Nagendra K. Sharma, *ChemBioChem*, **2019**, *20*, 1467-1475.

2. "Syntheses and Conformational Analysis of Aminopyrazolonyl Amino Acid (APA)/Peptides", Amarnath Bollu and Nagendra K. Sharma, *Eur. J. Org. Chem.*, **2019**, *2019*, 1286-1292.

3. "Cleavable Amide Bond: Mechanistic Insight of Cleavable 4-Aminopyrazolyloxy Acetamide at low pH", Amarnath Bollu and Nagendra K. Sharma, *J. Org. Chem.*, **2019**, *84*, 5596-5602.

Conferences

1. Poster presentation: "Light-dependent RNA interference with 2'-modified caged siRNA" in Photoinduced Processes in Nucleic Acids and Proteins: Faraday Discussion" (11-13)th January, 2018, Kerala, India.

2. Poster presentation: "Syntheses and Conformational Studies of Pyrazolone Derived Amino Acids/Peptides" in Inter IISER and NISER Chemistry Meet, IINCM-2017, (22-24)th December, 2017, NISER-Bhubaneswar, India.

3. Poster presentation: "Syntheses and Conformational Studies of Pyrazolone Derived Amino Acids/Peptides" in ACS on Campus, 23rd July, 2018, NISER-Bhubaneswar, India. *(Best Poster Awardee)*

4. Participant: International Conference in Chemistry, iconChem2018, (24-26)th May, 2018, IISER-Tirupati, India.

Others

1. "Syntheses of N^1-/N^6-Nitrobenzylated-*Caged*-2'-*tethered*-Adenosine Analogues and their incorporation into DNA/RNA", Amarnath Bollu and Nagendra K. Sharma (*communicated*)

B. Amarnath

AMARNATH BOLLU

DEDICATIONS

I dedicate this thesis for my supervisor, Dr. Nagendra K. Sharma for taking me a step closer to explore myself to the depth of scientific ideology.

In spirit of sciences

Universe is big with full of light,

cell is small with full of life,

with the weight of earth on the lids of eye,

fly to the light and life with the wings of eye…

ACKNOWLEDGEMENTS

I sincerely acknowledge National Institute of Science Education and Research (NISER) and its flourishing research environment. It is my privilege to be part of NISER and its community. I thank University Grant Commission (UGC) for Junior/Senior Research Fellowship and later financial support from NISER.

Am very much thankful to my supervisor, Dr. Nagendra K. Sharma for accepting me to work under his guidance. Am also grateful to my supervisor for providing indisputable research atmosphere to explore myself in path of scientific ideology. And certainly provided a strong foundation for my future research perspectives.

Am also thankful for Faculty-School of Chemical Sciences (SCS), my research lab members - Dr. G. Chandrasekar Reddy, Dr. C. Balachandra, Dr. Smitha Mohanlal, Bibhuti B Palai, Sagarika M, Chinmay K Jena, Manish K Gupta, Subhashree S Panda, Ramu, Saideep, and especially Nihar, Prajanandan Giri for their solidarity support during the course of hardship. I especially thankful to Dr. Manjusha Dixit and lab members- Khurshid and Dinesh from School of Biological Sciences (SBS) for their valuable scientific discussions and support. I also acknowledge Prof. A. Srinivasan, Dr. C S Purohit, Dr. Sanjib Kar, Dr. Himansu S Biswal, Dr. C. Gunanathan from NISER fraternity and also Dr. C Ganesh for their constant support at various research circumstances.

I thank my dear friends Rajendra Prasad, Durga Prasad, Dr. Ranjay Tiwari, Dr. A. Tirupati, Dr. S. Pardhasaradhi, Dr. M. Venkateswararao, Shubhayan C and PhD batch mates (Aug-2012, Chemistry). Immensely thankful to thesis reviewers for their valuable inputs in the present thesis. Without moral support from my family members, it is impossible for me to accomplish the research sovereignty.

CONTENTS

LIST OF SCHEMES

LIST OF TABLES

CHAPTER 1

Introduction to the Biological processes of Nucleic Acid/Proteins and their Therapeutic Applications

Chapter 1. Introduction to the Biological Processes Involving Nucleic Acid/Proteins and their Therapeutic Applications

1.1 Introduction

The unifying feature of all cellular life-forms on earth is their function at biochemical level. Though they appear different externally, they all use similar biological macromolecules and their building blocks in conjunction with small molecules to regulate the cellular functions. The macromolecules include nucleic acids, (DNA, RNA), proteins, lipids, carbohydrates. While the small molecules include water, oxygen, carbon dioxide, glucose, vitamins, and cholesterol are few specific examples. Herein, DNA, RNA, and proteins are dicussed.

DNA structure

In 1953, J.D. Watson (an American biologist) and F.H.C. Crick (a British Physicist) proposed the double helical structure of DNA (B-DNA).[1, 2] The helical DNA consists of a phosphate backbone, sugar and nitrogenous bases with minor and major grooves. The primary component in DNA is repeating units of monomers known as nucleotides, each nucleotide consists phosphate group, sugar and nitrogenous base. Each sugar unit is attached to four different types of nitrogenous bases. The four nitrogenous bases (nucleobases) are adenine (A), guanine (G) (purine bases), cytosine (C), and thymine (T) (pyrimidine bases). The DNA helical structure consists of complimentary single strands arranged in anti-parallel fashion (one strand 5'-3'; another strand 3'-5') which are held together with hydrogen bonding between the nucleobases to form base pairs and further stabilized by base stacking with adjacent nucleobases. This base pairing is specific and observed

only between purines and primidines, more specifically, A basepairs with T (A:T); G basepairs with C (G:C) with two and three hydrogen bonds respectively as shown in Figure 1.1.

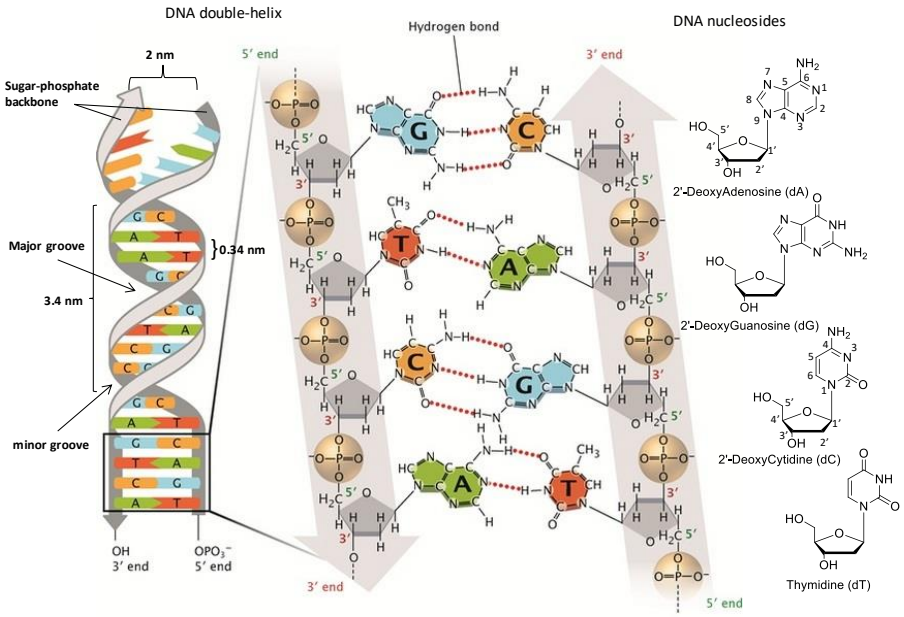


Figure 1.1. DNA structure, base pairing and nucleosides. Copyright 2013 Nature Education.

RNA structure

Similar to DNA, RNA is also a polymer made of nucleotides and mostly exists in single stranded form that often folds on itself. Unlike DNA, the sugar ring of RNA has a hydroxyl (OH) group at 2'-position. The four nitrogenous bases (nucleobases) are similar to DNA nucleobases except thymine is replaced with uracil. These are adenine (A), guanine (G) (purine bases) and cytosine (C), uridine (U) (pyrimidine bases) (Figure 1.2).[3]

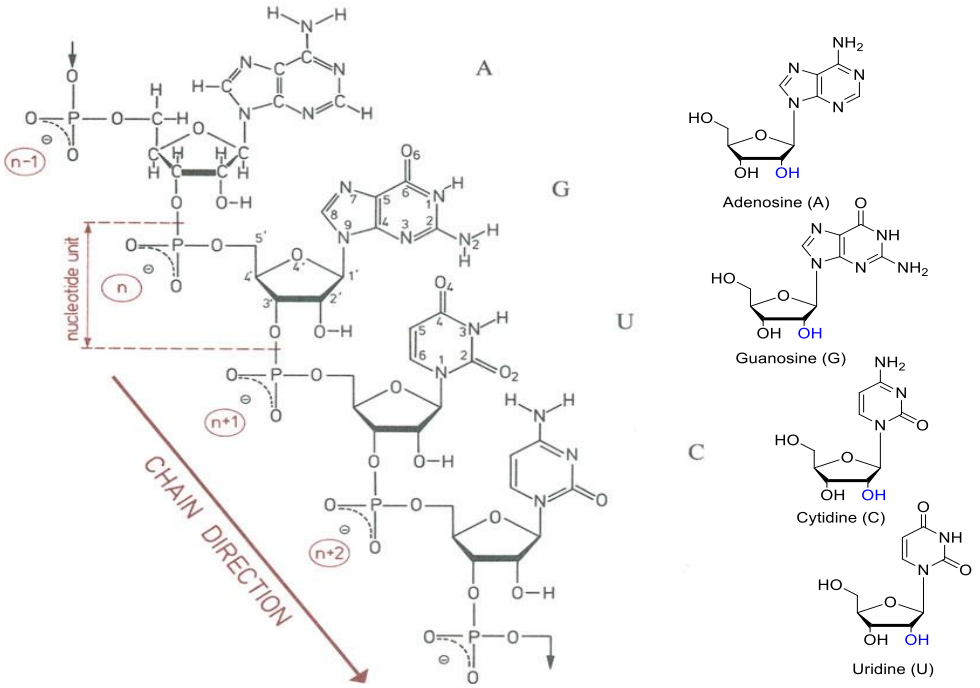


Figure 1.2. RNA structure and its components. Copyright 2013 Nature Education.

Proteins

Proteins are made of peptides, which are covalent assembly of 20 amino acid building blocks through amide bonds[4]. They have broad range of conformational diversity and functions, which are very specific from their self-assembly through non-covalent interactions, such as hydrogen bonding, van der Waals interactions and π-π stacking. Broadly, proteins acquire α-helix, β-turn type of conformations in their secondary structures.[5] The constituents of proteins are amino acid building blocks, each amino acid consist of amino group ($-NH_2$) and carboxylic acid group (-COOH) which are attached to the central carbon atom (α-carbon), and mostly amino acids exists in zwitter-ionic form ($-NH_3^+$, $-COO^-$). With different substituent at α-carbon (side chain), these twenty amino acids, except glycine, have stereogenic centre at α-carbon. All the naturally occurring proteins and peptides are made of amino acids with L- configuration (Figure 1.3).

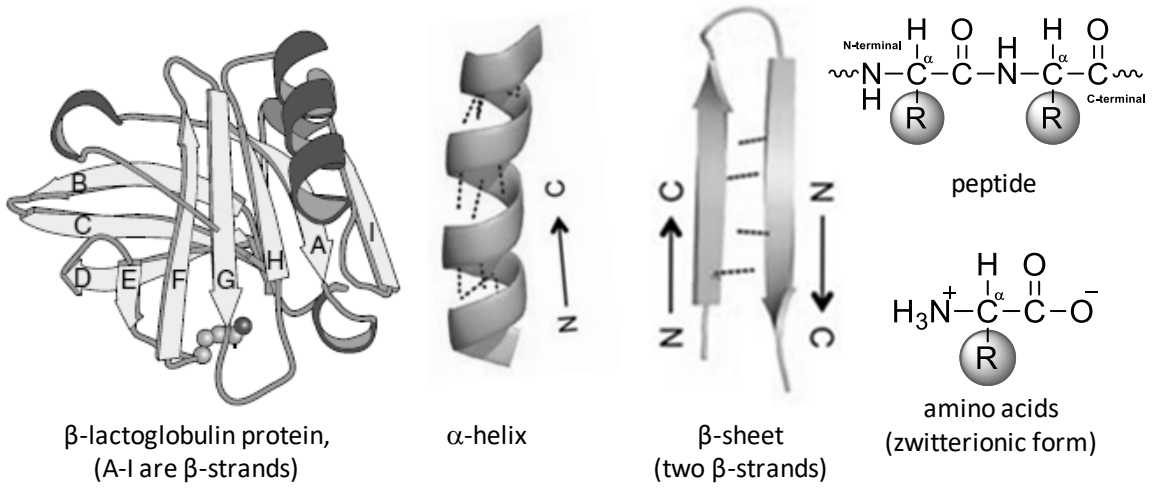


Figure 1.3. Structure and constituents of protein. (PDB code: 1beb); R-group defines naturally occurring L-amino acids

Central dogma of molecular biology

DNA is the hereditary material, which carries the genetic information from one generation to another generation. Genes are sections within the DNA with unique nucleotide sequences which store the genetic information. It is now believed that, humans contain 20,000-25,000 genes and each gene carries from few hundred DNA base pairs to few million DNA base pairs.[6] Broadly, the synthesis of proteins involves two steps[7]: (i) Transcription, process in which a portion of DNA is unwound with helicase enzyme to single strand with this strand acting as the template for mRNA synthesis. The enzyme RNA polymerase reads 3'end to 5'end of template DNA and makes complementary pre-messenger RNA (pre-mRNA) which consists of exons and introns, which is further processesed to mRNA which contains only exons in nucleus. (ii) Translation, the mRNA is transported from nucleus to ribosomes in the cytoplasm. The ribosomes read mRNA from 5'end to 3'end and mRNA acts as template for protein synthesis. Each trinucleotide (codon) of mRNA codes for specific amino acid according to the genetic code. Importantly, during protein synthesis,

each amino acid is carried by transfer RNA (tRNA) (Figure 1.4). The process of protein synthesis from specific genes on DNA through transcription and translation is known as gene expression. These proteins act as enzymes, which catalyse almost all the chemical transformations in living cells. In DNA, there are coding genes and noncoding genes. The coding genes will code for functional proteins. The noncoding genes will code for noncoding RNAs (ncRNA) which are not transformed into proteins, rather they involve in regulatory functions.

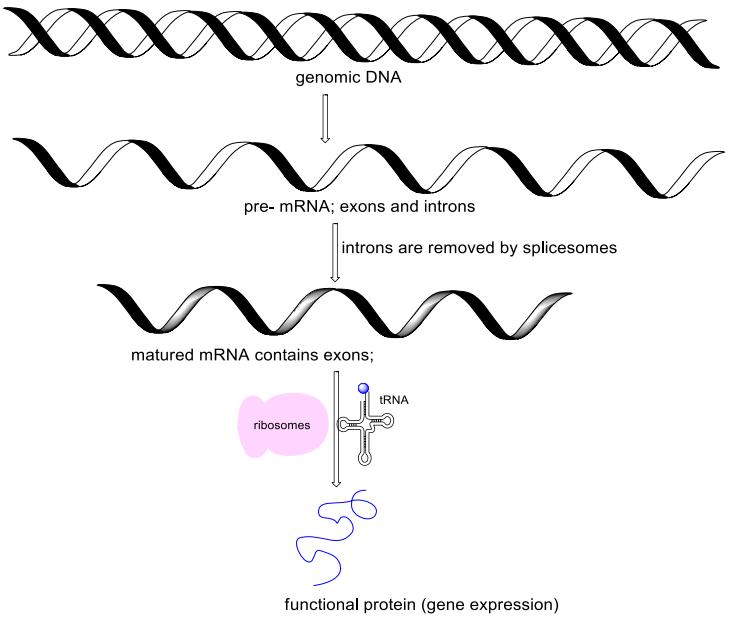


Figure 1.4. Central dogma in biochemistry.

RNA Interference (RNAi), double-stranded RNA-mediated gene regulation

RNA *Interference* (RNA*i*) is the gene regulation mechanism by small double-stranded, non-coding RNAs which were discovered in plants and animals.[8, 9] RNA interference (RNA*i*) involves 21-23 nucleotide-long, *micro-RNAs* (miRNA) which regulates protein coding genes.[10, 11] The mechanism of RNAi activity in mammals is as follows,[10, 12, 13] (Figure 1.5):

1. miRNA gene transcripts the primary miRNA (pri-miRNA) catalyzed by RNA polymerase II and III. This pri-miRNA is folded to form imperfect double stranded RNAs.
2. Drosha RNase III endonuclease enzyme cuts the pri-miRNA and release the 60-70 nucleotide stem-loop miRNA known as pre-miRNA. The base side of pre-miRNA has 5'-phosphate and 3'-has two nucleotide overhangs. This pre-miRNA is transported to cytoplasm by exportins.
3. In cytoplasm, Dicer recognizes the 5'-phosphate and 3'-has two nucleotide overhangs ends and cut the loops of other side of pre-miRNA. This also have 5'-phosphate and two nucleotide 3'overhangs on both the ends.
4. Essential argonaute proteins (Ago 1-4) binds with resulting complex to form RNA-induced silencing complex (RISC). The less thermodynamically stable 5' end of miRNA weakly binds with Ago protein in RISC and unwinds the miRNA; the unbound strand leaves the RISC and undergoes degradation.[14] The strand which sets loaded on the RISC is known as *guide strand* or *antisense strand* and the strand which leaves the RISC is known as *passenger strand* or *sense strand*.
5. The matured RISC down regulates the gene silencing in two ways depending on mRNA base pairing.
    a. mRNA cleavage: this occurs when the miRNA forms perfect by complementary base pairing. The activated RISC with complementary guide and mRNA cut at the 10-11 nucleotide base pair from 5' end of guide strand.[15, 16]
    b. mRNA suppression: this occurs when the miRNA does not find perfect match, and bind with mRNA with segmental complementary binding. This mostly occurs at 3'

untranslated region (3'-UTR) of mRNA. This will not cleave the mRNA but inhibit further translation.[17]

Importantly, a single miRNA can control several mRNA; conversely, several miRNA can control single mRNA.[18]

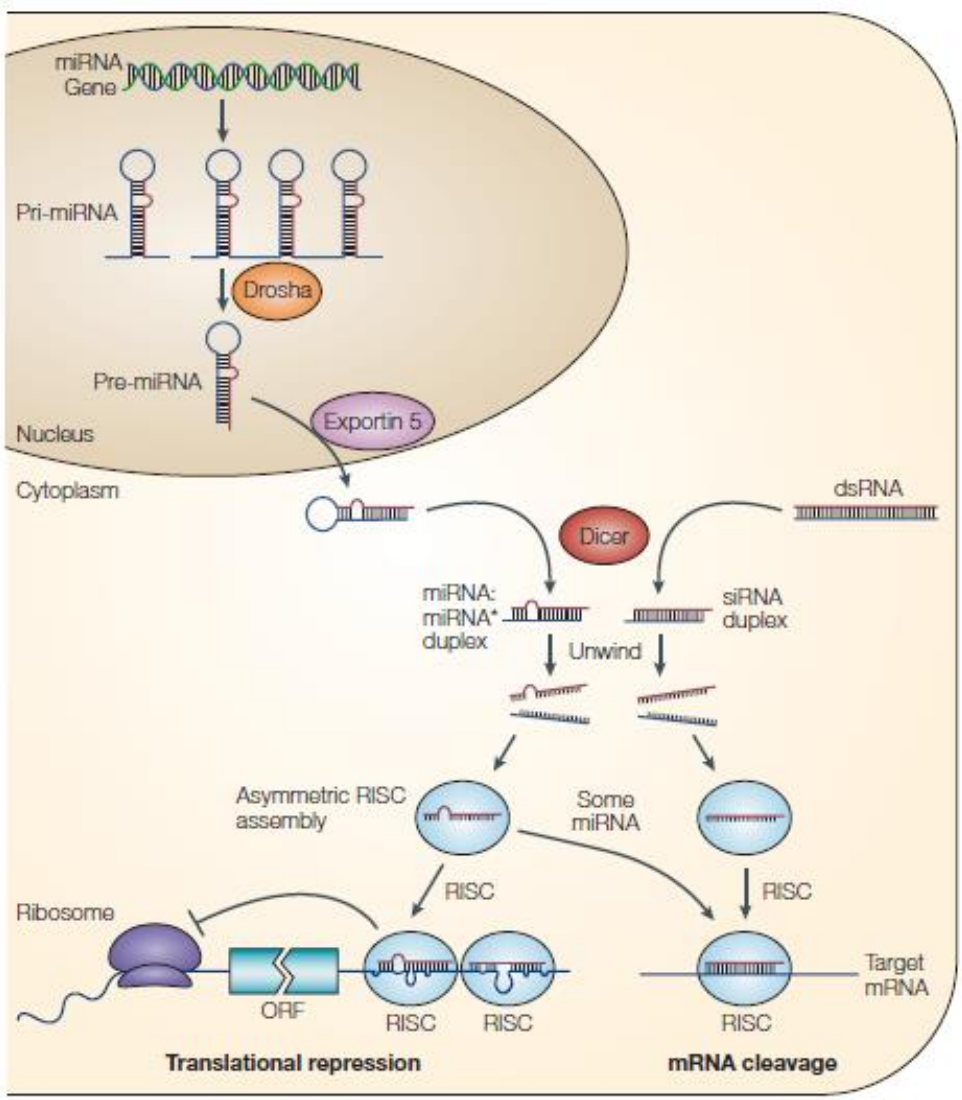


Figure 1.5. miRNA and siRNA induced RNA interference (RNAi).[19] Copyright 2004 Springer Nature

Nucleic acid mutations are associated with diseases

Due to rare mutations such as mismatch base pair (non-covalent), or methylations (covalent) in genes, some DNA, mRNA and their respective proteins are not regulated. This may lead to the up-regulation or down-regulation of their mRNA, protein expressions or they sometime become dormant.[20] If these mutations in gene are beyond the control of endogenous DNA repair mechanisms, then they may lead to the dysfunctioning of specific metabolic activity which results in disease such as genetic diseases or cancer.[21] The inhibition of regulatory functions of miRNAs also widely associated with genetic diseases or cancer. Here are few representative examples:

1. In humans, a *single nucleotide modification* in single motor neuron gene, *SMN1 and SMN2*, (exon 7, C to T) leads to severe spinal muscular atrophy (SMA).[22, 23] SMA occurs in approximately 1 in 10,000 live births.[23] The patients with SMA witness the muscle weakness, neuromuscular disorder, loss of respiratory function and death.[24]
2. miR-15/16 genes express microRNAs which are known to act as tumour suppressors. Low levels of miR-15 and miR-16 are associated with Chronic Lymphocytic Leukaemia (CLL).[25, 26]
3. Diminished levels of let-7 miR family associated with Human Non Small Cell Lung Cancer (HNSCLC).[27-30]
4. Duchenne's muscular dystrophy (DMD) is a rare and fatal disease caused by mutations in DMD genes. This is due to frameshift mutations which translate to non-functional proteins. The reduce levels of this proteins leads to cell death and loss of muscle function.
5. LDL cholesterol (low density lipoprotein) components are produced more due to over expression of apolipoprotien-B mRNA. This condition is known as homozygous familial

hypercholesterolaemia. This condition leads to increase the LDL cholesterol levels in serum, which causes fatal heart disease.[31, 32]

6. Due to mutations in gene encoding transthyretin, their mRNA transcripts translate transthyretin proteins which forms amyloid type of transthyretin fibrils which lead to polyneuropathy and cardiomyopathy.[33, 34]

Nucleic acid approach for genetic diseases and cancer

With this few representative examples, mutations in nucleic acids are certainly related to various kinds of disease. In finding cure for such diseases, application of synthetic oligonucleotides which are complementary to mRNA or pre-mRNA can interfere with protein translations and may induce the regulation of target gene. Currently two methods are exclusively studied for therapeutic application of oligonucleotides: antisense oligonucleotides (ASO) and small interfering RNAs (siRNA). The discovery of RNA interference (RNA*i*) have provided opportunities for development of gene based therapeutic drugs by silencing the expression of specific genes using double stranded small RNAs such as microRNA (miRNA) and small interfering RNA (siRNA)[35-37] (Figure 1.6). The miRNAs are produced inside cells from non-coding genes while siRNAs are synthetic mimics of miRNA. An alternate approach to alter the gene expression is explored by using chemically modified single strand antisense oligonucleotides (ASOs) as DNA/RNA/Chimeric forms. These oligonucleotides form duplex with complimentary sequence of messenger RNA (mRNA) and lead to mRNA degradation by relevant enzyme machinery to accomplish target-gene expression.[38] In addition, during the processing of pre-mRNA, splicing of exons and introns play a crucial role in gene expression. An alternate splicing mechanism may lead to the synthesis of non-functional protein isoforms and splicing mechanisms play an important

role in gene expression.[39] The antisense oligonucleotides which induce the alternative splicing mechanisms are also proven to be a good target.

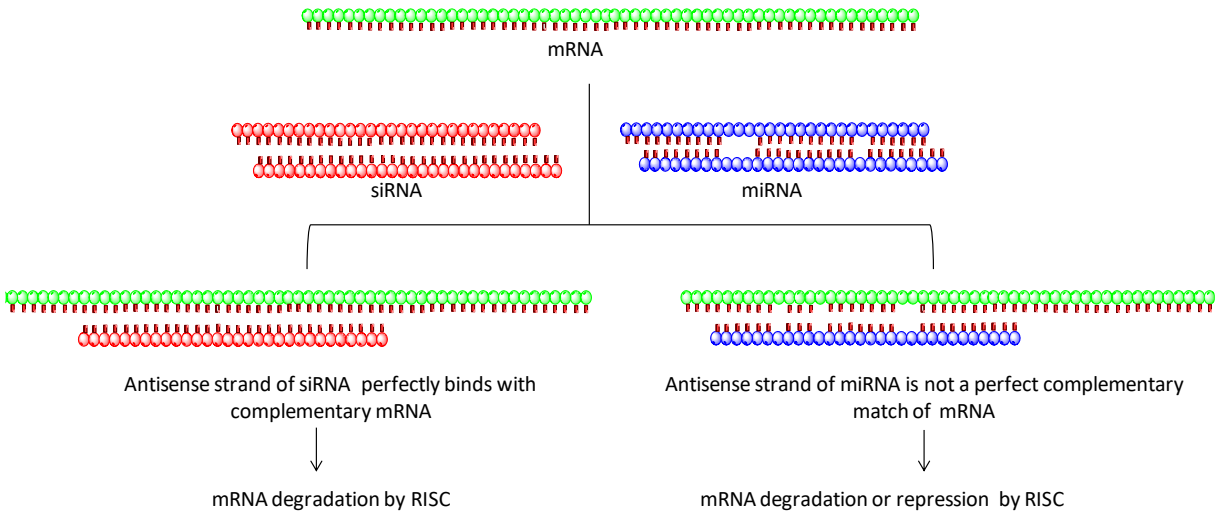


Figure 1.6. siRNA and miRNA mediated RNAi.

The extensive applications of synthetic oligonucleotides for screening the regulatory effect of target genes are amplified due to the advancements in oligonucleotide synthesis, automation and affordable cost of manufacturing. Due to automation of DNA/RNA-Synthesis, synthesis up to 30mer is routine, nevertheless, 70-80mers are also demonstrated. General synthetic steps involved in DNA/RNA automated synthesis is depicted in Figure 1.7. However, the use of unmodified oligonucleotides for therapeutic application is challenging due to poor cell permeability and *exo/endo* nucleases mediated oligonucleotide degradations.[40] These problems are addressed by chemically modifying the oligonucleotides to enhance the resistance for nucleases, improve cell permeability and diminished off-target effects.[41] Most of the chemically modified oligonucleotides are exclusively studied in antisense oligonucleotides, but these chemical modifications are also compatible in siRNAs.

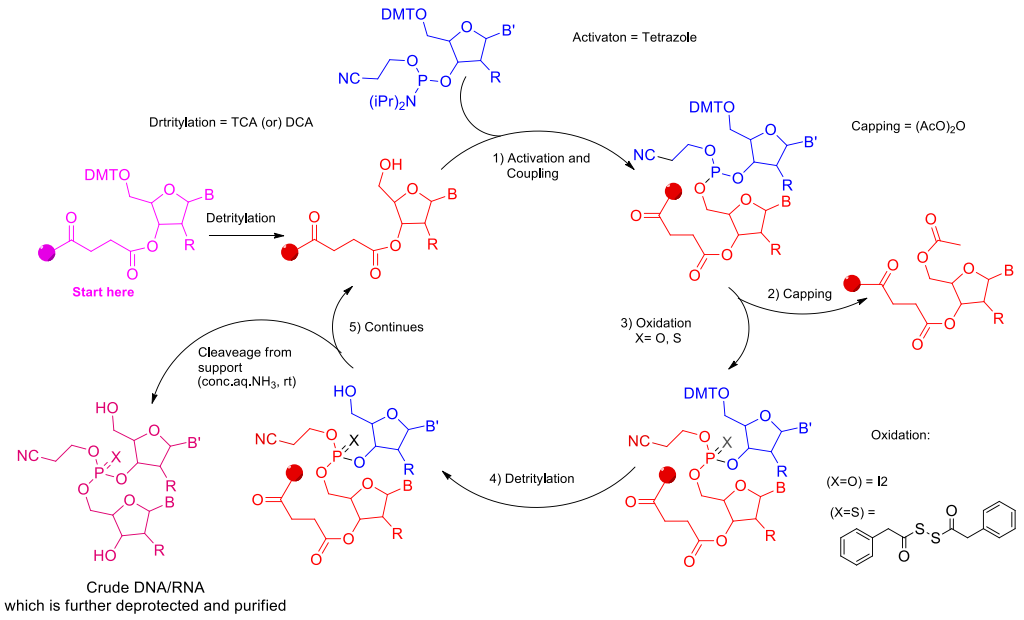


Figure 1.7. Synthetic steps involved in solid-phase DNA/RNA automated synthesis on DNA/RNA synthesizer.

Structural requirements of siRNA for RNAi activity

The 5'-OH of siRNA should be free for phosporylation, because the 5'-phosphate play a crucial role in identifying and loading the antisense strand on to RISC complex. The siRNA's antisense strand and mRNA should adopt A-form of duplex for RNAi activity. The 2'-OH of siRNA is not required for RNAi activity.[42, 43]

Chemically modified siRNA and antisense oligonucleotides

Due to the limitations to use natural oligonucleotides, several chemical modifications at different sites of oligonucleotides are synthesized and examined.

Phosphate modified oligonucleotides

The thiolated phosphate (PS) backbone modification to oligonucleotides improved the stability against exo/endo enzyme degradations and enhanced the cellular uptake.[41] This modification is

widely used for both in siRNA and antisense oligonucleotides. In other backbone modifications, the non-bridging phosphodiester oxygen of oligonucleotides is modified with borane, which is isoelectronic with oxygen and also carries negative charge (Figure 1.8). In boranophosphate backbone modified siRNAs, the nucleases stability is increased to 300-fold, and two folds more stable than phosphorothioate siRNAs.[44, 45]

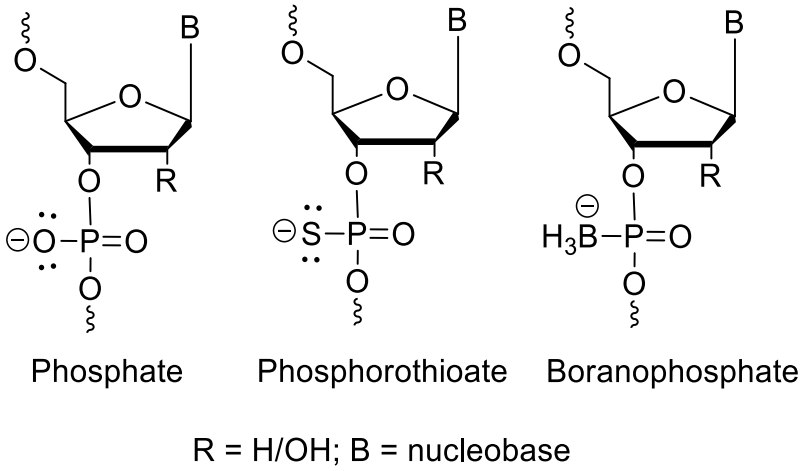


Figure 1.8. Phosphate backbone modified oligonucleotides. These modifications are isoelectronic and carry different negative charge on phosphate backbone.

Sugar ring modified oligonucleotides

Since 2'-OH is not involved in RNAi, this position is widely used for chemical modification. Importantly, the presence of 2'-*O*-methyl, (2'-*O*-Me) and 2'-*O*-methyloxyethyl (2'-*O*-MOE), at 2'-*O*H resists the sugar ring puckering and induces the formation of A-form duplex, hence increases the binding affinity.[35] These modifications exhibit high stability for nucleases. The 3'-*O*-phosphodiester bond cleavage catalyzed by nucleases is minimized due to the unavailability of free 2'-*O*H. The 2'-*OMe* and 2'-*OCH*$_2$*CH*$_2$*OM*e are widely used in chemical modifications of both siRNA and antisense oligonucleotides.[46] Additionally, few relevant 2'-*O*- chemical modifications

which have promising serum stability and binding affinity are depicted in Figure 1.9. The chemical modifications such as 2'-*O*-MOE, 2'-*O*-Me and locked nucleic acids (LNA) have used in successful RNA*i* and antisense oligonucleotide therapeutic applications.[41, 47, 48]

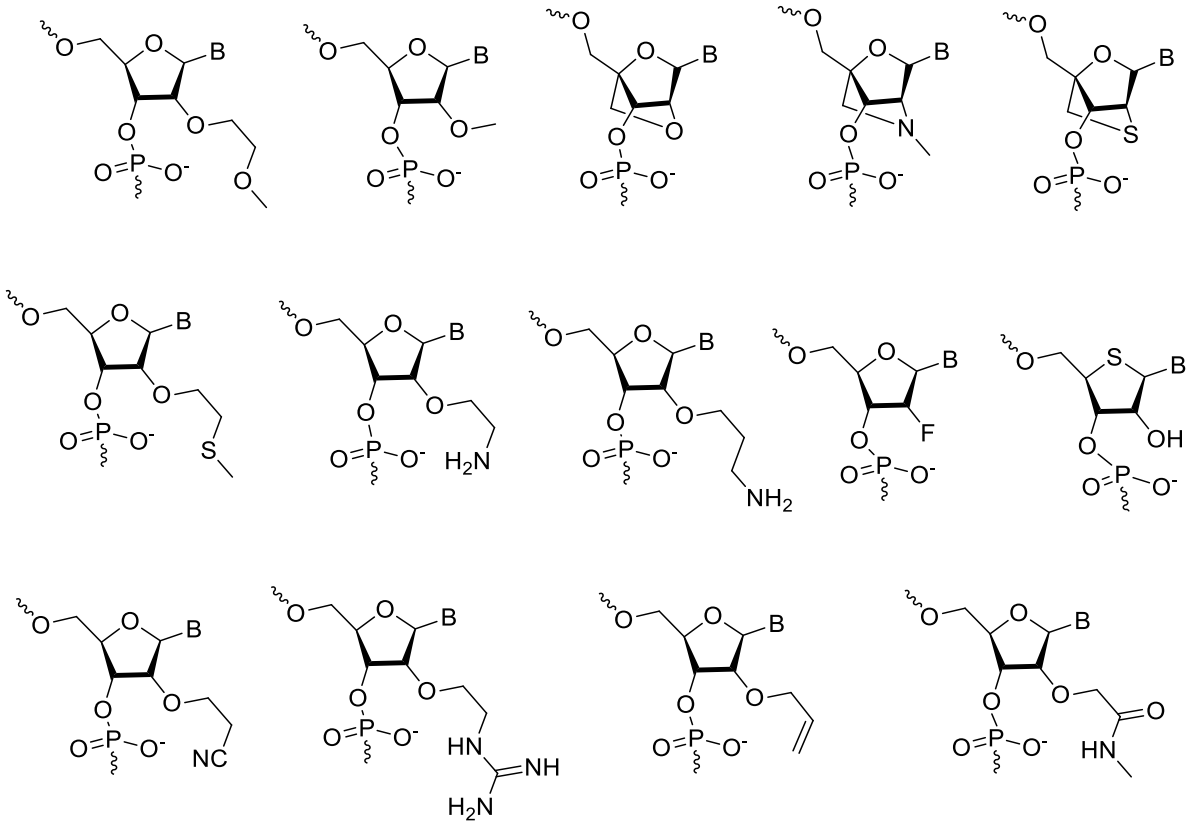


Figure 1.9. Few potent sugar ring chemical modification used in siRNAs and antisense oligonucleotides.

Nucleobase modified oligonucleotides

Modification to nucleobase will directly interfere with Watson-Crick hydrogen bonding in duplex structures. These modifications also influence the neighbouring base-stacking. With these properties nucleobase modification could directly influence the transcription and translation processes. More than 100 types of chemical modifications have occurred in cellular RNAs (mRNA, tRNA, and rRNA). The regulatory role of these naturally occurring non-classical

nucleobases is considerably increasing. Importantly, non-classical nucleobases and their association with cancer are also under consideration. For example, the recent reports show that me^6A RNA binding proteins are upregulated in human cancer tissue which increases oncotranscripts, cancer cell proliferation, tumor initiation, progression, and metastasis.[49] Regulatory roles of non-classical nucleobases is not fully understood, studies are significantly increasing to understand their role in cellular functions and disease. However, nucleobase modifications having chemical modification at C-5 position of pyrimidines and C-8 position of purines are synthesized using metal catalyzed coupling reactions and exclusively studied their biophysical and biochemical properties. These chemical modifications include ethynyl, pyrenyl, thiophene conjugated purines and pyrimidine nucleosides are few examples. The pyrenyl, thiophene, ethynyl pyrene modified nucleosides are fluorescent and are demonstrated extensively as fluorescent probes. Since these are involved in non-covalent interactions, such as H-bonding, π-π stacking, they become sensitive to the changes in this microenvironment, which results changes in their photophysical properties, such as fluorescence properties, and used them as fluorophore probes to understand the microenvironment of oligonucleotides. Eric T. Kool pioneered the exclusive work on fluorescent nucleoside and oligonucleotide probes. Importantly, the bicylo pyrrolopyrimidins are also known to interfere with Watson-Crick base pairing.[50] The detailed fluorescence behaviour of these fluorescent nucleosides/ oligonucleotides and other nucleobase modification are illustrated in many reviews.[51-54] Overveiw of some chemically modified nucleosides are depicted in Figure 1.10.

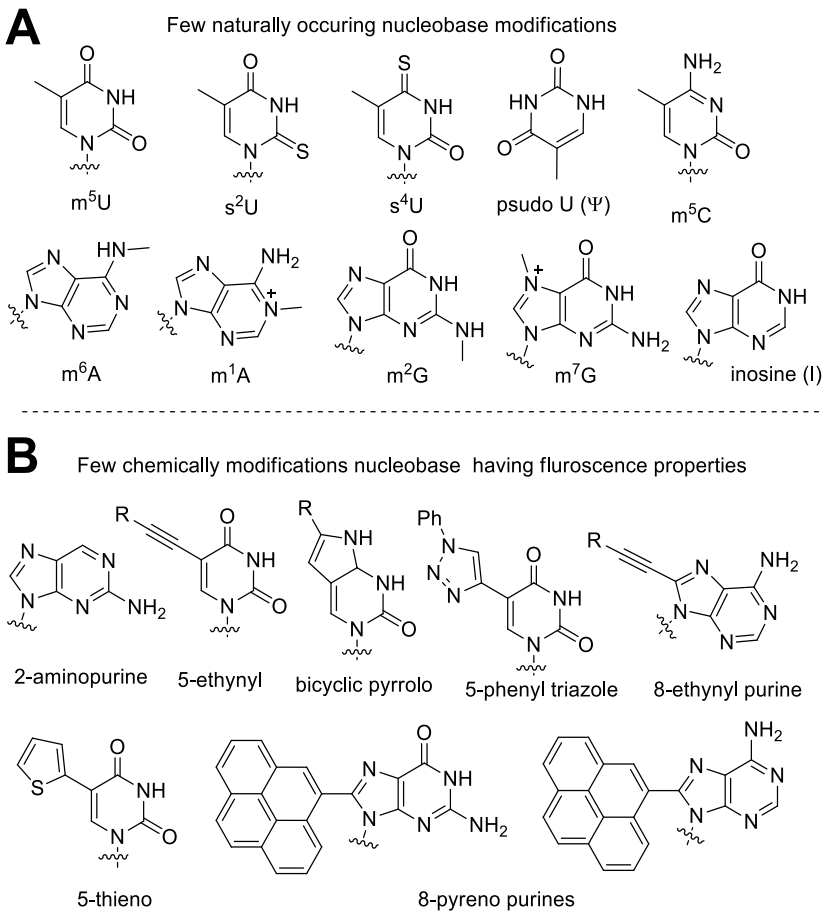


Figure 1.10. Nucleobase modified nucleoside and oligonucleotides. Naturally available non-classical nucleobases (A) and few chemically modified nucleobases, used as fluorescent probes (B).

Regulation of gene expression with light

Covalently attaching a photo labile group to a molecule and releasing that molecule with light at will is of long practice in organic synthesis.[55-57] Many photo labile groups have been developed and examined for efficient photo release of substrate using light at different wavelengths.[58] The use of photo labile protecting groups (PLPG) in bio-molecules to control their biological activity with light is an effective method to regulate their functions. The biological activities are engaged with chemical transformations at biochemical level. This involves specific functional group

transformations at the active site of interaction. If these functional groups are covalently protected with PLPG, then the biological activity can be controlled with light. In another approach, the active site of bio-molecule is attached with PLPG to induce the conformational changes and develop steric constraints which may block their function (*caging*). When the PLPG is photo cleaved with light, then the bio-molecule retains to its normal conformation and involve in biological activities (*uncaging*). Sometimes modification to bio-molecules may not affect the biological activity. Therefore, the site of PLPG incorporation and the tolerance of PLPG at the site of interactions play a key role in temporarily blocking the desired activity after caging.

The photo labile groups are also applied in siRNA and antisense oligonucleotides to regulate the gene expression with light.[59-61] The PLPGs exclusively used in oligonucleotides are depicted in Figure 1.11.

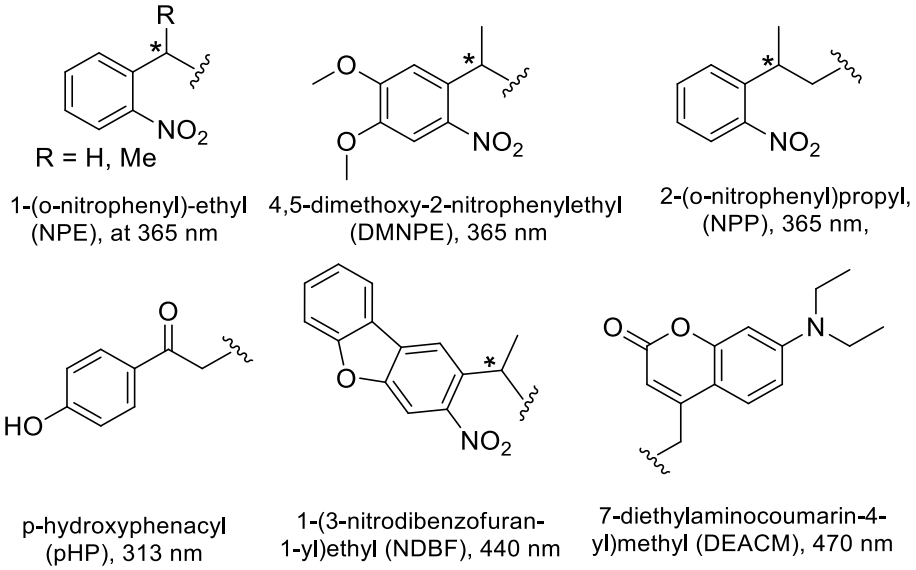


Figure 1.11. Overview of PLPGs used in exclusive study of oligonucleotides.

To control the RNAi acitivity, the siRNAs are also modified with photo labile groups (*caged* siRNA), which desired to block temporarily the interaction with RNAi enzyme machinery.

Caged phosphate analogues

For example, Friedman and Co-workers[61] treated GFP siRNAs with 4,5-dimethoxy-2-nitroacetophenone hydrazone reagent and generated caged siRNA with DMNPE as PLPG (Figure 1.12). This approach to attach the PLPG to siRNA is non-specific and believed to have PLPG on phosphate backbone in both sense and antisense strands. Moreover, this method will not provide the number of DMNPE groups present on caged siRNA.

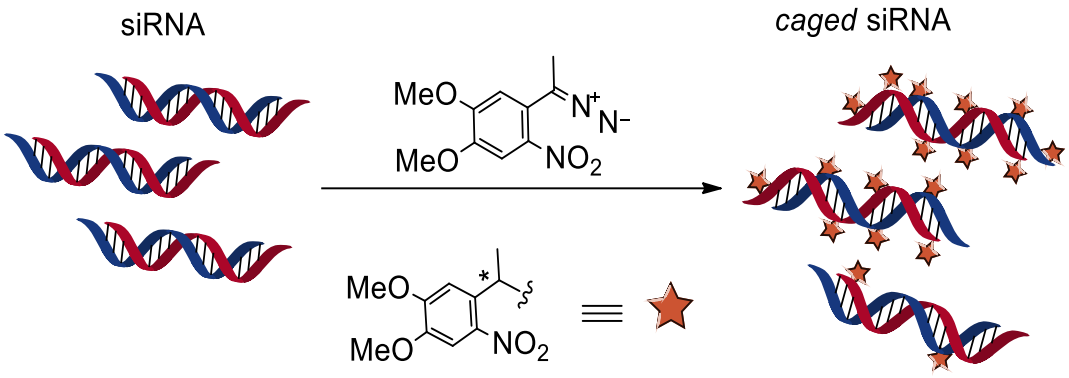


Figure 1.12. Direct ligation of DMNPE group (PLPG) onto GFP siRNA. Both sense strand and antisense strands are caged in non-specific manner.

Regulatory effect of these caged siRNA targeting the GFP transcripts are tested in HeLa cells and their results are depicted in Figure 1.13A. When the siRNA are caged, the expression of GFP is more (less RNAi activity), but the cells treated with light at 365 nm, siRNAs are uncaged and the GFP expression is reduced like unmodified control siRNA (retained RNAi activity). Similarly, Monroe and Co-workers[62] also demonstrated light controllable RNAi in BHK cells. In this study they also prepared caged siRNA having DMNPE group as PLPG, however they incorporate PLPG groups on sense and antisense strands separately and then mixed with complementary strands to generate strand specific caged siRNAs. By this approach only one strand contains the PLPG group, additionally they also prepared both sense and antisense modified caged siRNA. Moreover, these

siRNAs are fully modified with 2'-fluro modification (FNAs). These are tested for GFP expression in BHK cells and their results are provided in Figure 1.13B. When the antisense strand contains PLPG (AS-caged), the non-UV treated cells have very low RNAi activity (more GFP expression), while UV treated cells have more RNAi activity. A similar result was observed with sense strand containing PLPG (S-caged), but the UV-treated cells have more RNAi activity than AS-uncaged siRNA. Moreover sense and antisense modified caged siRNA also have similar results. All the cases, a significant difference exists between all photoexposed and non-photoexposed cells. Importantly in both the studies, the caged siRNAs with PLPGs also displayed a noticeable RNAi activity without treating with light.

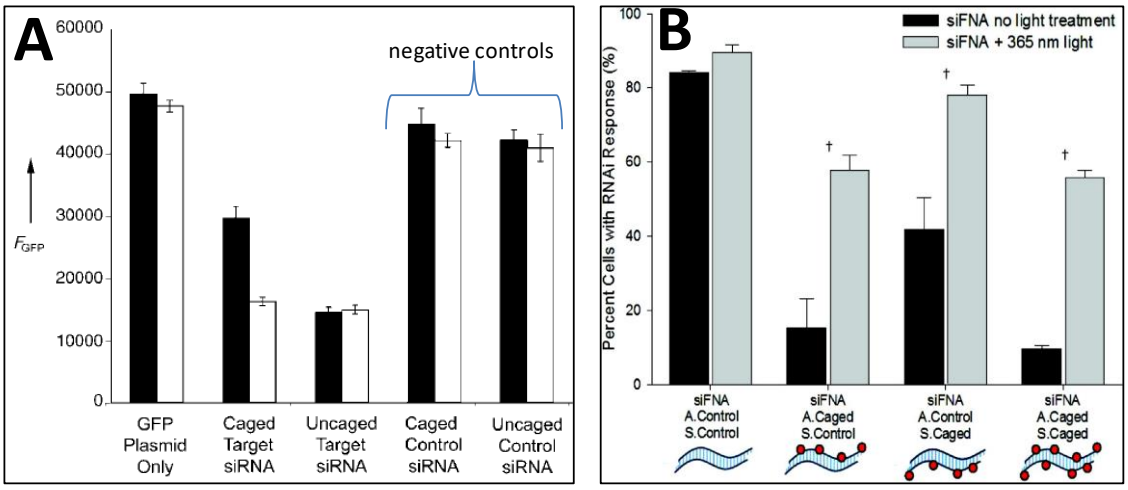


Figure 1.13. DMNPE-caged siRNA silencing of GFP, RNAi acitivity. Black bars represent no light treatment and white/gray bars are light (365 nm) treated cells. A significant difference exists between irradiated and non-irradiated samples. Friedmann et al work (A) Monroe et al work (B). Copyright 2005 John Wiley & Sons, Inc. (A) and 2008 Royal Society of Chemistry (B)

5'-Caged phosphate backbone oligonucleotides

In siRNA pathway, the 5'-OH is necessary to be free for phosporylation, the 5'phosphate play a key role in RISC complex. McMaster and Co-workers used a single biotin-labeled caged phosphate modification at the 5'-phosphate of antisense strand. The caged siRNA is transfected to HeLa cells for glyceraldehyde 3-phosphate dehydrogenase (GAPDH) assay, and demonstrated the light controllable binding of siRNA to RISC complex. The activation of siRNA was achieved after exposure with 365 nm light.[63]

Nucleobase caged oligonucleotide analogues

The site-specific PLPG modification was demonstrated by Heckel and Co-workers.[64] They synthesized the nucleobase caged deoxynucleosides and incorporated into GFP siRNA (Figure 1.14). They examined the preferential strand selectivity and positional effect of caging group for RNAi activity. The presence of NPP PLPG at Watson-Crick H-bonding interface, the caged antisense strand displayed mismatch base pairing with the complimentary mRNA and forms a bulge. The presence of bulge, the RNAi activity has been reduced. After treatment with light at 365 nm, PLPG is cleaved from Watson-Crick base pairing interface and forms a perfect match with mRNA and as a resultant the RNAi acticity is retained. This study reveals, the caging at antisense strand of siRNA has more RNAi activity and importantly, the modification at seed region within the antisense strand is more effective.

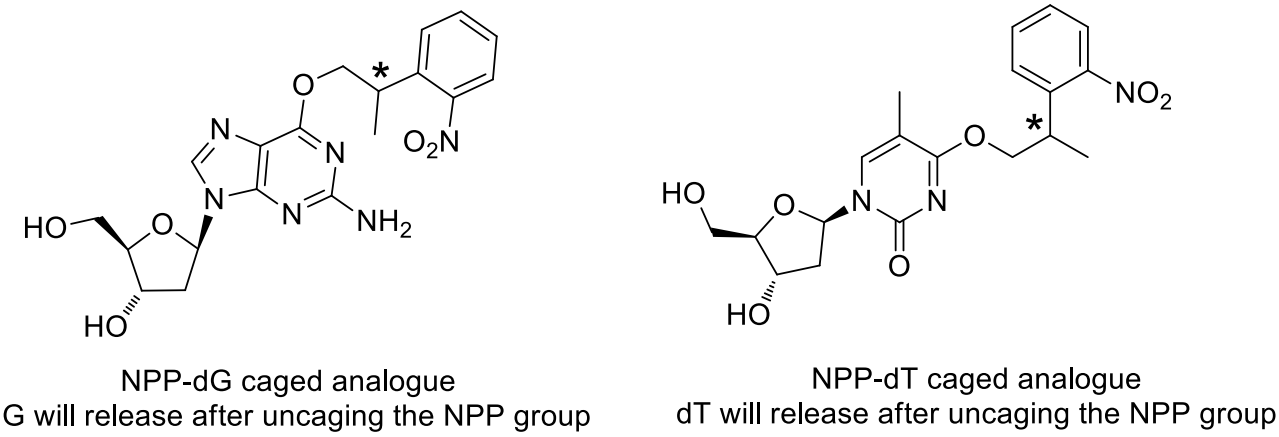


Figure 1.14. Nucleobase modified caged nucleosides used by Heckel and Co-workers to demonstrate light controlled RNAi activity.

2'-caged oligonucleotide analogues

The mRNA transcripts, encoded from DNA have introns and exons. During the process of pre-mRNA to mature mRNA, introns are removed by splicing mechanisms which are catalyzed by spliceosomes (Figure 1.15A). The matured mRNA contains only exons and these are translated to proteins. The biochemical events during splicing mechanisms are depicted in Figure 1.15B. Importantly, in primary step, the 2'-OH of adenosine attack on phosphodiester bond and form unusual 2'-5'-phosphodiester bond. Macmillan and Co-workers for the first time synthesized caged pre-mRNA transcripts and demonstrated the light controllable mRNA-splicing.[65, 66] In this study they synthesized 2'-*O*-caged adenosine nucleoside having *o*-nitrobenzyl (NBn) as photo labile group at 2'-OH. This caged adenosine is incorporated at the primary cleavage site of pre-mRNA. When the 2'-OH is protected with PLPG, NBn group, the primary cleavage step is not observed (*caging*). Upon treating with light, the PLPG is deprotected (*uncaging*) and 2'-OH is involved in primary cleavage step.

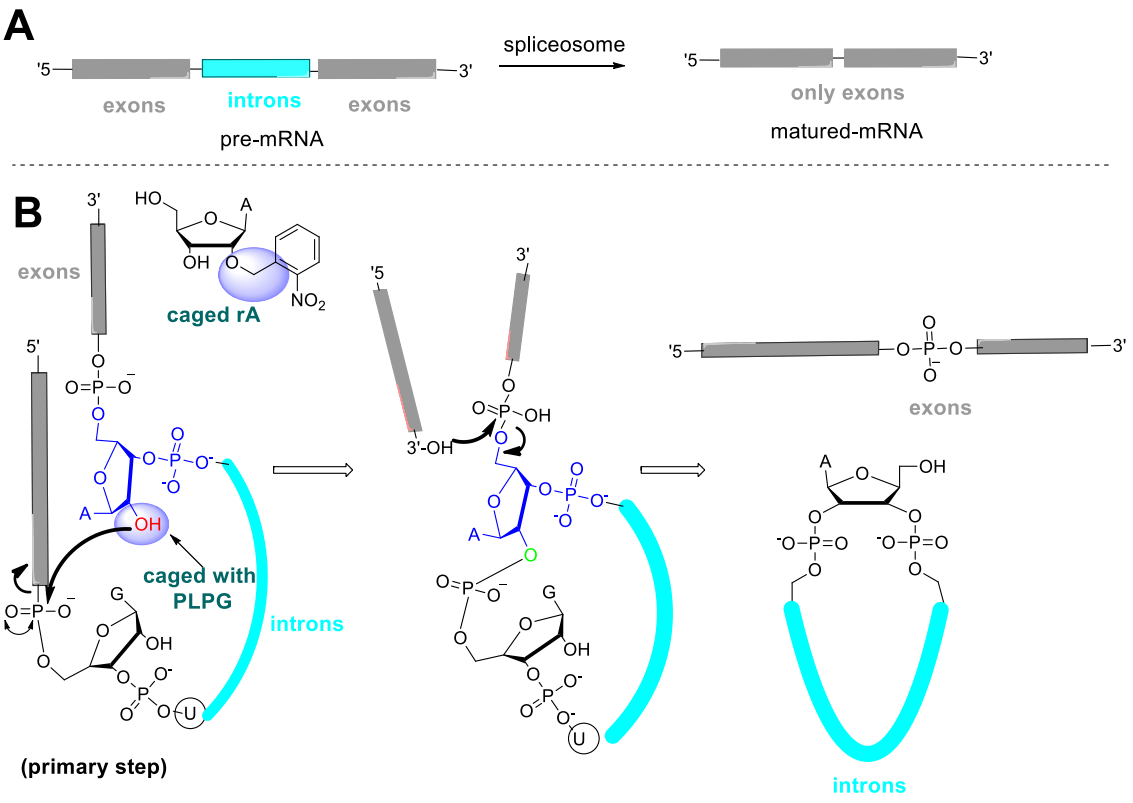


Figure 1.15. Overview of pre-mRNA processing to mature mRNA, catalyzed by spliceosome (A); biochemical events (B), in the first step, the 2′-OH of the adenosine (**A** in blue) within the intron carries out a nucleophilic displacement at the 5′-splice site followed by other events.

Challenges to use synthetic oligonucleotides as therapeutic agents

The stability of oligonucleotides has been addressed over the time by chemically modifying the oligonucleotides. The formulations of phosphorothioate and 2'-modifications such as 2'-OMe, 2'-MOE, 2'-fluro and LNA have successfully demonstrated for improved nuclease stability.[67] Despite overcoming the nucleases stable oligonucleotides by these chemical modifications, the problem of drug safety and delivery remained still challenging.

When oligonucleotide delivered in to a cell, toll-like receptors (TLRs) present in the cell, may assume these oligonucleotides as invaded foreign genetic material similar to viral/bacterial-

RNA/DNA and activates the innate immune responses.[68-71] These immune responses may lead to the dysfunctioning of various cellular mechanisms, and even apoptosis. Different TLRs recognize certain motifs in oligonucleotides and trigger the immune responses accordingly.[72-74] Judge and Co-workers demonstrated that the GU rich siRNAs having UGU motif within the siRNAs are highly predictable to TLRs recognition and triggering immune responses.[75-78] The siRNA sequences which lack GU rich regions are less predictable for activating the immune system.[75, 79, 80] On basis of these findings it is possible to design siRNA or antisense oligonucleotide with chemical modifications at specific sites to have minimal or inhibitory immune responses.

Large molecular weights and highly hydrophilic nature of oligonucleotide result in low cellular intake. Most of the siRNA are accumulated in liver or filtered-off to urine in kidneys.[81] To overcome these limitations the siRNA are formulated with carrier molecules such as lipid nanoparticles (LNP), cyclodextrin nanoparticles (CNP) with adamantine PEG as additive to improve the cellular uptake and efficacy for *in vivo* delivery.[82-84] Triantennary GalNAc–siRNA (N-Acetylgalactosamine) formulations have shown a promising delivery system for delivery of many siRNA for disease treatment, these are administered subcutaneously[84] (Figure 1.16) The GalNAc residue strongly binds to the asialoglyco-protein receptors (ASGPR) on hepatocytes in liver and obserbed into the cell mediated by endocytosis; importantly the release of siRNA is extremely pH sensitive.[85, 86] The development of new siRNA/antisense oligonucleotide drug delivery systems are rapidly growing.

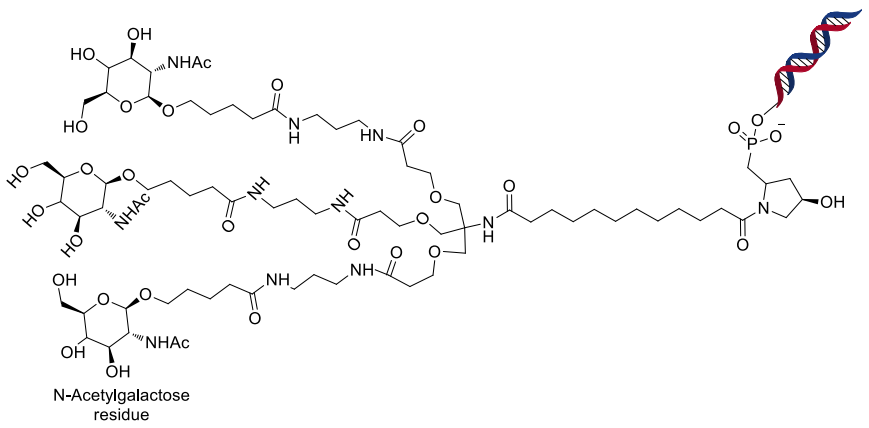


Figure 1.16. Triantennary GalNAc siRNA. Structure of GalNAc conjugate used in several siRNA drug deliveries.

Food and Drug Administration (FDA) approved oligonucleotide drugs

From many years the small drug molecules are used in treatment of many diseases. However, the scientific progress in understanding the biological processes, preferentially the role of nucleic acid in connection with various diseases has been center of attraction. Many disease-associated genes are also identified. Using desired genomic information, the oligonucleotide approach for drugs design is certainly ingenious. This approach is empowered over traditional small molecule drug design, because the oligonucleotides will selectively and specifically bind to the complimentary target sequence using Watson-Crick base pairing. Even after having the genomic information, the traditional approach for designing, identifying lead molecules itself is challenging and moreover it is expensive. Recent advancements in designing potent chemical modifications to oligonucleotides and their target specific delivery certainly increased the applications of oligonucleotides as therapeutic drug. There are more than 10 siRNAs and over 20 antisense oligonucleotides are being tested at various phases of clinical trials[87-89] (Figure 1.17).

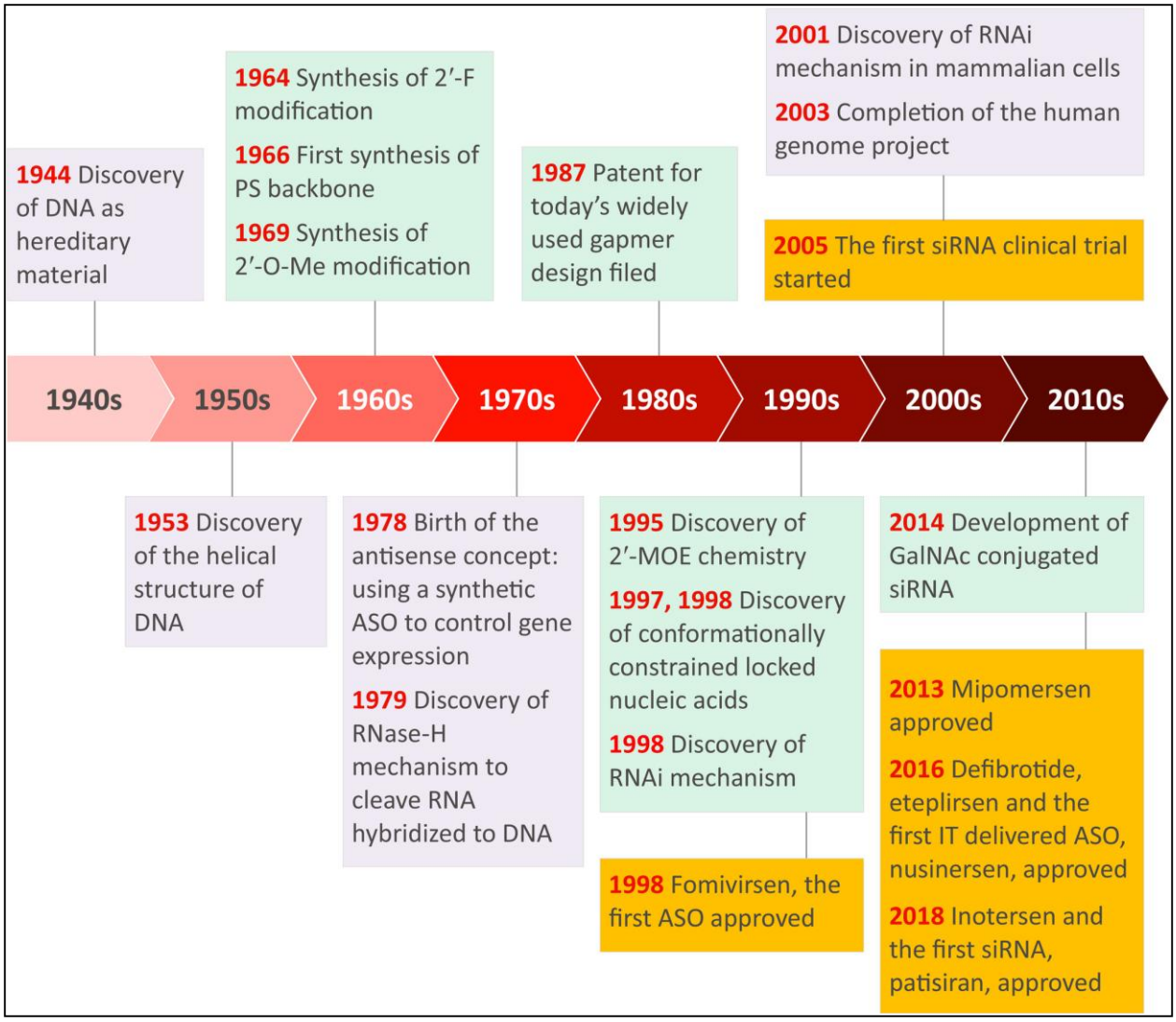


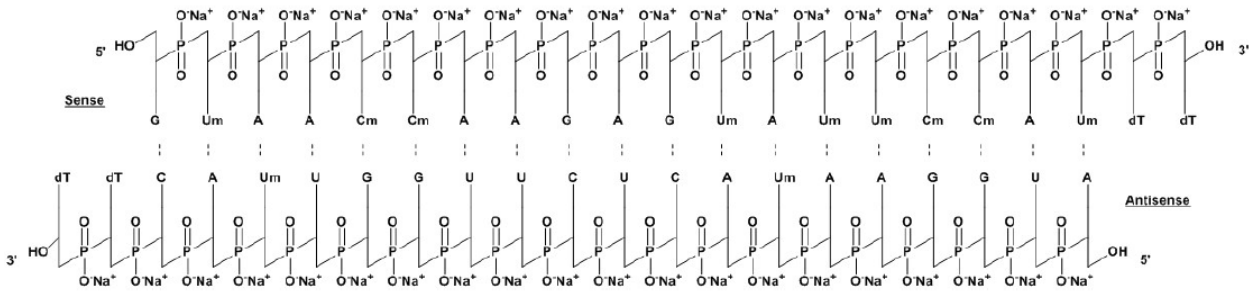
Figure 1.17. Overview of oligonucleotides development as therapeutics drugs. Copyright 2019 Wiley Periodicals, Inc.

FDA approved oligonucleotide drugs, there structure and nature of chemical modification are discussed in below.

Fomivirsen: Fomivirsen is 21mer- DNA oligonucleotide (*P*-thio;5'-GCGTTTGCTCTTCTTCTTGCG-3') with full-phosphorothioate backbone modification. It blocks the translation of viral mRNA by binding to the complementary mRNA of a key

human cytomegalovirus, CMV gene UL123, which encodes the CMV protein IE2. It is the first antisense antiviral drug approved by the FDA.[38]

Patisiran: Due to mutations in gene encoding transthyretin, mutant mRNA transcripts encode transthyretin protein which leads to polyneuropathy and cardiomyopathy. Patisiran is the first siRNA based drug approved by the FDA and the first drug approved to treat polyneuropathy. It is a gene silencing drug that interferes with mutant mRNA transcripts and induces RNAi.[90] Chemistry- the 2'-OMe is used as chemical modification with phosphate backbone. (Figure 1.18)



A, adenosine; C, cytidine; G, guanosine; U, uridine; Cm, 2'-*O*-methylcytidine; Um, 2'-*O*-methyluridine; dT, thymidine

Figure 1.18. Structure and position of chemical modifications are depicted in patisiran,

Nusinersen: In humans, a mutation in single motor neuron gene, *SMN1 and SMN2*, (exon 7, C to T) leads to severe spinal muscular atrophy (SMA). The patients with SMA witness the muscle weakness, neuromuscular disorder, loss of respiratory function and death.[24] Nusinersen, 18mer-RNA became the first approved drug used in treating this disorder. Nusinersen binds to *SMN2* transcript within the intron of the pre-mRNA and induce the alternate splicing mechanism. These events lead to the expression of *SMN2* functional proteins.[41, 91] Chemistry- the phosphate backbone is fully modified with phosphorothioates and the 2'-*O*H groups of the all the sugar rings are

modified with 2'-*O*-methoxyethyl groups. The uridine and cytosine nucleosides are replaced with 5-methyl uridine and 5-methyl cytosine respectively. (Figure 1.19)

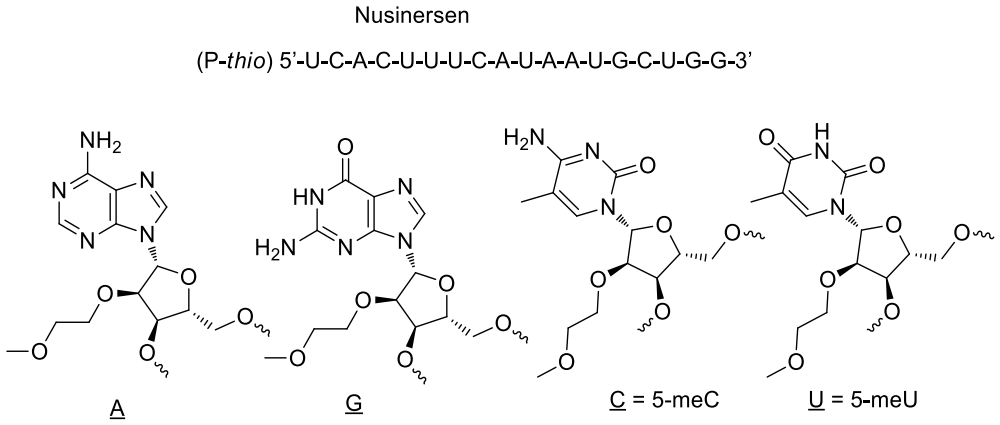


Figure 1.19. Sequence and chemical modifications of Nusinersen

Mipomersen: LDL cholesterol (low density lipoprotein) components are produced more due to over expression of apolipoprotien-B mRNA. This condition leads to increase the LDL cholesterol levels in serum. Mipomersen (20mer-chimeric DNA/RNA) binds to the apolipoprotien-B mRNA transcript. As a consequence, the mRNA is degraded by the ribonuclease-H enzyme, and apolipoprotien-B is regulated.[41] Chemistry- it contains both ribonucleosides with 2'-*O*-methoxyethyl modification and deoxy ribonucleosides (Figure 1.20)

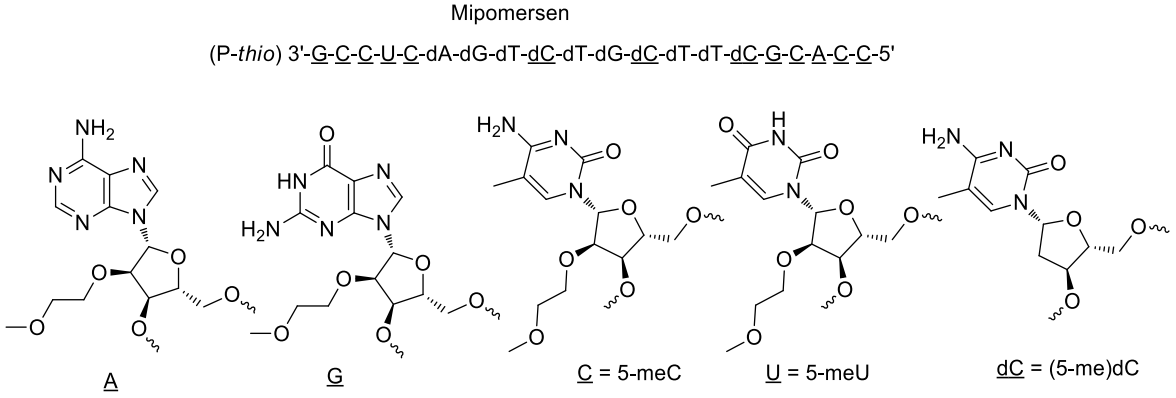


Figure 1.20. Sequence and chemical modifications of Mipomersen.

1.2 Conclusions and objectives of current thesis

Oligonucleotides are widely used for various therapeutic applications and are of great importance in treating diseases which are otherwise not treatable. From this brief introduction to chemically modified oligonucleotides for therapeutic applications, but still the regulation of gene activity with photo labile modified oligonucleotides and interaction of modified oligonucleotides with metallozymes are hardly addressed.

The photo labile oligonucleotides studied in literature will generate unmodified oligonucleotides with free 2'-OH after uncaging, which are prone to degradation via 3'-phosphodiester mediated cleavage. Hence, the development of photo labile oligonucleotides with 2'-modifications which can resist the 3'-phosphodiester mediated cleavage could be of more effective and certainly necessary to study the regulation of the gene activity with light.

Moreover, most of the siRNAs and antisense oligonucleotides interact with RNA transcripts which are mediated by enzymes having metals (metalloenzymes), therefore the role of metal ions are critical in the gene regulation. Thus modified oligonucleotides containing specific metal chelating residues could interact with those metalloenzymes and can control the protein expression. The synthesis of oligonucleotides containing metal chelating and environmental sensitive chromophores/fluorophores would be of interest in controlling the gene expression.

Further, proteins are encoded from RNA transcripts and most of the oligonucleotides are designed to target these RNA transcripts. However, proteins have unique conformational diversity and functionality from self assembly. The modified unnatural amino acids/peptides and their interaction with nucleic acid also marked their role in therapeutics and emerging rapidly.9, 10.

Development of small peptides targeting the specific motifs within the target protein is also of great utility.

The scope of the present thesis will provide understanding of photo labile *caging* and *uncaging* of siRNA and RNA/DNA oligonucleotides. *Tropolone* conjugated DNA and *Aminopyrazolone* based unnatural amino acids/small peptides will be studied.

Chapter-2, emphasizes the development of *photolabile siRNA* and *RNA/DNA oligonucleotides*, In native RNA, the 2'-*Ó*H is known to involve in cleavage of phosphordiester bond. To prevent this, the 2'-*O*H is extended with small tether group, 2'-*O*-hydroxyethyl group and photo labile group (Figure 1.21). Site-specifically incorporated the photo labile group (caging) in siRNA to sterically block the RNA interference (RNAi) prevents the degradation of its target mRNA. Only upon irradiation would remove photo labile group and release the active siRNA and initiate the RNA interference (RNAi) which degrades the target mRNA. A similar function of these modifications is hypothesized for single strand modified RNA/DNA oligonucleotides for ASO applications. The photo labile group is attached at various positions at 2'-*O*-hydroxyethyl modified nucleosides and incorporated into siRNA, RNA and DNA oligonucleotides (Figure 1.21A/B). Though there are many photo labile groups are developed, certainly the choice of photo labile group is important for application in biological system. Since the light which is used to remove the photo labile group should not damage the DNA/RNA or interfere with the photo biochemical pathways (Figure 1.21C). Under physiological conditios, the ideal range of photo cleavage is λ350nm-400nm because the damage of DNA/RNA may trigger below 350nm, and may interfere in the exisiting photo biochemical pathways such as photoisomerization of rhodopsin and bilirubins above

400nm.[92] The potential photo labile groups, which cleaves at this range, are *o*-nitrobenzyl group and its derivatives (Figure 1.21D). We choose simple *o*-nitrobenzyl group as compare to other potent *o*-nitrobenzyl derivatives.

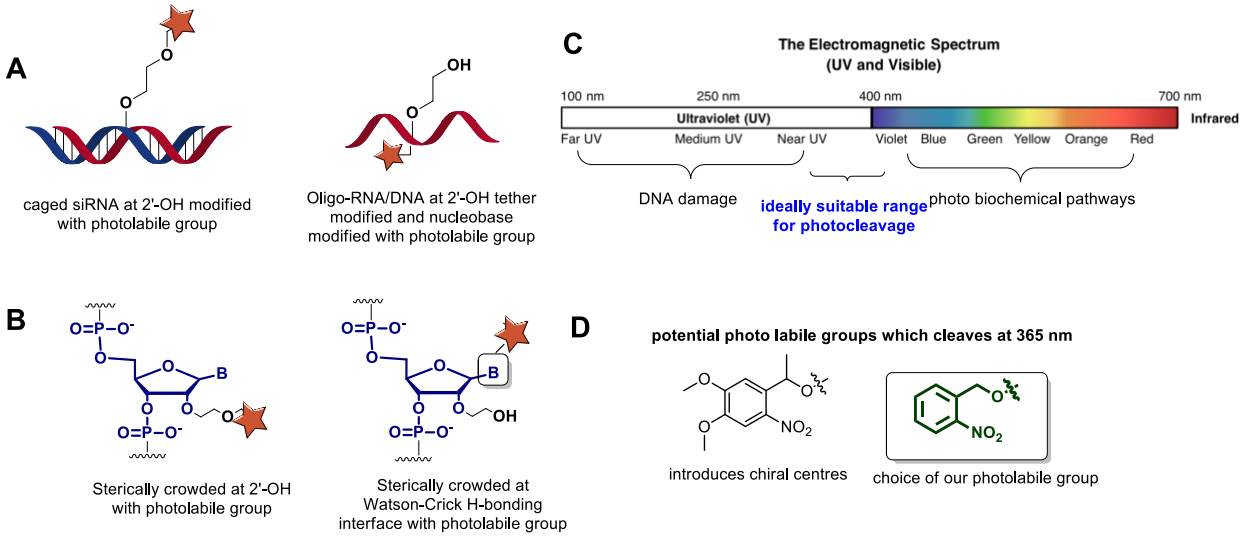


Figure 1.21. Chemical modifications and photo labile group used in this thesis. *Caged* siRNA/RNA/DNA oligonucleotides with position of modifications (A/B). Examination of electromagnetic spectrum for potential photo cleavage wavelength for photo labile group (C). Potential photo labile groups and our choice of photo labile group (D).

Herein, we describe the 2'-*O*-hydroxyethyl and *o*-nitrobenzyl group modified siRNA, RNA and DNA oligonucleotide in two sections.

A. Section-A: Synthesis and biochemical evaluation of *caged*-nucleoside and caged-siRNA analogues, which contain photo labile protected ribonucleoside, 2'-*O*-(*o*-nitrobenzyloxy)-ethyl uridine are discussed. These *caged* siRNAs tested for RNA *interference* targeting GFP transcripts in HEK-293T cell lines are illustrated.

B. Section-B: Address the challenging chemical synthesis and photo cleavable assessment of *caged* adenosine nucleosides and *caged*-RNA/DNA oligonucleotides analogues. Three *caged* adenosine modified analogues: 2'-*O*-(2-nitrobenzyloxy)ethyl adenosine, and two regiosiomers namely, N^6-nitrobenzyl 2'-*O*-hydroxyethyl adenosine and N^1-nitrobenzyl 2'-*O*-hydroxyethyl adenosine nucleoside analogues are demonstrated.

The chapter-3 describes the DNA oligonucleotides conjugated with versatile metal chelating chromophores, *tropolone* and its *tropoaminoiminyl* derivative. Tropolone and tropoaminoimines are structurally similar but fundamentally their electronic properties are reasonably different[93, 94] (Figure 1.22). For example, the dipole of tropolone is away from the ring and exhibit electropositive ring current which experimentally supports the involvement of 6π electrons in aromatic ring (Figure 1.22A-C). In contrast, the tropoaminoimines enclose the dipole into the ring, exhibit electronegitive ring current which experimentally supports the involvement of 10π electrons in aromatic ring (four from the double bonds and a lone pair of amino group) as demosnstrated in Figure 1.22D-F. Thus unique structural and biocompatibility properties of tropolone encourage us to introduce the tropolonyl/aminotropolonyl residue into DNA (Nucleosides/oligonucleotides).

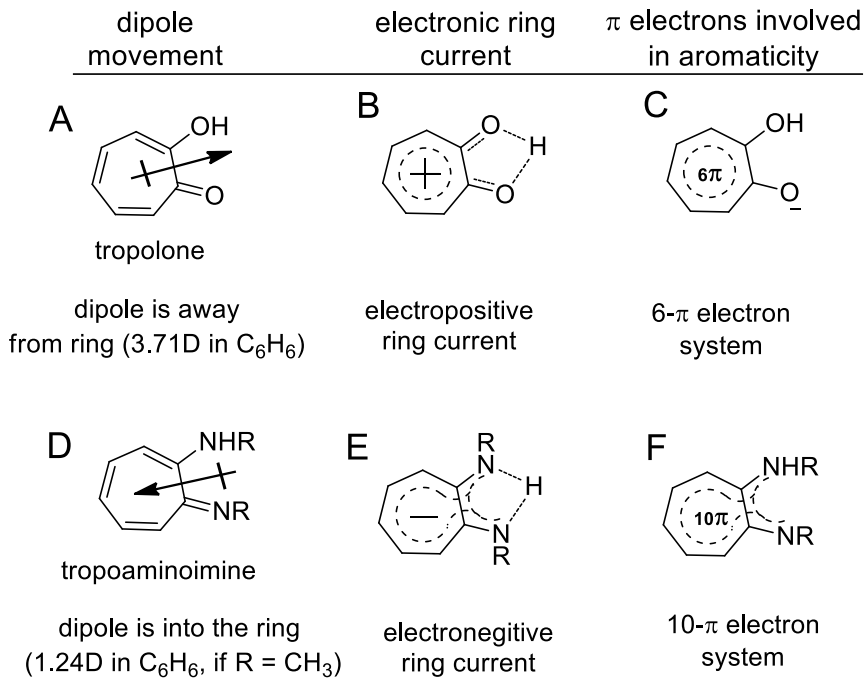


Figure 1.22. Overview of tropolone and tropoaminoimine structural properties. Tropolone (A-C) and tropoaminoimine (D-F).

Herein, we describe the tropolone and tropoaminoimine conjugated DNA oliogonucleotides in two parts. In Part-A, tropolone conjugated DNA oligonucleotides are described and in Part-B, tropoaminoimine derivatives conjugated DNA oliogonucleotides are described.

A. Section-A: The synthesis, biophysical and photophysical studies of *tropolone*-conjugated 2'-deoxyuridine nucleoside analogue and corresponding DNA oligonucleotides.

B. Section-B: The synthesis of *tropoaminoiminyl* conjugated 2'-deoxyuridine nucleoside analogues and their DNA oligonucleotides.

The chapter-4 describes the synthesis and conformational studies of *N*-phenyl/*N*-alkyl *aminopyrazolonyl* amino acids (APA) and its small hybrid peptides. Aminopyrazolone and related

derivies are bioactive molecules and considered potential drug candidates. For example, Ampyrone is analgesic drug, though it has severe side effect. Currently, ampyrone is being used as biochemical reagent (Figure 1.23A). Aminopyrazolone derivatives have strong ability to exhibit non-covalent interactions such as hydrogen bonding and π-π interactions, and thus these are considered to introduce into amino acids for tuning the structure/conformation of peptides. Herein new unnatural amino acids are rationally designed containing the *N*-phenyl (or aminoethyl) substituted pyrazolone derivatives are designed and coupled with natural peptides to explore the role of aminopyrazolone residue in the structural changes of peptides (Figure 1.23B/C). These studies are described in two different sections: Section-A/B

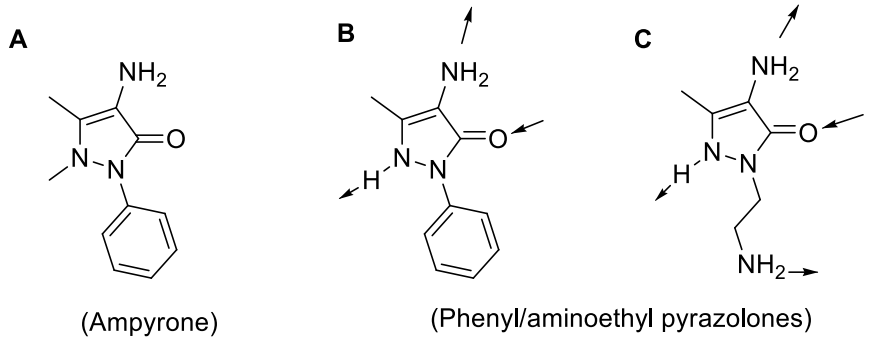


Figure 1.23. Ampyrone (A) and rationally designed phenyl/aminoethyl pyrazolone derivatives with potential binding sites (B/C).

Section-A: It describes the synthesis of *N*-phenyl/ *N*-alkyl aminopyrazole based regioisomeric *4-aminopyrazolonyl amino acid* (APA) derivatives, unnatural aromatic amino acids. The formation of three regiosiomers and two regiosiomers from *N*-phenyl/ *N*-alkyl 4-*aminopyrazolone derivative* are emphasized. The *N*-phenyl derivative, *O*-alkylalted APA derivatives are exclusively illustrated. These hybrid peptides are investigated for the formation of *organogels*.

Section-B: It demonsterates the unusual activities of amide/acid/ester bond such as the cleavage of hybrid APA analogues via formation of lactam under mild acidic conditions. It also demonstrates the methanolysis of amide bond of that lactam.

1.3 Reference

1. Watson, J. D.; Crick, F. H., Molecular structure of nucleic acids. *Nature* **1953,** *171*, 737-738.
2. Crick, F. H. C.; Watson, J. D., The complementary structure of deoxyribonucleic acid. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **1954,** *223*, 80-96.
3. Holley, R. W.; Apgar, J.; Everett, G. A.; Madison, J. T.; Marquisee, M.; Merrill, S. H.; Penswick, J. R.; Zamir, A., Structure of a ribonucleic acid. *Science* **1965**, 1462-1465.
4. Whitford, D., *Proteins: structure and function*. John Wiley & Sons2005.
5. Loveday, S. M.; Sawyer, L., β-Lactoglobulin: Nomenclature, Structure, and Function. *Reference module in food sciences* **2016**.
6. Sequencing, H. G., C. Finishing the euchromatic sequence of the human genome. *Nature* **2004,** *431*, 931-945.
7. Nelson, D. L.; Lehninger, A. L.; Cox, M. M., *Lehninger principles of biochemistry*. 7th ed.; WH Freeman, Macmillan2008.
8. Carthew, R. W.; Sontheimer, E. J., Origins and mechanisms of miRNAs and siRNAs. *Cell* **2009,** *136*, 642-655.
9. Shabalina, S. A.; Koonin, E. V., Origins and evolution of eukaryotic RNA interference. *Trends Ecol. Evol.* **2008,** *23*, 578-587.

10. Bartel, D. P., MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **2004,** *116*, 281-297.

11. Mello, C. C.; Conte, D., Revealing the world of RNA interference. *Nature* **2004,** *431*, 338-342.

12. Lee, Y.; Ahn, C.; Han, J.; Choi, H.; Kim, J.; Yim, J.; Lee, J.; Provost, P.; Rådmark, O.; Kim, S., The nuclear RNase III Drosha initiates microRNA processing. *Nature* **2003**, 415-419.

13. Lee, Y.; Jeon, K.; Lee, J. T.; Kim, S.; Kim, V. N., MicroRNA maturation: stepwise processing and subcellular localization. *EMBO. J.* **2002,** *21*, 4663-4670.

14. Khvorova, A.; Reynolds, A.; Jayasena, S. D., Functional siRNAs and miRNAs exhibit strand bias. *Cell* **2003,** *115*, 209-216.

15. Elbashir, S. M.; Lendeckel, W.; Tuschl, T., RNA interference is mediated by 21-and 22-nucleotide RNAs. *Genes & development* **2001,** *15*, 188-200.

16. Elbashir, S. M.; Martinez, J.; Patkaniowska, A.; Lendeckel, W.; Tuschl, T., Functional anatomy of siRNAs for mediating efficient RNAi in Drosophila melanogaster embryo lysate. *EMBO. J.* **2001,** *20*, 6877-6888.

17. Hutvágner, G.; Zamore, P. D., A microRNA in a multiple-turnover RNAi enzyme complex. *Science* **2002,** *297*, 2056-2060.

18. MacFarlane, L.-A.; R Murphy, P., MicroRNA: biogenesis, function and role in cancer. *Curr. Genomics* **2010,** *11*, 537-561.

19. He, L.; Hannon, G. J., MicroRNAs: small RNAs with a big role in gene regulation. *Nat. Rev. Genet.* **2004,** *5*, 522-531.

20. Modrich, P.; Lahue, R., Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annu. Rev. Biochem.* **1996,** *65*, 101-133.

21. Sancar, A.; Lindsey-Boltz, L. A.; Ünsal-Kaçmaz, K.; Linn, S., Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. *Annu. Rev. Biochem.* **2004,** *73*, 39-85.

22. Lorson, C. L.; Hahnen, E.; Androphy, E. J.; Wirth, B., A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. *Proc. Natl. Acad. Sci. U S A* **1999,** *96*, 6307-6311.

23. Coovert, D. D.; Le, T. T.; McAndrew, P. E.; Strasswimmer, J.; Crawford, T. O.; Mendell, J. R.; Coulson, S. E.; Androphy, E. J.; Prior, T. W.; Burghes, A. H., The survival motor neuron protein in spinal muscular atrophy. *Hum. Mol. Genet.* **1997,** *6*, 1205-1214.

24. Crawford, T. O.; Pardo, C. A., The neurobiology of childhood spinal muscular atrophy. *Neurobiol. Dis.* **1996,** *3*, 97-110.

25. Calin, G. A.; Dumitru, C. D.; Shimizu, M.; Bichi, R.; Zupo, S.; Noch, E.; Aldler, H.; Rattan, S.; Keating, M.; Rai, K., Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc. Natl. Acad. Sci. U S A* **2002,** *99*, 15524-15529.

26. Aqeilan, R.; Calin, G. A.; Croce, C., miR-15a and miR-16-1 in cancer: discovery, function and future perspectives. *Cell. Death Differ.* **2010,** *17*, 215-220.

27. Esquela-Kerscher, A.; Trang, P.; Wiggins, J. F.; Patrawala, L.; Cheng, A.; Ford, L.; Weidhaas, J. B.; Brown, D.; Bader, A. G.; Slack, F. J., The let-7 microRNA reduces tumor growth in mouse models of lung cancer. *Cell cycle* **2008,** *7*, 759-764.

28. Li, J.; Yang, H.; Chen, L.; Li, Y.; Zhu, Y.; Dai, Y.; Chen, K.; Ai, J.; Zeng, T.; Mao, X., Establishment and characterization of human non-small cell lung cancer cell lines. *Mol. Med. Rep.* **2012,** *5*, 114-117.

29. Lamichhane, S. R.; Thachil, T.; De Ieso, P.; Gee, H.; Moss, S. A.; Milic, N., Prognostic role of microRNAs in human non-small-cell lung cancer: a systematic review and meta-analysis. *Dis. Markers* **2018,** *2018*.

30. Lu, J.; Zhan, Y.; Feng, J.; Luo, J.; Fan, S., MicroRNAs associated with therapy of non-small cell lung cancer. *Int. J. Biol. Sci.* **2018,** *14*, 390.

31. Akdim, F.; Stroes, E. S.; Sijbrands, E. J.; Tribble, D. L.; Trip, M. D.; Jukema, J. W.; Flaim, J. D.; Su, J.; Yu, R.; Baker, B. F., Efficacy and safety of mipomersen, an antisense inhibitor of apolipoprotein B, in hypercholesterolemic subjects receiving stable statin therapy. *J. Am. Coll. Cardiol.* **2010,** *55*, 1611-1618.

32. Austin, M. A.; Hutter, C. M.; Zimmern, R. L.; Humphries, S. E., Familial hypercholesterolemia and coronary heart disease: a HuGE association review. *Am. J. Epidemiol.* **2004,** *160*, 421-429.

33. Coelho, T., Familial amyloid polyneuropathy: new developments in genetics and treatment. *Curr. Opin. Neurol.* **1996,** *9*, 355-359.

34. Rapezzi, C.; Quarta, C. C.; Riva, L.; Longhi, S.; Gallelli, I.; Lorenzini, M.; Ciliberti, P.; Biagini, E.; Salvi, F.; Branzi, A., Transthyretin-related amyloidoses and the heart: a clinical overview. *Nat. Rev. Cardiol.* **2010,** *7*, 398-408.

35. Prakash, T. P.; Allerson, C. R.; Dande, P.; Vickers, T. A.; Sioufi, N.; Jarres, R.; Baker, B. F.; Swayze, E. E.; Griffey, R. H.; Bhat, B., Positional effect of chemical modifications on short interference RNA activity in mammalian cells. *J. Med. Chem.* **2005,** *48*, 4247-4253.

36. Setten, R. L.; Rossi, J. J.; Han, S.-p., The current state and future directions of RNAi-based therapeutics. *Nat. Rev. Drug Discov.* **2019,** *18*, 421-446.

37. Elbashir, S. M.; Harborth, J.; Lendeckel, W.; Yalcin, A.; Weber, K.; Tuschl, T., Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **2001,** *411*, 494-498.

38. Opalinska, J. B.; Gewirtz, A. M., Nucleic-acid therapeutics: basic principles and recent applications. *Nat. Rev. Drug Discov.* **2002,** *1*, 503-514.

39. Hua, Y.; Vickers, T. A.; Okunola, H. L.; Bennett, C. F.; Krainer, A. R., Antisense masking of an hnRNP A1/A2 intronic splicing silencer corrects SMN2 splicing in transgenic mice. *Am. J. Hum. Genet.* **2008,** *82*, 834-848.

40. Sands, H.; Gorey-Feret, L. J.; Cocuzza, A. J.; Hobbs, F. W.; Chidester, D.; Trainor, G. L., Biodistribution and metabolism of internally 3H-labeled oligonucleotides. I. Comparison of a phosphodiester and a phosphorothioate. *Mol. Pharmacol.* **1994,** *45*, 932-943.

41. Levin, A. A., Treating disease at the RNA level with oligonucleotides. *N. Engl. J. Med.* **2019,** *380*, 57-70.

42. Chiu, Y.-L.; Rana, T. M., siRNA function in RNAi: a chemical modification analysis. *RNA* **2003,** *9*, 1034-1048.

43. Chiu, Y.-L.; Rana, T. M., RNAi in human cells: basic structural and functional features of small interfering RNA. *Mol. Cell* **2002,** *10*, 549-561.

44. Hall, A. H.; Wan, J.; Shaughnessy, E. E.; Ramsay Shaw, B.; Alexander, K. A., RNA interference using boranophosphate siRNAs: structure–activity relationships. *Nucleic Acids Res.* **2004,** *32*, 5991-6000.

45. Shaw, B. R.; Sergueev, D.; He, K.; Porter, K.; Summers, J.; Sergueeva, Z.; Rait, V., Boranophosphate backbone: a mimic of phosphodiesters, phosphorothioates, and methyl phosphonates. *Methods Enzymol.* **2000,** *313*, 226-257.

46. Prakash, T. P., An overview of sugar-modified oligonucleotides for antisense therapeutics. *Chem. Biodivers.* **2011,** *8*, 1616-1641.

47. Passini, M. A.; Bu, J.; Richards, A. M.; Kinnecom, C.; Sardi, S. P.; Stanek, L. M.; Hua, Y.; Rigo, F.; Matson, J.; Hung, G., Antisense oligonucleotides delivered to the mouse CNS ameliorate symptoms of severe spinal muscular atrophy. *Sci. Transl. Med.* **2011,** *3*, 72ra18.

48. van Ravesteyn, T. W.; Dekker, M.; Fish, A.; Sixma, T. K.; Wolters, A.; Dekker, R. J.; te Riele, H. P., LNA modification of single-stranded DNA oligonucleotides allows subtle gene modification in mismatch-repair-proficient cells. *Proc. Natl. Acad. Sci. U S A* **2016,** *113*, 4122-4127.

49. Lan, Q.; Liu, P. Y.; Haase, J.; Bell, J. L.; Hüttelmaier, S.; Liu, T., The critical role of RNA m6A methylation in cancer. *Cancer Res.* **2019,** *79*, 1285-1292.

50. Woo, J.; Meyer Jr, R. B.; Gamper, H. B., G/C-modified oligodeoxynucleotides with selective complementarity: synthesis and hybridization properties. *Nucleic Acids Res.* **1996,** *24*, 2470-2475.

51. Dai, N.; Kool, E. T., Fluorescent DNA-based enzyme sensors. *Chem. Soc. Rev.* **2011,** *40*, 5756-5770.

52. Saito, Y.; Hudson, R. H., Base-modified fluorescent purine nucleosides and nucleotides for use in oligonucleotide probes. *J. Photochem. Photobiol. C* **2018,** *36*, 48-73.

53. Xu, W.; Chan, K. M.; Kool, E. T., Fluorescent nucleobases as tools for studying DNA and RNA. *Nat. Chem.* **2017,** *9*, 1043.

54. Jahnz-Wechmann, Z.; Framski, G. R.; Januszczyk, P. A.; Boryski, J., Base-modified nucleosides: Etheno derivatives. *Front Chem.* **2016,** *4*, 19.

55. Patchornik, A.; Amit, B.; Woodward, R., Photosensitive protecting groups. *J. Am. Chem. Soc.* **1970,** *92*, 6333-6335.

56. Barltrop, J.; Plant, P.; Schofield, P., Photosensitive protective groups. *Chem. Commun. (London)* **1966**, 822-823.

57. Barton, D.; Chow, Y.; Cox, A.; Kirby, G., 654. Photochemical transformations. Part XIX. Some photosensitive protecting groups. *J. Chem. Soc.* **1965**, 3571-3578.

58. Mayer, G.; Heckel, A., Biologically active molecules with a "light switch". *Angew. Chem. Int. Ed. Engl.* **2006,** *45*, 4900-4921.

59. Menge, C.; Heckel, A., Coumarin-caged dG for improved wavelength-selective uncaging of DNA. *Org. Lett.* **2011,** *13*, 4620-4623.

60. Schäfer, F.; Joshi, K. B.; Fichte, M. A.; Mack, T.; Wachtveitl, J.; Heckel, A., Wavelength-selective uncaging of dA and dC residues. *Org. Lett.* **2011,** *13*, 1450-1453.

61. Shah, S.; Rangarajan, S.; Friedman, S. H., Light-activated RNA interference. *Angew. Chem. Int. Ed. Engl.* **2005,** *44*, 1328-1332.

62. Blidner, R. A.; Svoboda, K. R.; Hammer, R. P.; Monroe, W. T., Photoinduced RNA interference using DMNPE-caged 2′-deoxy-2′-fluoro substituted nucleic acids in vitro and in vivo. *Mol. Biosyst.* **2008,** *4*, 431-440.

63. Nguyen, Q. N.; Chavli, R. V.; Marques, J. T.; Conrad II, P. G.; Wang, D.; He, W.; Belisle, B. E.; Zhang, A.; Pastor, L. M.; Witney, F. R., Light controllable siRNAs regulate gene suppression and phenotypes in cells. *Biochim. Biophys. Acta.* **2006,** *1758*, 394-403.

64. Mikat, V.; Heckel, A., Light-dependent RNA interference with nucleobase-caged siRNAs. *RNA* **2007,** *13*, 2341-2347.

65. Chaulk, S. G.; MacMillan, A. M., Synthesis of oligo-RNAs with photocaged adenosine 2′-hydroxyls. *Nat. Protoc.* **2007,** *2*, 1052-1058.

66. Chaulk, S. G.; MacMillan, A. M., Caged RNA: photo-control of a ribozyme reaction. *Nucleic Acids Res.* **1998,** *26*, 3173-3178.

67. Behlke, M. A., Chemical modification of siRNAs for in vivo use. *Oligonucleotides* **2008,** *18*, 305-320.

68. Diebold, S. S.; Kaisho, T.; Hemmi, H.; Akira, S.; e Sousa, C. R., Innate antiviral responses by means of TLR7-mediated recognition of single-stranded RNA. *Science* **2004,** *303*, 1529-1531.

69. Heil, F.; Hemmi, H.; Hochrein, H.; Ampenberger, F.; Kirschning, C.; Akira, S.; Lipford, G.; Wagner, H.; Bauer, S., Species-specific recognition of single-stranded RNA via toll-like receptor 7 and 8. *Science* **2004,** *303*, 1526-1529.

70. Hornung, V.; Guenthner-Biller, M.; Bourquin, C.; Ablasser, A.; Schlee, M.; Uematsu, S.; Noronha, A.; Manoharan, M.; Akira, S.; de Fougerolles, A., Sequence-specific potent induction of IFN-α by short interfering RNA in plasmacytoid dendritic cells through TLR7. *Nat. Med.* **2005,** *11*, 263-270.

71. Hartmann, G.; Weeratna, R. D.; Ballas, Z. K.; Payette, P.; Blackwell, S.; Suparto, I.; Rasmussen, W. L.; Waldschmidt, M.; Sajuthi, D.; Purcell, R. H., Delineation of a CpG phosphorothioate oligodeoxynucleotide for activating primate immune responses in vitro and in vivo. *J. Immunol.* **2000,** *164*, 1617-1624.

72. Lund, J. M.; Alexopoulou, L.; Sato, A.; Karow, M.; Adams, N. C.; Gale, N. W.; Iwasaki, A.; Flavell, R. A., Recognition of single-stranded RNA viruses by Toll-like receptor 7. *Proc. Natl. Acad. Sci. U S A* **2004,** *101*, 5598-5603.

73. Alexopoulou, L.; Holt, A. C.; Medzhitov, R.; Flavell, R. A., Recognition of double-stranded RNA and activation of NF-κB by Toll-like receptor 3. *Nature* **2001,** *413*, 732-738.

74. Sioud, M., Induction of inflammatory cytokines and interferon responses by double-stranded and single-stranded siRNAs is sequence-dependent and requires endosomal localization. *J. Mol. Biol.* **2005,** *348*, 1079-1090.

75. Judge, A. D.; Sood, V.; Shaw, J. R.; Fang, D.; McClintock, K.; MacLachlan, I., Sequence-dependent stimulation of the mammalian innate immune response by synthetic siRNA. *Nat. Biotechnol.* **2005,** *23*, 457-462.

76. Robbins, M.; Judge, A.; Liang, L.; McClintock, K.; Yaworski, E.; MacLachlan, I., 2′-O-methyl-modified RNAs act as TLR7 antagonists. *Mol. Ther.* **2007,** *15*, 1663-1669.

77. Diebold, S. S.; Massacrier, C.; Akira, S.; Paturel, C.; Morel, Y.; Reis e Sousa, C., Nucleic acid agonists for Toll-like receptor 7 are defined by the presence of uridine ribonucleotides. *Eur. J. Immunol.* **2006,** *36*, 3256-3267.

78. Judge, A.; Maclachlan, I., Overcoming the innate immune response to small interfering RNA. *Hum. Gene. Ther.* **2008,** *19*, 111-124.

79. Morrissey, D. V.; Lockridge, J. A.; Shaw, L.; Blanchard, K.; Jensen, K.; Breen, W.; Hartsough, K.; Machemer, L.; Radka, S.; Jadhav, V., Potent and persistent in vivo anti-HBV activity of chemically modified siRNAs. *Nat. Biotechnol.* **2005,** *23*, 1002-1007.

80. Judge, A. D.; Bola, G.; Lee, A. C.; MacLachlan, I., Design of noninflammatory synthetic siRNA mediating potent gene silencing in vivo. *Mol. Ther.* **2006,** *13*, 494-505.

81. Huang, Y.; Hong, J.; Zheng, S.; Ding, Y.; Guo, S.; Zhang, H.; Zhang, X.; Du, Q.; Liang, Z., Elimination pathways of systemically delivered siRNA. *Mol. Ther.* **2011,** *19*, 381-385.

82. Mishra, S.; Heidel, J. D.; Webster, P.; Davis, M. E., Imidazole groups on a linear, cyclodextrin-containing polycation produce enhanced gene delivery via multiple processes. *J. Control Release* **2006,** *116*, 179-191.

83. Davis, M.; Pun, S.; Bellocq, N.; Reineke, T.; Popielarski, S.; Mishra, S.; Heidel, J., Self-assembling nucleic acid delivery vehicles via linear, water-soluble, cyclodextrin-containing polymers. *Curr. Med. Chem.* **2004,** *11*, 179-197.

84. Kanasty, R.; Dorkin, J. R.; Vegas, A.; Anderson, D., Delivery materials for siRNA therapeutics. *Nat. Mater.* **2013,** *12*, 967-977.

85. Rozema, D. B.; Lewis, D. L.; Wakefield, D. H.; Wong, S. C.; Klein, J. J.; Roesch, P. L.; Bertin, S. L.; Reppen, T. W.; Chu, Q.; Blokhin, A. V., Dynamic PolyConjugates for targeted in vivo delivery of siRNA to hepatocytes. *Proc. Natl. Acad. Sci. U S A* **2007,** *104*, 12982-12987.

86. Esko, J. D.; Kimata, K.; Lindahl, U., *Essentials of glycobiology*. 2009; p 321-33.

87. Yin, W.; Rogge, M., Targeting RNA: a transformative therapeutic strategy. *Clin. Transl. Sci.* **2019,** *12*, 98-112.

88. Crooke, S. T.; Witztum, J. L.; Bennett, C. F.; Baker, B. F., RNA-targeted therapeutics. *Cell Metab.* **2018,** *27*, 714-739.

89. Springer, A. D.; Dowdy, S. F., GalNAc-siRNA conjugates: leading the way for delivery of RNAi therapeutics. *Nucleic Acid Ther.* **2018,** *28*, 109-118.

90. Adams, D.; Gonzalez-Duarte, A.; O'Riordan, W. D.; Yang, C.-C.; Ueda, M.; Kristen, A. V.; Tournev, I.; Schmidt, H. H.; Coelho, T.; Berk, J. L., Patisiran, an RNAi therapeutic, for hereditary transthyretin amyloidosis. *N. Engl. J. Med.* **2018,** *379,* 11-21.

91. Rigo, F.; Hua, Y.; Krainer, A. R.; Bennett, C. F., Antisense-based therapy for the treatment of spinal muscular atrophy. *J. Cell Biol.* **2012,** *199*, 21-25.

92. Dugave, C.; Demange, L., Cis− trans isomerization of organic molecules and biomolecules: implications and applications. *Chem. Rev.* **2003,** *103*, 2475-2532.

93. Benson, R., N, N-DISUBSTITUTED-1-AMINO-7-IMINO-1, 3, 5-CYCLOHEPTATRIENES, A NON-CLASSICAL AROMATIC SYSTEM. *J. Am. Chem. Soc.* **1960,** *82*, 5948-5949.

94. Brasen, W.; Holmquist, H.; Benson, R., N, N'-Disubstituted-1-amino-7-imino-1, 3, 5-cycloheptatrienes, a Non-classical Aromatic System. *J. Am. Chem. Soc.* **1961,** *83*, 3125-3135.

CHAPTER 2

Part-A

Synthesis and Biochemical Evaluation of Site-specific 2'-*O*-(*o*-nitrobenzyoxy)ethyl-uridine (U-NBnOE) modified *caged* siRNA for RNA *interference*

.

Chapter 2A. Synthesis and Biochemical Evaluation of Site-specific 2'-*O*-(*o*-nitrobenzyoxy)ethyl-uridine (U-NBnOE) modified *caged* siRNA for RNA *interference*

2A.1 Introduction

RNA interference (RNA*i*) is a biological process which involves 21-23 nucleotide-long double-stranded RNAs, named as micro-RNA (*mi*RNA) in regulating protein coding genes.[1-3] The biogenesis of miRNA in controlling the protein expression sets a foundation for the development of RNA based therapeutics. For therapeutic applications, RNAi activity can also have accomplished by using rationally designed synthetic *small-interfering* RNA (siRNA) for target-gene control.[4, 5] The mechanism of RNAi activity involves the degradation of specific mRNA with complementary strand (*antisense* strand) of siRNA and RNA induced silencing complex (RISC) protein and always cut at 10^{th} and 11^{th} nucleotide from siRNA's 5'end.[4, 6] For binding and processing of siRNA with target mRNA in RISC 5'-OH of antisense strand in siRNA should be free for 5'-OH phosphorylation. Moreover, A-type of helix is required for siRNA to enhance RNAi activity.[7, 8] The synthesis and application of unmodified siRNA is challenging and particularly, they are prone to *exo/endo* nucleases mediated degradations and they have low cell permeability.

To increase the RNAi efficacy and stability of siRNAs, chemical modifications to siRNA are employed. These chemical modifications to siRNA are employed to protect from *exo/endo* enzyme degradation, enhance cell permeability, and diminish off-target effects. The thiorylated phosphate backbone modification to siRNA improved stability against exo/endo enzyme degradations and enhanced cell permeability.[9] The chemical modifications at 2'-*OH* group of ribose-sugar ring in

siRNA also improved the structural and conformational requirement of A-type of helix without disturbing the Watson-Crick base pairing.[10] The 2'-modified siRNAs, such as *2'-tethered* nucleoside derivatives (-OCH_2CH_2OMe/-$OCH_2CH_2NH_2$, -$OCH_2CH_2N_3$) are synthesized and investigated their role in RNAi activity[11] These analogues have shown inspiring results especially in synthesis of stable/cell-permeable siRNA and fluorophore labeled siRNA. The chemical modifications such as 2'-*O*-methyloxyethyl (2'-*O*-MOE), 2'-*O*-methyl, (2'-*O*-Me) and Locked Nucleic Acids (LNA) are exemplified in successful RNA*i* therapeutic applications.[9, 12, 13]

To contol the RNAi acitivity, the siRNAs are also modified with photo labile groups (*caged* siRNA), which temporarily block the interaction with RNAi enzyme machinery. For example, Friedman and co-workers demonstrated the light activated RNAi activity with non-specific modification of siRNA with photo labile group (4,5-dimethoxy-2-nitrophenylethyl, DMNPE group).[14] McMaster and co-workers temporarily blocked the 5'-phosphate of antisense strand in siRNA with photo labile group, which is required for RNAi activity, and controlled the binding of siRNA and RISC with light.[15] Heckel and Co-workers synthesized the site-specific photo labile siRNAs (*caged* siRNAs), which temporarily blocked the Watson-Crick hydrogen bonding interface of siRNA:mRNA duplex within the RISC complex.[16] Moreover, 2'-*O*H of RNA is caged with photo labile group (*o*-nitrobenzyl) for controlling the cleaving properties of hammerhead ribozyme with light.[17, 18]

2A.2 Objective of our work

From concise overview of chemically modified siRNAs reveals that the 2'-tethered siRNAs and photo labile siRNAs (caged siRNAs) are employed for the regulation of specific-gene

expression through RNAi pathway. Inspiring from these studies, we designed site-specific *caged* siRNAs with 2'-tether alcohol capped with photo labile protecting group. We hypothesized that the introduction of bulky PLPG at tethered alcohol of siRNA could block the RNAi activity temporarily because of significant changes in the conformation of sugar ring/nucleobase, due to steric effects. This could result in temporarily blocking of the RNAi pathway. The binding affinity of those *caged*-siRNAs with RISC complex/complimentary mRNA can be regulated with light. Herein, we have planned to synthesize the *caged*-siRNA, by caging 2'-tether alcohol of uridine with bulky PLPG, *o*-nitrobenzyl group, and then employed for the light dependent controllable GFP's gene expression in HEK293T cell lines (Figure 2A.1).

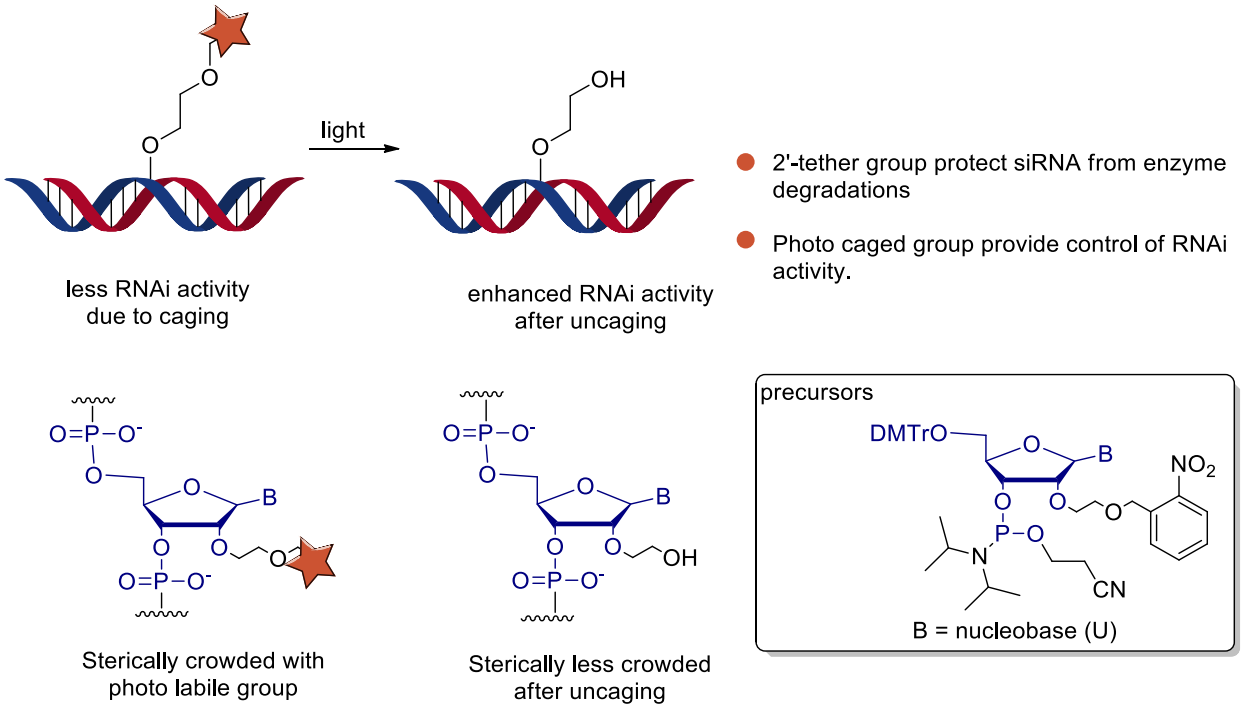


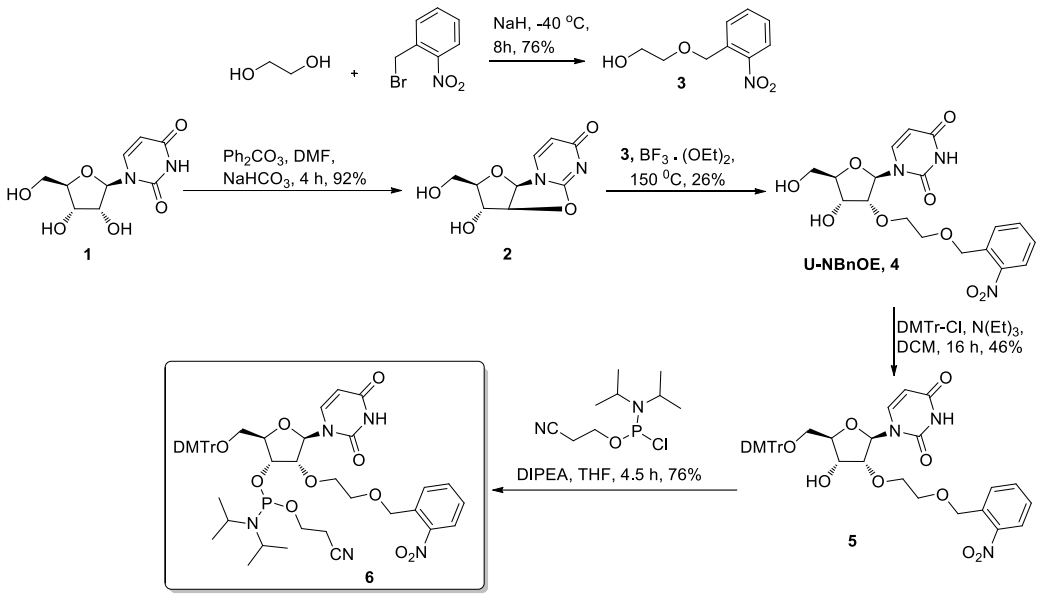
Figure 2A.1. Proposed hypothesis of light controllable RNAi with caged siRNA and designed nucleoside precursor.

Challenges in proposed hypothesis. After overcoming the synthetic challenges, the major problem we can anticipate that most of the 2'-*O*- modification itself can show RNAi activity. And 2'-*O*- modified caged siRNAs within the sense and antisense strands may also have RNAi activity with photo labile group itself (RNAi leakage). [19-21]

2A.3 Results and Discussion

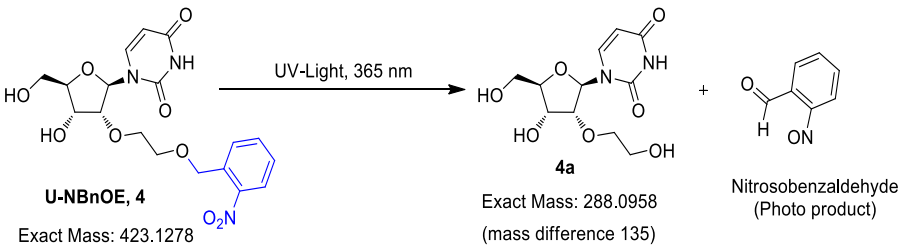
We started the synthesis of designed *caged*-uridine phosphoramidite from commercially available Uridine nucleoside (**1**). The synthesis is outlined in Scheme 2A.1, Uridine (**1**) was converted into reactive cyclic intermediate (**2**) by following previously reported procedure.[22] Intermediate **2** was treated with *o*-nitrobenzyloxyethanol **3** for nucleophilic addition reaction at C-2 of ribose sugar ring. As resultant, 2'-*O*-(*o*-nitrobenzyloxyethyl)-uridine nucleoside **4** (*caged* nucleoside, U-NBnOE nucleoside) was prepared. We recorded, the ^{1}H-^{1}H COSY and ^{1}H-^{1}H NOESY experiments and assigned all the protons in caged-nucleoside (U-NBnOE) **4**. The *o*-nitrobenzyloxyethanol **3** was prepared by following previously reported procedure.[23] The caged-nucleoside (U-NBnOE) **4** was subjected to the protection of its 5'-hydroxyl group (5'-OH) by treating with DMTr-Cl, which produced DMTr-derivative **5**. For phosphorylation at 3'-OH, DMTr-derivative **5** was treated with 2-Cyanoethyl *N,N*-diisopropylchlorophosphoramidite to produce desired *caged*-uridine phosphoramidite (**6**). The characterization data (^{1}H-/^{13}C-/^{1}H-^{1}H COSY and ^{1}H-^{1}H NOESY NMR spectra and ESI-HRMS) are provided in Appendix.

Scheme 2A.1. Synthesis of 2'-*caged* uridine phosphoramidite



Cleavage of PLPG with light. Before incorporating U-NBnOE nucleoside into siRNA, we tested and optimized the photo cleavage (uncaging) of photo labile protecting group (*o*-nitrobenzyl group) in U-NBnOE nucleoside **4** with exposure of light (monochromatic UV light, 365 nm). This photo cleavage reaction is monitored by ^{1}H-NMR and Mass techniques. The anticipated reaction products after photo cleavage are provided in Scheme 2A.2.

Scheme 2A.2. Photo cleavage of U-NBnOE nucleoside **4** with light (365 nm)



The time dependent ^{1}H-NMR experiments of U-NBnOE nucleoside **4** with exposure of the light (365 nm) at different time intervals is provided in Figure 2A.2. Before exposure of light (0 min),

the chemical shift (δ) of *o*-nitrobenzyl -*C**H**₂*- appears as a singlet at δ4.9 ppm. With increasing the exposure time, subsequently -*C**H**₂*- of *o*-nitrobenzyl group disappear, and a new peak at δ11.3 ppm appears. The peak at 11.3 ppm belongs to aldehyde proton of nitrosobenzyladehyde. In literature the photo cleavage of *o*-nitrobenzyl group is extensively studied, the photo cleavage of *o*-nitrobenzyl group occurs with the disappearance of *o*-nitro*benzylic* -*C**H**₂*- and known to form *o*-nitrosobenzaldehyde. After the photo labile group cleavage, we isolated and characterized the uncaged nucleoside **4a** by ^{1}H-/^{13}C-/^{1}H-^{1}H COSY and ^{1}H-^{1}H NOESY NMR experiments and assigned all the protons in **4a** (NMR spectra provided in Appendix).

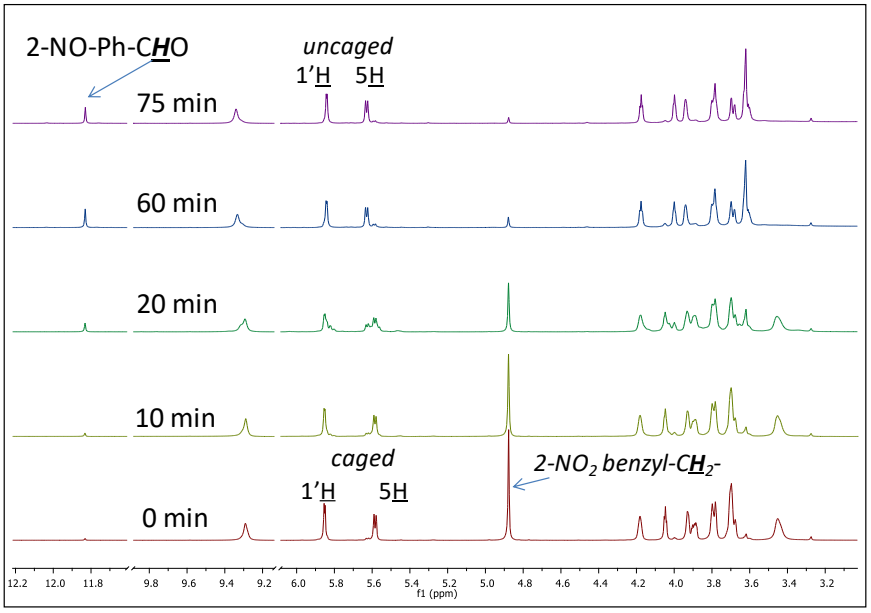


Figure 2A.2. Photo cleavage of U-NBnOE nucleoside **4** by ^{1}H-NMR at different exposure times in CD_3CN at 365 nm.

Since the presence of bulkier caged group (*o*-nitrobenzyl group) in U-NBnOE nucleoside **4**, the conformation of U-NBnOE nucleoside **4** could affect when compared to the conformation of uncaged nucleoside **4a**. The ^{1}H-NMR spectra before and after removal of photo labile group is shown in Figure 2A.3, we observed that only small conformational changes are observed before

(*caged*) and after removal (*uncaged*) of photo labile group, but not characteristic changes. This indicates that the presence of bulky PLPG group (*o*-nitrobenzyl group) is not affecting the conformation of nucleoside monomer.

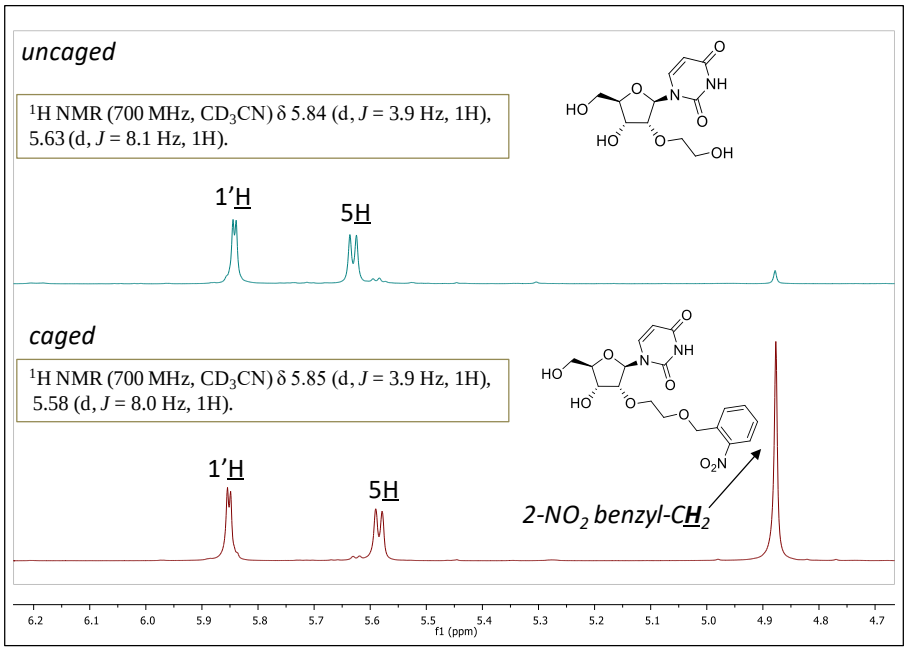


Figure 2A.3. Chemical shifts and coupling constants for 1'H and 5H protons in U-NBnOE nucleoside **4** before (*caged*) and after (*uncaged*) exposure of light (365 nm) in CD_3CN.

The photo cleavage of reaction of U-NBnOE nucleoside **4** was also monitored by ESI-MS, mass spectra, before and after irradiation with light (365 nm) is provided in Figure 2A.4. The mass spectra of U-NBnOE nucleoside **4** exhibit mass peak at m/z 446.119 [M+Na^+] before exposure, which belongs to the *caged* U-NBnOE nucleoside **4** (Figure 2A.4A). After 40 min exposure of light (365 nm), the mass spectra of **4** exhibit a new peak at m/z 311.0874, which belongs to the molecular ion peak of *uncaged* nucleoside **4a** (Figure 2A.4B). These NMR/Mass experimental analyses confirm the photo cleavage of *o*-nitrobenzyl group from U-NBnOE nucleoside **4**.

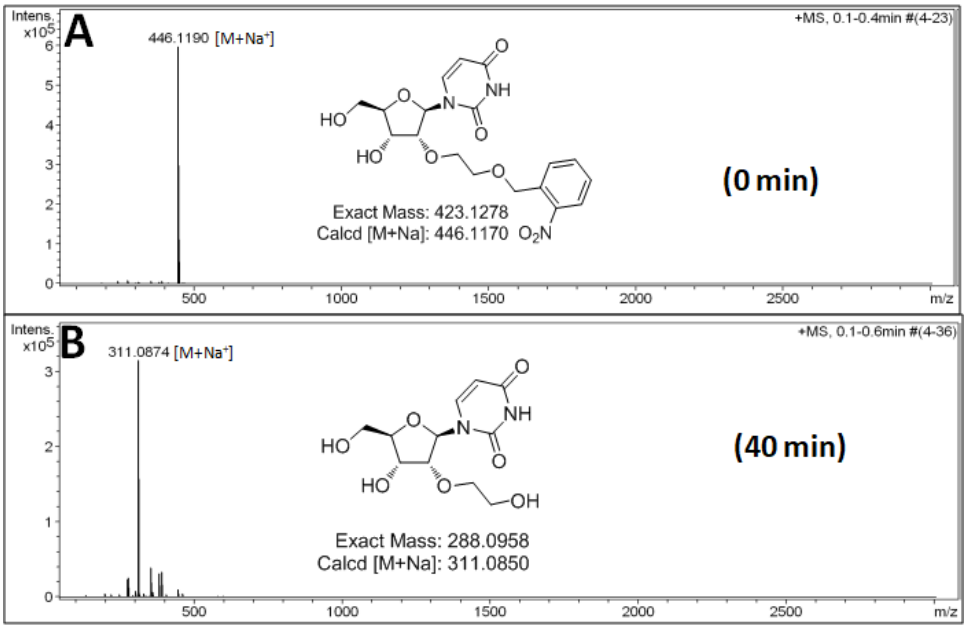


Figure 2A.4. ESI-Mass spectra of U-NBnOE nucleoside **4** before, *caged* (A) and after, *uncaged* (B) with exposure of light at 365 nm.

Caged U-NBnOE RNA synthesis. After optimizing the reaction conditions for the photo cleavage, of U-NBnOE phosphoramidite (**6**), we planned to incorporate U-NBnOE residue in siRNA which target the GFP reporter mRNA. By using, ABI-394 DNA/RNA synthesizer, we incorporated the U-NBnOE phosphoramidite (**6**) into a series of sense (S) and antisense (AS) RNA single strands which target the designed EGFP reporter mRNA. The RNAs were synthesized under standard RNA synthesis conditions, except the coupling time of U-NBnOE phosphoramidite (**6**) is increased to 18 min. We also synthesized unmodified sense and antisense RNA strands for control studies. After solid phase synthesis, all the RNA oligonucleotides were cleaved from resin with ammonium hydroxide, 40% aqueous methyl amine (1:1, v/v). The 2'-*O*- Silyl (TBDMS/TOM) protection groups were removed with TEA.HF and purified by preparative 20% PAGE. After extracting the RNAs from gel and desaltations, we characterized these RNAs by ESI-MS; their mass spectra are provided in the Appendix. The sequence of synthesized GFP siRNA single strands, position of U-

NBnOE modifications are provided in Table 2A.1, (Column 3). In the mass spectra of RNA oligonucleotides, the mass of AS1 and S1-S6 RNAs (Table 2A.1, Entry 1-2 and 5-9) appeared in charge states z = 5 and z = 6, while the mass of RNA oligonucleotides, AS2 and AS3 (Table 2A.1, Entry 3-4) appeared from charge state z = 5 to z = 10 with Na^+/K^+ adducts; the deconvoluted mass data is provided in Table 2A.1 (Column 5). The observed molecular mass of RNAs is almost equal to their calculated mass (Entry 1-9, Table 2A.1). For example, the calculated mass of RNA, S2-S5 (one U-NBnOE incorporated) is 6769.13 (m/z), and their observed mass is 6770.1 (m/z) (Table 2A.1, Entry 5-8). The observed mass difference (~8 units) is probably due to the protonation of respective RNAs, because these RNAs were equilibrated in 0.1% formic acid in acetonitrile, water (1:1) mixture for ESI-MS analysis. In literature, the presence of small amount of acid will improve the mass sensitivity of RNA oligonucleotides and to some extent reduce the formation of charge states and adducts.[24] Similar mass results were obtained for other synthesized RNAs. These mass analyses confirm the successful synthesis of U-NBnOE containing RNA oligonucleotides.

Table 2A.1. Unmodified RNA and 2'-modified caged-RNA (21-mer)

Entry	RNA	RNA Sequence (5'-----3')	Calculated mass (*M*)	Observed mass
1	AS1	UCC UUG AAG AAG AUG GUG CTT	6710.06	6741.01, $[M + 1Na^+ + 8H^+]^{-1}$
2	S1	GCA CCA UCU UCU UCA AGG ATT	6590.06	6597.83, $[M + 8H^+]^{-1}$
3	AS2	UCC UUG AAG AAG A**X**G GUG CTT	6889.23	7020.61, $[M + 4Na^+ + 1K^+]^{-1}$
4	AS3	UCC UUG AAG AAG AUG G**X**G CTT	6889.23	7020.62, $[M + 4Na^+ + 1K^+]^{-1}$
5	S2	GCA CCA **X**CU UCU UCA AGG ATT	6769.13	6777.09, $[M + 8H^+]^{-1}$
6	S3	GCA CCA UCU UC**X** UCA AGG ATT	6769.13	6777.08, $[M + 8H^+]^{-1}$
7	S4	GCA CCA UCU UCU **X**CA AGG ATT	6769.13	6777.07, $[M + 8H^+]^{-1}$
8	S5	GCA CCA UCU **X**CU UCA AGG ATT	6769.13	6777.10, $[M + 8H^+]^{-1}$
9	S6	GCA CCA **X**CU UCU **X**CA AGG ATT	6948.21	6957.30, $[M + 8H^+]^{-1}$

S: Sense strand; AS: Antisense strand; X: 2'-modify caged-Uridine incorporated.

Photo cleavage of o-nitrobenzyl groups (uncaging) from U-NBnOE RNAs. After successful synthesis of U-NBnOE RNA oligonucleotides, we studied the photo cleavage of photo labile RNA

sense strands by ESI-MS. The presence of photo labile *o*-nitrobenzyl group adds a mass of m/z 136 to caged RNA. When the photo labile group is removed by exposure of light, there is a reduction in mass of m/z 136 to *uncage* the RNA for each unit of *o*-nitrobenzyl group, as shown in Figure 2A.5.

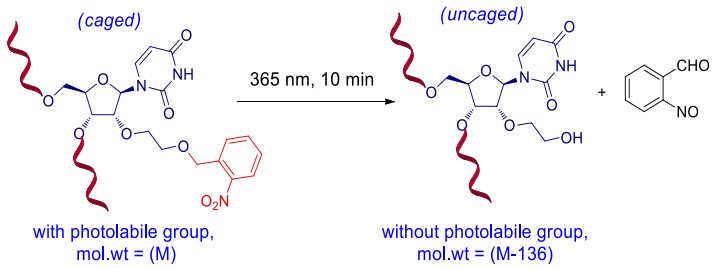


Figure 2A.5. Anticipated photo cleavage mass differences in U-NBnOE modified nucleoside before and after treatment of light.

Each desalted RNA oligonucleotide (500 pmol) was exposed at 365 nm for 10 min. The ESI-MS measurements were carried out before and after exposure of light for U-NBnOE RNA sense strands (S2-S6) and control RNAs (S1, AS1). Within 10 min of exposure of light (365 nm) the photo labile groups (o-nitrobenzyl group) were cleaved completely. The mass spectra of U-NBnOE RNA oligonucleotides after exposure of light appeared mostly in charge state z = 5, their mass spectra are provided in APPENDIX and deconvoluted mass data is provided in Table 2A.2, Column 5. In RNA oligonucleotides, S2/S3/S4/S5 have a single photo labile group (*caged*) which are differ in their position of modification, have mass m/z 6777 before exposure to light, after exposure to light, this mass is reduced to m/z 6641 (Table 2A.2, Entry 3-6). The observed mass difference of m/z 136 is equivalent to the *o*-nitrobenzyl group. These mass results confirm the photo cleavage on nitrobenzyl group from U-NBnOE RNA sense stands. Whereas U-NBnOE RNA, S6 has two U-NBnOE modifications and have a mass difference of m/z 272 before and after

exposure of light. The mass difference of m/z 272 is due to the two *o*-nitrobenzyl groups (136x2 units), Table 2A.2, Entry 7. While control RNA, S1 and AS1 has no significant changes in their mass (Table 2A.2, Entry 1-2), this is expected because they don't have any photo labile groups. These mass results indicate the presence of photolabile group in synthesized U-NBnOE RNA oligonucleotides and undergo complete photo cleavage with exposure of light (at 365 nm) within 10 min.

Table 2A.2. ESI-MS mass data before and after exposure to light at 365 nm for 10 min.

Entry	RNA	RNA sequence	Observed mass (0 min)	Observed mass (365 nm, 10 min)	Difference in mass
1	AS1	5' UCC UUG AAG AAG AUG GUG CTT 3'	6741.01	6717.19	26.8 (Na^+)
2	S1	5'-GCA CCA UCU UCU UCA AGG ATT-3'	6597.83	6597.05	0.8
3	S2	5'-GCA CCA **X**CU UCU UCA AGG ATT-3'	6777.09	6641.05	136.04
4	S3	5'-GCA CCA UCU UC**X** UCA AGG ATT-3'	6777.08	6641.05	136.03
5	S4	5'-GCA CCA UCU UCU **X**CA AGG ATT-3'	6777.07	6641.05	136.01
6	S5	5'-GCA CCA UCU **X**CU UCA AGG ATT-3'	6777.10	6641.10	135.99
7	S6	5'-GCA CCA **X**CU UCU **X**CA AGG ATT-3'	6957.30	6685.13	272.17 (136.08 * 2)

X: U-NBnOE modified nucleoside.

Photo cleavage study of PLPG in U-NBnOE RNA, S6, by UV-Visible spectroscopy. After conformation of photo cleavage (*uncaging*) of nitrobenzyl group in U-NBnOE RNAs, we performed time dependent UV-Visible experiment with similar conditions for at least U-NBnOE RNA S6 at different exposure times. First, the UV spectra of U-NBnOE RNA, S6 is recorded before exposure of light (0 min), then U-NBnOE RNA S6 is irradiated with light (365 nm) and recorded UV spectra for every 5 min (Figure 2A.6). Before exposure of light (0 min) S6 RNA has absorption at λ260 nm. However, after exposure of light, there are no significant changes in its

UV spectra. Though S6 RNA has two photo labile (nitrobenzyl) groups, we could not notice any major changes in their UV-spectra after *uncaging*. We assume that, due to the strong RNA nucleobase absorption overlapping with comparatively weak absorption of nitrobenzyl chromophore and its photoproduct (*o*-nitrosobenzaldehyde) before and after photo cleavage.

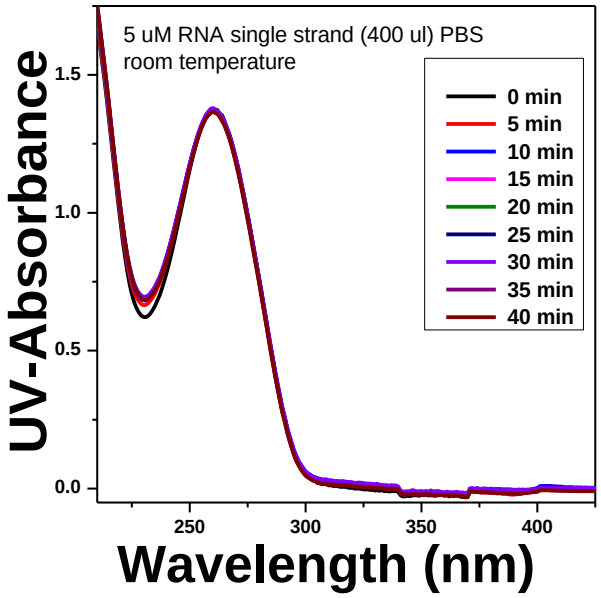


Figure 2A.6. UV- absorbance spectra of RNA S6, 5'-GCA CCA XCU UCU XCA AGG ATT-3', 5 μm in PBS treated with 365 nm with 5 min interval for 50 min.

Caged siRNA duplexes. After successful synthesis and characterization of siRNA single strands, siRNA duplexes were prepared by annealing equimolar concentrations of complementary sense (RNA S1-S6) and antisense (RNA AS1-AS3) single strands. The sequence and position of modification within the sense and antisense strands of *caged*-siRNAs (siRNA-2 to siRNA-8) and control siRNAs (siRNA-1) are provided in Table 2A.3.

Table 2A.3. Sequences of caged and control siRNA duplexes

Entry	Code	siRNA duplex	modified strand[a]	modification site from 5'-end
1	siRNA-1	5'-UCC UUG AAG AAG AUG GUG C TT-3 ' 3'-TT AGG AAC UUC UUC UAC CAC G-5'	no	no
2	siRNA-2	5'-UCC UUG AAG AAG A**X**G GUG C TT-3 ' 3'-TT AGG AAC UUC UUC UAC CAC G-5'	AS	14
3	siRNA-3	5'-UCC UUG AAG AAG AUG G**X**G C TT-3 ' 3'-TT AGG AAC UUC UUC UAC CAC G-5'	AS	17
4	siRNA-4	5'-UCC UUG AAG AAG AUG GUG C TT-3 ' 3'-TT AGG AAC UUC **X**UC UAC CAC G-5	S	10
5	siRNA-5	5'-UCC UUG AAG AAG AUG GUG C TT-3 ' 3'-TT AGG AAC U**X**C UUC UAC CAC G-5'	S	12
6	siRNA-6	5'-UCC UUG AAG AAG AUG GUG C TT-3 ' 3'-TT AGG AAC **X**UC UUC UAC CAC G-5'	S	13
7	siRNA-7	5'-UCC UUG AAG AAG AUG GUG C TT-3 ' 3'-TT AGG AAC UUC UUC **X**AC CAC G-5'	S	7
8	siRNA-8	5'-UCC UUG AAG AAG AUG GUG C TT-3 ' 3'-TT AGG AAC **X**UC UUC **X**AC CAC G-5'	S	7, 13

X = photo caged uridine, [a] S = sense strand, AS = antisense strand

CD studies of siRNA duplexes. CD spectroscopy is widely used to study the DNA/RNA duplexes structures. It has been noticed that siRNA reportedly forms characteristic A-type duplex, which contains CD signals at 240 nm (minima) and 250-290 nm (maxima).[25] The formation of A-type duplex is critical for binding with RISC protein for RNAi activity.[8] We also investigated the duplex formation by control siRNA (siRNA-1) and U-NBnOE caged-siRNA (siRNA-2 to siRNA-8) by CD spectroscopy. Their CD spectra are provided in Figure 2A.7. The CD spectra of siRNA-1 (control) exhibit typical siRNA duplex CD structure. Which indicate the formation of A-type RNA duplex structure with positive Cotton effect bands (maxima) at λ267 nm/ λ225 nm and negative Cotton effect band (minima) at λ210 nm/ λ240 nm (Figure 2A.7). While the CD signals of U-NBnOE incorporated antisense strand, *caged* siRNA-2 is also similar to control siRNA-1, but to some extent perturbed from A-type of duplex (from λ250-210 nm), (Figure 2A.7). The CD spectra of U-NBnOE incorporated antisense strand, *caged* siRNA-3 and U-NBnOE incorporated sense

strand siRNAs (siRNA4-8) are depicted in Figure 2A.7B/C, and the CD spectra of these siRNAs are also similar to the control siRNA-1. Typically, all the CD spectra of U-NBnOE incorporated *caged*-siRNAs (antisense and sense strand modified), indicate the formation of A-type RNA duplex structure with positive Cotton effect bands at ~λ260/λ225 nm and negative Cotton effect bands at ~λ210/λ240 nm. From these CD studies, we assume that the presence of *caged* U-NBnOE modification is not affecting the conformation of siRNA duplex structure.

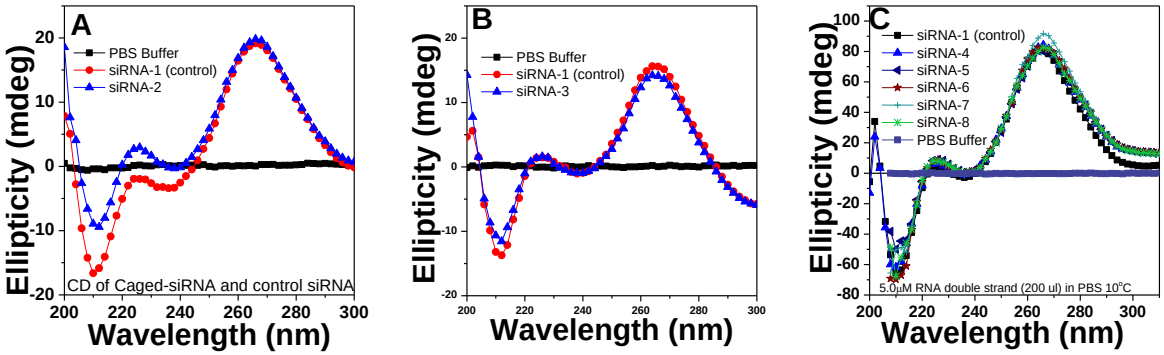


Figure 2A.7. CD spectra of control si-RNA (S:AS), caged-siRNA (S*-AS), UV-exposed caged-siRNA (at different interval of time): 14th position of sense strand (A) and 17th position of sense strand.

To examine the conformational changes of caged siRNAs after removal of *o*-nitrobenzyl group, we recorded the CD spectra of siRNAs (siRNA-2 to siRNA-8) with photo labile group (*caged* siRNA) and without photo labile group (*uncaged* siRNA). The CD spectra of siRNA-2/siRNA-3 are depicted in Figure 2A.8A/B, while the CD spectra of siRNA-4/5/6/7/8 are provided in Appendix. The CD spectra of *caged* siRNA-2, before exposure of light (0 min) is slightly perturbed from the CD spectra of control siRNA-1. But after exposure of light (365 nm, 5 min), the CD spectra of *uncaged* siRNA-2 is almost similar to unmodified control siRNA-1 (Figure 2A.8A). The

CD spectra of siRNA-2 have no characteristics changes before (*caged*) and after (*uncaged*) exposure of light (365 nm, 5 min), however, the CD spectra of both caged and uncaged siRNA-2 are similar to the CD spectra of control siRNA-1 (Figure 2A.8B). Similarly, in all the other caged siRNAs (siRNA-4/5/6/7/8), the CD structure is similar to the control unmodified siRNA before and after exposure of light (365 nm, 5 min), their CD spectra are provided in Appendix. These CD studies infer that all caged and uncaged siRNAs adopt A-type of double helical structures, which is required for RNA*i* activity and caging and uncaging of photo labile group (*o*-nitrobenzyl group) is not interfering with the siRNA conformation.

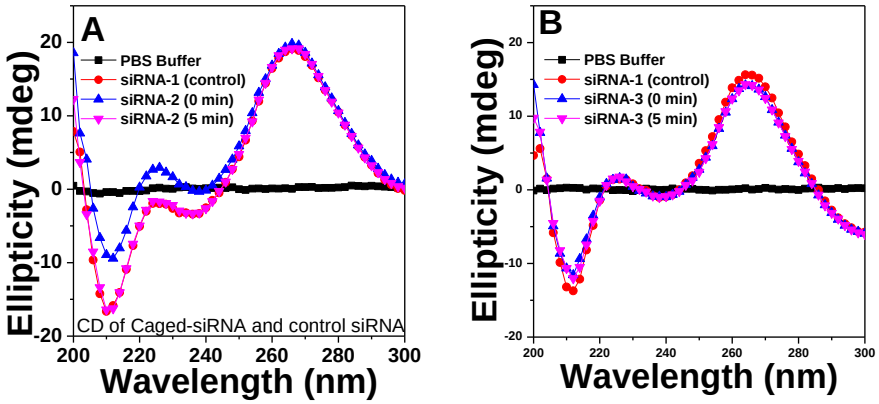


Figure 2A.8. CD spectra of U-NBnOE modified caged siRNA-2/3 along with control siRNA-1 at different exposure times (365 nm)

RNAi studies. To test the hypothesis of light controllable *caged* siRNAs for *RNAi* activity in regulating gene expression, we choose green fluorescence protein (GFP) reporter assay as reliable model system. We investigated this model system in HEK293T cell lines. We choose to test in this model system because we can easily image the cells and efficiently quantify the green fluorescence for RNAi activity. The expression of GFP is observed under both conditions: (a) non-UV treated and (b) after UV treated. The HEK293T cells were co-transfected with EGFP plasmid,

lipofectamine-3000, corresponding control/*caged*-siRNAs. The cells were incubated in the transfection media for 7 h. For non-UV treatment, cells were kept in dark. For UV treatment, after 7 h of post-transfection, cells were irradiated with 365 nm (8W) for 40 min at a distance of 10 cm from source, after 36 h the cells were examined for RNAi. For qualitative analysis, the cells were imaged for green fluorescence under fluorescence microscope. For imaging cells, DAPI is used for staining the cell nucleus. The number of cells expressing the GFP can be easily compared from the total number of cells having nuclear DAPI staining. GFP is observed in FITC filter-sets. While for the quantitative analysis, GFP expression in cells was quantified by flow-cytometry.

The RNAi studies with antisense strand modified U-NBnOE siRNAs (siRNA-2/3) along with unmodified control siRNA (siRNA-1) are provided in Figure 2A.9. Cells treated with only EGFP reporter gene without siRNA have expressed the GFP in both, non-UV-treated and UV-treated cells. Importantly, the GFP expression levels are not significantly affected with UV-treatment (Figure 2A.9, GFP panel). The cells treated with control siRNA-1 knocked down the expression of GFP in both non-UV-treated and UV-treated cells as expected (Figure 2A.9, control siRNA-1 panel). In case of single U-NBnOE modification, the presence of photo labile group within the antisense strands of caged siRNA-2 and siRNA-3 (14^{th} and 17^{th} position from 5'end respectively), the non-UV-treated cells exhibit GFP expression but in UV-treated cells, GFP expression is knocked down like control siRNA-1(Figure 2A.9, siRNA-2/3 panels).

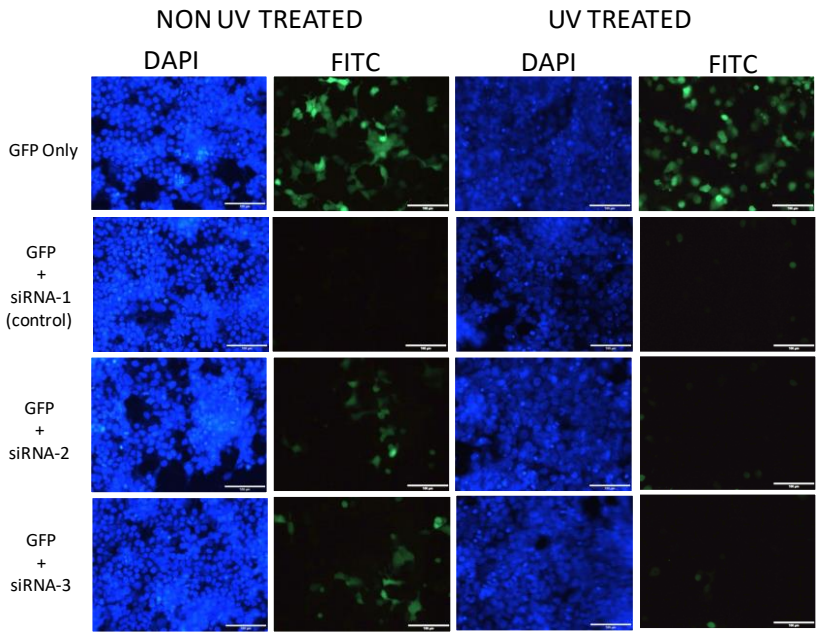


Figure 2A.9. Light regulated RNAi activity for GFP gene expression in HEK293T; control siRNA (siRNA-1) caged siRNA modified within antisense strand (siRNA-2/3) under no light treatment (non-UV treated cells) and light treatment (UV treated cells) conditions.

The RNAi studies with sense strand modified U-NBnOE siRNAs (siRNA-4/5/6/7/8) along with unmodified control siRNA (siRNA-1) are provided in Figure 2A.10. In non-UV treated cells, the U-NBnOE modifications within the sense strands of caged siRNA-4/5/6/7/8 have various levels of GFP expression, however, in their UV treated cells, the GFP expression levels are more than non-UV-treated cells, which indicate, low RNAi activity after uncaging. Though siRNA-8 have two U-NBnOE modifications (two photo labile groups), it resulted in almost similar RNAi activity like single U-NBnOE modified (one photo labile group) siRNAs. It is noteworthy that U-NBnOE modified caged siRNAs within the sense (siRNA 2-3) and antisense (siRNA 4-8) strands also have RNAi activity when compared with cells treated with only GFP. For example, caged siRNA-6 and siRNA-7 significantly knockdown the GFP expression levels (more RNAi activity) in

presence of photo labile group and less RNAi activity after uncaging, but still the caging and uncaging is effective in controlling the GFP gene expression levels by these siRNAs

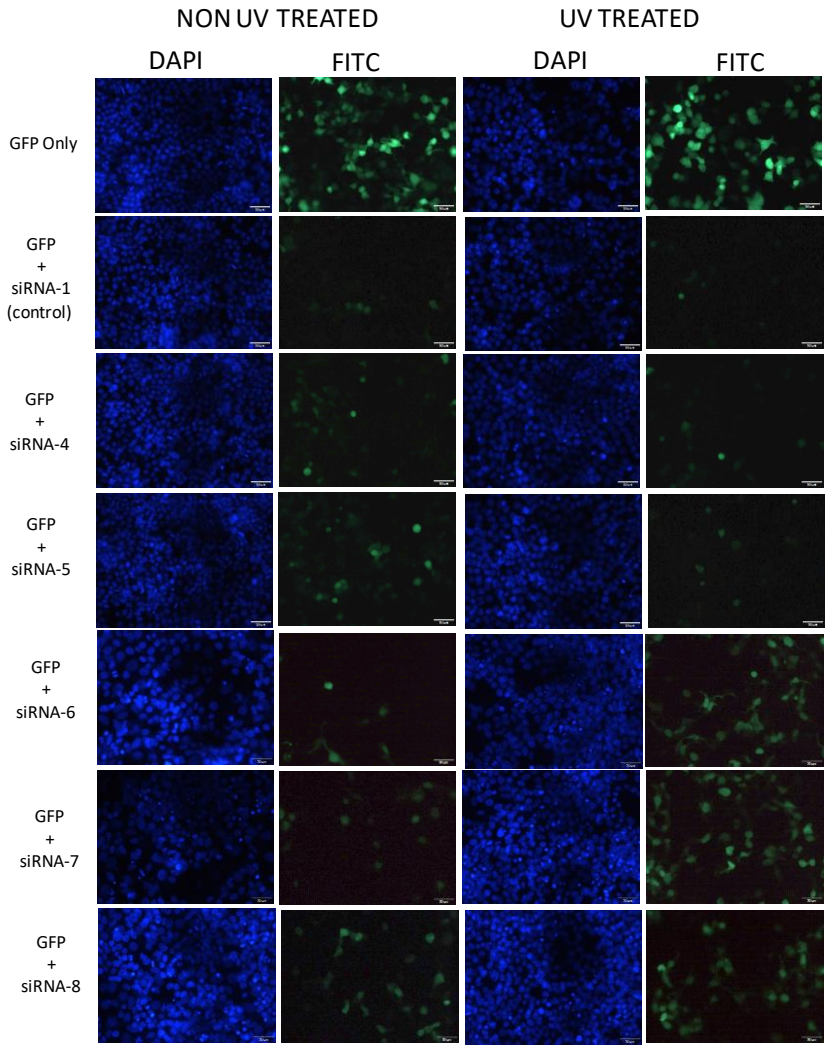


Figure 2A.10. Light regulated RNAi activity for GFP gene expression in HEK293T; control siRNA (siRNA-1) caged siRNA modified within sense strand (siRNA-2/3) under no light treatment (non-UV treated cells) and light treatment (UV treated cells) conditions.

FACS studies. Fluorescence activated cell sorting (FACS) is a sensitive and valuable technique to characterize and quantify fluorescence in cellular subpopulations and this is widely applied in GFP expression levels in determining RNAi activity. Herein, from the flow-cytometry results, we

calculated the GFP knockdown percentage with non-UV treated control siRNA-1 as our reference and plotted in normalized percentage (average value of triplicates). These results are provided in Figure 2A.11. The control siRNA-1 knockdown the GFP expression levels up to ~95% in both UV-treated and non-UV-treated cells.

The U-NBnOE modified antisense strands, siRNA-2 and siRNA-3 (14^{th} and 17^{th} position from 5'end respectively) have low RNAi activity (up to 55% and 38%) in non-UV treated cells and in their UV treated cells, have relatively high RNAi activity (up to 80% and 66%). This indicates, the caged siRNAs (siRNA-2 and siRNA-3) hindered the RNAi activity to an extent due to the presence of bulky photo labile group (nitrobenzyl group) and resulted in lower RNAi activity. When this bulky photo labile group (nitrobenzyl group) is removed (UV treated cells) RNAi activity is increased.

The sense strand modified siRNA-4 (10^{th} position from 5' end) exhibit high levels of GFP knockdown in both non-UV and UV treated cells (RNAi activity up to 80%). The modification at this position is ineffective for the presence (caged) and absence (uncaged) of photo labile group. The siRNA-5 (sense strand modified at 12^{th} position from 5' end) displayed low levels of GFP knockdown (up to 40%) in non-UV treated cells and improved activity (up to 55%) in UV treated cells. Drastic changes in GFP gene control by caging and uncaging has been resulted from modifications in sense strand of siRNAs, siRNA-6/7/8 (7^{th}, 10^{th} and 7^{th} /10^{th} positions from 5'end respectively). At these modification sites, the siRNAs displayed high levels of GFP knockdown (RNAi activity up to 85-90%) in non-UV treated cells like control siRNA-1, and UV treated cells have reduced activity (up to 40-50%). The caging and uncaging within the sense strands of siRNA-1 at these positions (7^{th}, 10^{th} and 7^{th} /10^{th} positions from 5'end respectively) effectively resulted in

controlling the GFP expression. Thus by modifying at these positions with photo labile group resulted in light controllable GFP gene silencing effectively like control siRNA.

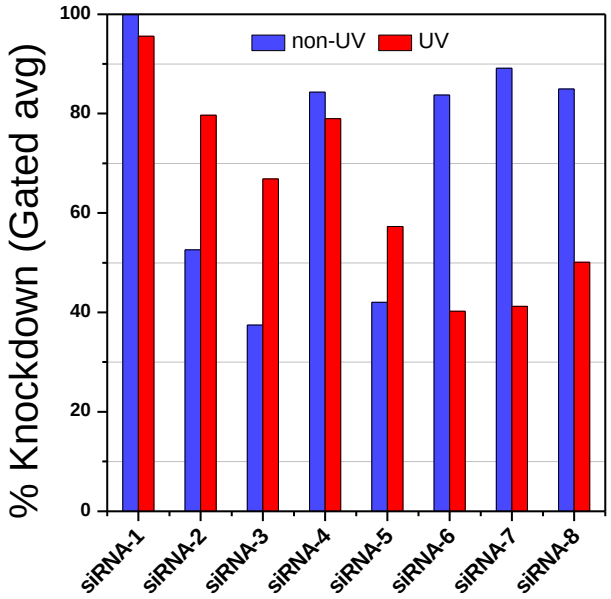


Figure 2A.11. FACS results of GFP expression in HEK293T cells to determine the RNA*i* activity. siRNA-1 (unmodified control); siRNA 2/3 (antisense strand caged); siRNA 4-8 (sense strand caged).

Serum stability. The chemical modifications to siRNA ought to enhance the siRNA duplex stability and functional activity against *endo/exo* nucleases present in the serum. The phosphorothioate backbones, 2′-OH modifications such as, 2′-O-MOE, 2′-O-Me, 2′-fluoro and LNAs have enhanced the serum stability.[11] Importantly, the improved serum stability of siRNAs alone will not govern the enhancement of RNAi activity. For *in-vitro* studies, in maintaining the cells in culture, 10% fetal bovine serum (FBS) is a key component and added as a growth supplement which contains nucleases. The stability of siRNA in this serum will distress the RNA*i* activity before reaching the

target or even during transfection. To examine the stability and tolerance of photo caged modification in caged siRNA against *endo/exo* nucleases in this serum, we performed the serum stability assay for unmodified (siRNA-1) and photo caged siRNAs (siRNA-2/3) with 10% FBS mimicking the transfection conditions (Figure 2A.12). Under these conditions these siRNAs are stable up to 24 h like unmodified control siRNA-1.

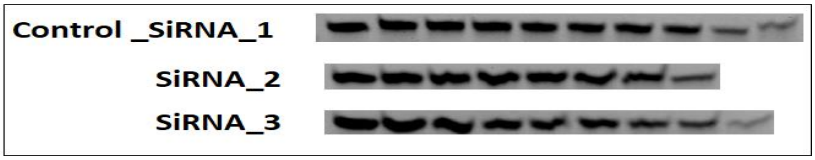


Figure 2A.12. Serum stability of control and photo caged siRNA-2/3, incubated at 37 ^{o}C with 5% CO_2 in 10% fetal bovine serum (FBS) and examined at different time intervals up to 24 h. 16% PAGE; for imaging gel was stained with ethidium bromide.

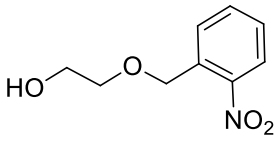
2A.4 Conclusion

We successfully synthesized and characterized rationally designed caged uridine nucleoside, U-NBnOE nucleoside and optimized the uncaging conditions at monomer level. After successful synthesis of U-NBnOE phosphoramidite we incorporated these phosphoramidites into sense and antisense strands of GFP siRNAs and confirmed the photo cleavage (*uncaging*) by ESI-MS. *Caged* siRNA's (sense and anti-sense strands) with photo labile protecting (PLP) group and its role in the light dependent controllable RNAi activity at cellular level for versatile GFP gene expression in HEK293T cell line has been tested. The stability of 2'- photo labile modified siRNA's is also tested in fetal bovine serum and found stable up to 24 h like control siRNA.

2A.5 Experimental section

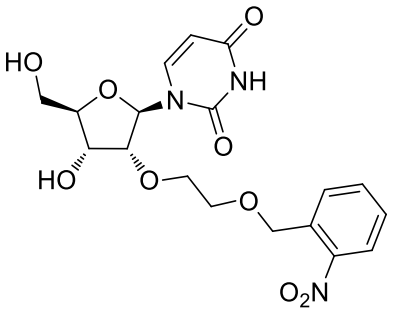
Materials. All the required chemicals were obtained from commercial suppliers and used without any further purification. Pyridine and methylene chloride solvents were distilled over calcium hydride and stored under 4Å molecular sieves prior using for reactions. Tetrahydrofuran was freshly distilled over sodium and benzophenone for reactions. All the reactions were monitored by thin layer chromatography, visualized by UV and Ninhydrin. Column chromatography was performed with 230-400 mesh silica. MS and HRMS were obtained from Bruker micrOTOF-Q II Spectrometer. $^{1}H/^{13}C/^{31}P$ NMR spectra were recorded on Bruker AV-400 or 700 MHz at 298 K. ^{1}H and ^{13}C NMR chemical shifts were recorded in ppm downfield from tetramethyl silane or residual solvent peak. Splitting patterns are abbreviated as: s, singlet; d, doublet; dd, doublet of doublet; t, triplet; q, quartet; dq, doublet of quartet; m, multiplet.

2-((2-nitrobenzyl)oxy)ethanol (**3**). Experimental procedure is followed from literature report. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.04 (dd, *J* = 8.2, 0.8 Hz, 1H), 7.78 (d, *J* = 7.8 Hz, 1H), 7.65 (dd, *J* = 11.0, 4.2 Hz, 1H), 7.44 (t, *J* = 7.7 Hz, 1H), 4.93 (s, 2H), 3.92 – 3.76 (m, 2H), 3.76 – 3.62 (m, 2H), 2.48 (s, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 147.36, 134.61, 133.62, 128.76, 128.14, 124.67, 72.40, 69.72, 61.75. ESI-HRMS $[M+H]^+$ calcd for $C_9H_{12}N_1O_4$ 198.0761, found 198.0821.

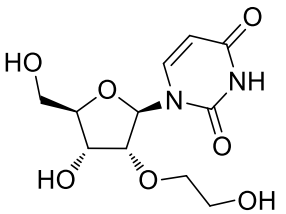


2'-*O*-(2-nitrobenzyloxy)ethyl) uridine (**4**). Compound **2** (2.30 g, 10.1 mmol) was co-evaporated with anhydrous pyridine for 2 times and dissolved in dimethyl acetamide (16 ml) and added $BF_3 \cdot OEt_2$ (3.23 ml, 0.3 mmol) dropwise. The solution was heated to 110 ^{0}C to get a clear solution then added *o*-nitrobenzyloxy ethyl alcohol, (**3**) (11.70 g, 61.06 mmol) drop wise and refluxed at 150 ^{0}C for 12 h. The crude reaction mixture concentrated in *vacuo*, the resultant oily residue was

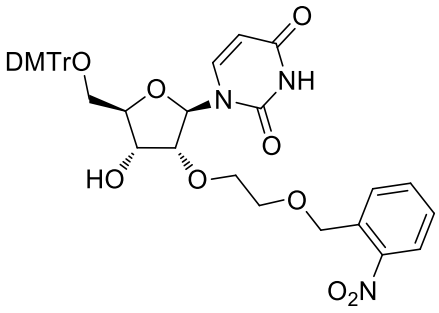
coevaporated 3 times with methanol. The resulting oil was purified by column chromatography from 2 - 6% MeOH in CH_2Cl_2 to yield 1.12 g (26%) of nucleoside **4** as white foam. ^{1}H NMR (700 MHz, DMSO) δ 11.34 (s, 1H), 8.05 (d, *J* = 8.0 Hz, 1H), 7.93 (d, *J* = 8.1 Hz, 1H), 7.76 (s, 2H), 7.56 (s, 1H), 5.89 (d, *J* = 4.7 Hz, 1H), 5.61 (d, *J* = 8.0 Hz, 1H), 5.15 (t, *J* = 4.7 Hz, 1H), 5.09 (d, *J* = 5.6 Hz, 1H), 4.84 (s, 2H), 4.13 (dd, *J* = 4.9 Hz, 1H), 4.01 (t, *J* = 4.7 Hz, 1H), 3.88 (s, 1H), 3.79 (dd, *J* = 11.0 Hz, 1H), 3.73 (dd, *J* = 4.6 Hz, 1H), 3.65 (s, 3H), 3.57 (d, *J* = 11.5 Hz, 1H). ^{13}C NMR (176 MHz, DMSO) δ 163.08, 150.58, 147.18, 140.44, 134.36, 133.87, 128.69, 128.45, 124.47, 101.78, 86.21, 85.03, 81.45, 69.74, 69.04, 68.68, 68.35, 60.47. ESI-HRMS $[M+Na]^+$ calcd for $C_{18}H_{21}N_3O_9Na$ 446.117, found 446.1190.

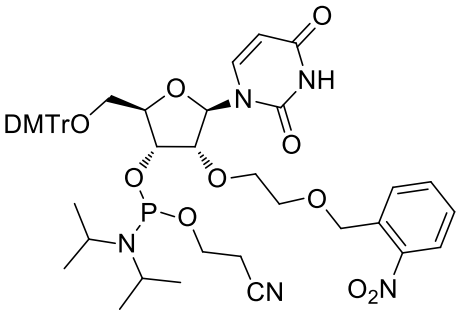


2'-*O*-((hydroxy)ethyl) uridine (**4a**). Compound **4** (15 mg, 0.0034 mmol) was dissolved in acetonitrile-d3 (0.5 ml) and irradiated at 365 nm. The completion of the reaction was monitored by proton NMR and once the reaction completed the crude reaction mixture was concentrated in *vacuo*. The resultant oily residue purified was by column chromatography from 5 - 15% MeOH in CH_2Cl_2. ^{1}H NMR (700 MHz, DMSO) δ 11.33 (s, 1H), 7.95 (d, *J* = 8.1 Hz, 1H), 5.84 (d, *J* = 5.0 Hz, 1H), 5.64 (d, *J* = 8.1 Hz, 1H), 5.23 (s, 1H), 5.10 (d, *J* = 5.0 Hz, 1H), 4.79 (s, 1H), 4.12 (dd, *J* = 9.7, 4.9 Hz, 1H), 3.95 (t, *J* = 4.9 Hz, 1H), 3.85 (dd, *J* = 7.4, 3.1 Hz, 1H), 3.61 (m, 2H), 3.59 – 3.52 (m, 2H), 3.52 – 3.47 (m, 2H), 3.16 (d, *J* = 3.7 Hz, 1H). ^{13}C NMR (176 MHz, DMSO) δ 163.15, 150.58, 140.53, 101.81, 86.19, 84.96, 81.44, 71.57, 68.51, 60.45, 60.15. ESI-HRMS $[M+Na]^+$ calcd for $C_{11}H_{16}N_2O_7Na$ 311.0852, found 311.0874.



5'-*O*-(4,4-Dimethoxytrityl)-2'-*O*-(2-nitrobenzyloxy)ethyl) uridine (**5**). Compound **4** (1.12 g, 2.6 mmol) was co-evaporated with anhydrous pyridine for 3 times and dissolved in dry CH_2Cl_2 (20 ml). To the resultant solution anhydrous $N(Et)_3$ (1.87 ml, 13.27 mmol) followed by DMTr-Cl (1.35 g, 3.98 mmol) was added and stirred at room temperature for 16 h. The crude reaction mixture evaporated to dryness and then subjected to column chromatography with prewashed silica with $N(Et)_3$ (0.5% in 65% EtOAc/n-Hexanes) to yield 1.13 g (59%) of compound **5** as white foam. ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.09 (s, 1H), 8.09 – 7.99 (m, 2H), 7.72 (d, *J* = 7.3 Hz, 1H), 7.63 (t, *J* = 7.6 Hz, 1H), 7.44 (t, *J* = 7.8 Hz, 1H), 7.41 – 7.37 (m, 2H), 7.33 – 7.22 (m, 8H), 6.87 – 6.81 (m, 4H), 5.97 (d, *J* = 1.5 Hz, 1H), 5.28 (d, *J* = 8.1 Hz, 1H), 4.93 (s, 2H), 4.48 (dd, *J* = 13.2, 7.9 Hz, 1H), 4.18 (ddd, *J* = 11.2, 5.1, 2.5 Hz, 1H), 4.07 (d, *J* = 7.7 Hz, 1H), 4.00 (dd, *J* = 5.1, 1.5 Hz, 1H), 3.94 – 3.86 (m, 1H), 3.83 – 3.74 (m, 8H), 3.55 (qd, *J* = 11.2, 2.3 Hz, 2H), 3.07 (d, *J* = 8.8 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 163.34, 158.86, 158.82, 150.30, 147.62, 144.52, 140.16, 135.46, 135.24, 134.41, 133.80, 130.35, 130.25, 128.85, 128.41, 128.30, 128.14, 127.28, 124.88, 113.43, 113.42, 102.21, 87.88, 87.19, 83.44, 83.15, 70.32, 70.04, 68.69, 61.43, 55.39, 29.82. ESI-HRMS $[M+Na]^+$ calcd for $C_{39}H_{39}N_3O_{11}Na$ 748.2477, found 748.2464.





5'-*O*-(4,4-Dimethoxytrityl)-2'-*O*-(2-nitrobenzyloxy)ethyl)uridine-3'-O-(2-Cyanoethyl diisopropylphosphoramidite) (**6**). To a stirred solution of compound **5** (0.2 g, 0.27 mmol) in anhydrous THF (1.4 ml) and N,N-Diisopropyl ethylamine (0.287 ml, 1.6 mmol) was treated with 2-cyanoethyl diisopropylchlorophosphoramidite (0.122 ml, 0.5 mmol) at room temperature. After

stirred for 2 h, the crude reaction mixture subjected to column chromatography using silica gel. Silica-gel was prewashed with 1% $N(Et)_3$ in 60% EtOAc in n-Hexanes and eluted with 60% EtOAc in n-Hexanes to yield 0.193 g (76%) as white foam. ^{31}P NMR (162 MHz, $CDCl_3$) δ 150.69, 150.52. ESI-HRMS $[M+H]^+$ calcd for $C_{48}H_{57}N_5O_{12}P$ 926.3736, found 926.3756.

Solid-phase RNA Oligonucleotide synthesis. All the reagents were purchased from Sigma-Aldrich and standard 2'-TOM/TBDMS-phosphoramidites from ChemGenes Inc... All the RNA oligonucleotides were synthesized using ABI 394 DNA/RNA Synthesizer on multiple 0.2 µmol with a double coupling and increased coupling time of 10 min for modified amidites with DMT off. The standard amidites were dissolved in dry CH_3CN for 0.10 M and for modified amidite was 0.13 M concentration. After the synthesis, solid support (polystyrene resin beads) was transferred in to screw cap vial for cleavage and deprotection using ammonium hydroxide and 40% aqueous methyl amine (1:1) at 60 ^{0}C for 20 min. After cooling at 0 ^{0}C centrifuged and supernatant was transferred in to 2.0 ml eppendorf tube and concentrated. The residue was dissolved in anhy. DMSO (100 ul) and triethylamine trifluoride (150 ul) and heated at 65 ^{0}C for 2.5 h. The crude siRNA's was cooled to room temperature and precipitated. The precipitated oligonucleotides were purified by 20% denaturating PAGE using 1xTBE (pH 8.3). The purified RNA was desalted using NAP-G25 columns using sterile water and quantified (OD260) using UV-Vis spectroscopy, the fractions containing RNA oligonucleotides were lyophilized and used for further experiments.

Preparation of duplex siRNA. 500 pmol of each single strand RNA oligonucleotides were mixed in phosphate-buffered saline (PBS) and annealed in waterbath at 95 ^{0}C for 5 min then switched off the bath temperature and allowed to reach room temperature then kept at 4 ^{0}C for 1 h.

CD spectra. CD spectra were recorded on a Jasco-1500 spectropolarimeter (Jasco GmbH) at 20 °C using 10 mm super micro quartz cuvettes (Thorlabs). All the siRNA were diluted to either 1µM/5µM in PBS buffer and Samples were irradiated at 365 nm at room temperature for 5 min at each time points. Spectra were recorded from 200-300 nm with 2 nm data pitch and 2 nm data interval.

Serum stability. 25 pmol of stock SiRNA was mixed with DMEM media and added 10 ul of neat fetal bovine serum (FBS), finally adjusted the total volume to 100 ul with DMEM to maintain overall of 10% FBS (v/v). An aliquot of 10ul immediately taken out and quenched with 10ul of loading buffer (0.2% bromophenol, 10xTBE, sterile water, pH 8.28 (8:1:1)) and snap freeze on dry ice and stored at -80 ^{0}C. Remaining sample was incubated at 37 ^{0}C with 5% CO_2; remaining samples were made at given time points similarly. The samples were analyzed by 16% native PAGE with 1XTBE as running buffer (pH 8.27). The gel was fixed with 10% acetic acid in water for 10 min and stained with ethidium bromide (0.1% in 1xTBE) for 5 min and finally washed the gel with 1x TBE. The gel images were recorded using ChemDoc and density plots were made using ImageJ software.

Transfection and Cell Imaging. Human embryonic cell lines, HEK293T were cultured in DMEM (Pan-biotech) supplemented with 10% FBS (Pan-biotech), 1% PSA (penicllin, streptomycin and amphotericin). One day before transfection, cells were seeded at a density of 10,000 cells/well in a 96 well plate in 100 µl complete media. For transfection, the medium in each well was replaced with a mixture containing 65 ng p-EGFP-C1 plasmid (invitrogen) , 3.12 pmol siRNA and 0.5 µL lipofectamine-3000 (Invitrogen) in 120 µL serum-free DMEM (according to manufacture recomandations). Cells were incubated in the transfection media for 6 h and the transfection media was then replaced with 100 µL complete media. For UV treatment, after incubation for 7 h in

transfection media, cells were exposed to UV light at 365 nm from a UV lamp (8W) at a distance of 10 cm from the lamp for 40 min. The cells were then incubated for 36 h at 37 ^{0}C with 5% CO_2 in a CO_2 incubator (Galaxy 170R, New Brunswick) Upon incubation, cells were fixed with 4% PFA (Paraformaldehyde) for 10 mins followed by DAPI staining. GFP expression was checked using, 485 nm excitation filter and a 535 nm emission filter.

Flow cytometry. Human embryonic cell lines, HEK293T were cultured as described above. One day before transfection, cells were seeded equally in 6 well-plates. On the day of transfection, when the cells reached 70% confluency the medium in each well was replaced with a mixture containing 2.5 µg p-EGFP-C1 plasmid (invitrogen), 75 pmol siRNA and 7.5 µL lipofectamine-3000 (company name) in 1 ml serum-reduced DMEM (Invitrogen). Cells were incubated in the transfection media for 8 h and the transfection media was then replaced with 3 ml complete media. UV treatment was carried out as described above. After 36 hours incubation, cells were collected from each well upon trypsinisation and finally resuspended in 100 µl PBS. GFP expression is these cells were acquired with FACS Calibur (BD Biosciences) and data were analysed using CELL QUEST PROsoftware (BD Biosciences).

2A.6 Reference

1. Fire, A.; Xu, S.; Montgomery, M. K.; Kostas, S. A.; Driver, S. E.; Mello, C. C., Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **1998,** *391*, 806-811.
2. Mello, C. C.; Conte, D., Revealing the world of RNA interference. *Nature* **2004,** *431*, 338-342.
3. MacFarlane, L.-A.; R Murphy, P., MicroRNA: biogenesis, function and role in cancer. *Curr Genomics* **2010,** *11*, 537-561.

4. Elbashir, S. M.; Harborth, J.; Lendeckel, W.; Yalcin, A.; Weber, K.; Tuschl, T., Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **2001,** *411*, 494-498.

5. Setten, R. L.; Rossi, J. J.; Han, S.-p., The current state and future directions of RNAi-based therapeutics. *Nat. Rev. Drug. Discov.* **2019,** *18*, 421-446.

6. Elbashir, S. M.; Martinez, J.; Patkaniowska, A.; Lendeckel, W.; Tuschl, T., Functional anatomy of siRNAs for mediating efficient RNAi in Drosophila melanogaster embryo lysate. *EMBO. J.* **2001,** *20*, 6877-6888.

7. Chiu, Y.-L.; Rana, T. M., RNAi in human cells: basic structural and functional features of small interfering RNA. *Mol. Cell* **2002,** *10*, 549-561.

8. Chiu, Y.-L.; Rana, T. M., siRNA function in RNAi: a chemical modification analysis. *RNA* **2003,** *9*, 1034-1048.

9. Levin, A. A., Treating disease at the RNA level with oligonucleotides. *N. Engl. J. Med.* **2019,** *380*, 57-70.

10. Prakash, T. P.; Allerson, C. R.; Dande, P.; Vickers, T. A.; Sioufi, N.; Jarres, R.; Baker, B. F.; Swayze, E. E.; Griffey, R. H.; Bhat, B., Positional effect of chemical modifications on short interference RNA activity in mammalian cells. *J. Med. Chem.* **2005,** *48*, 4247-4253.

11. Behlke, M. A., Chemical modification of siRNAs for in vivo use. *Oligonucleotides* **2008,** *18*, 305-320.

12. Adams, D.; Gonzalez-Duarte, A.; O'Riordan, W. D.; Yang, C.-C.; Ueda, M.; Kristen, A. V.; Tournev, I.; Schmidt, H. H.; Coelho, T.; Berk, J. L., Patisiran, an RNAi therapeutic, for hereditary transthyretin amyloidosis. *N. Engl. J. Med.* **2018,** *379*, 11-21.

13. van Ravesteyn, T. W.; Dekker, M.; Fish, A.; Sixma, T. K.; Wolters, A.; Dekker, R. J.; te Riele, H. P., LNA modification of single-stranded DNA oligonucleotides allows subtle gene modification in mismatch-repair-proficient cells. *Proc. Natl. Acad. Sci. USA* **2016,** *113*, 4122-4127.

14. Shah, S.; Rangarajan, S.; Friedman, S. H., Light-activated RNA interference. *Angew. Chem., Int. Ed.* **2005,** *44*, 1328-1332.

15. Nguyen, Q. N.; Chavli, R. V.; Marques, J. T.; Conrad II, P. G.; Wang, D.; He, W.; Belisle, B. E.; Zhang, A.; Pastor, L. M.; Witney, F. R., Light controllable siRNAs regulate gene suppression and phenotypes in cells. *Biochim. Biophys. Acta.* **2006,** *1758*, 394-403.

16. Mikat, V.; Heckel, A., Light-dependent RNA interference with nucleobase-caged siRNAs. *RNA* **2007,** *13*, 2341-2347.

17. Chaulk, S. G.; MacMillan, A. M., Caged RNA: photo-control of a ribozyme reaction. *Nucleic Acids Res.* **1998,** *26*, 3173-3178.

18. Chaulk, S. G.; MacMillan, A. M., Synthesis of oligo-RNAs with photocaged adenosine 2′-hydroxyls. *Nat. Protoc.* **2007,** *2*, 1052-1058.

19. Blidner, R. A.; Svoboda, K. R.; Hammer, R. P.; Monroe, W. T., Photoinduced RNA interference using DMNPE-caged 2′-deoxy-2′-fluoro substituted nucleic acids in vitro and in vivo. *Mol. Biosyst.* **2008,** *4*, 431-440.

20. Matsushita-Ishiodori, Y.; Ohtsuki, T., Photoinduced RNA interference. *Acc. Chem. Res* **2012,** *45*, 1039-1047.

21. Yu, L.; Jing, N.; Yang, Z.; Zhang, L.; Tang, X., Caged siRNAs with single folic acid modification of antisense RNA for photomodulation of exogenous and endogenous gene expression in cells. *Org. Biomol. Chem.* **2018,** *16*, 7029-7035.

22. Jin, S.; Miduturu, C. V.; McKinney, D. C.; Silverman, S. K., Synthesis of amine-and thiol-modified nucleoside phosphoramidites for site-specific introduction of biophysical probes into RNA. *J. Org. Chem.* **2005,** *70*, 4284-4299.

23. Arimitsu, K.; Kimura, H.; Kajimoto, T.; Ono, M.; Ohmomo, Y.; Yamashita, M.; Node, M.; Saji, H., Novel design and synthesis of a radioiodinated glycolipid analog as an acceptor substrate for N-acetylglucosaminyltransferase V. *J. Labelled Comp. Radiopharm.* **2013,** *56*, 562-572.

24. Cheng, X.; Gale, D. C.; Udseth, H. R.; Smith, R. D., Charge state reduction of oligonucleotide negative ions from electrospray ionization. *Anal. Chem.* **1995,** *67*, 586-593.

25. Gray, D. M.; Hung, S.-H.; Johnson, K. H., Absorption and circular dichroism spectroscopy of nucleic acid duplexes and triplexes. *Methods Enzymol.* **1995,** *246*, 19-34.

2A.7 Appendix

1.^{1}H-/ ^{13}C-/ESI-MS/HRMS spectra of compound 3

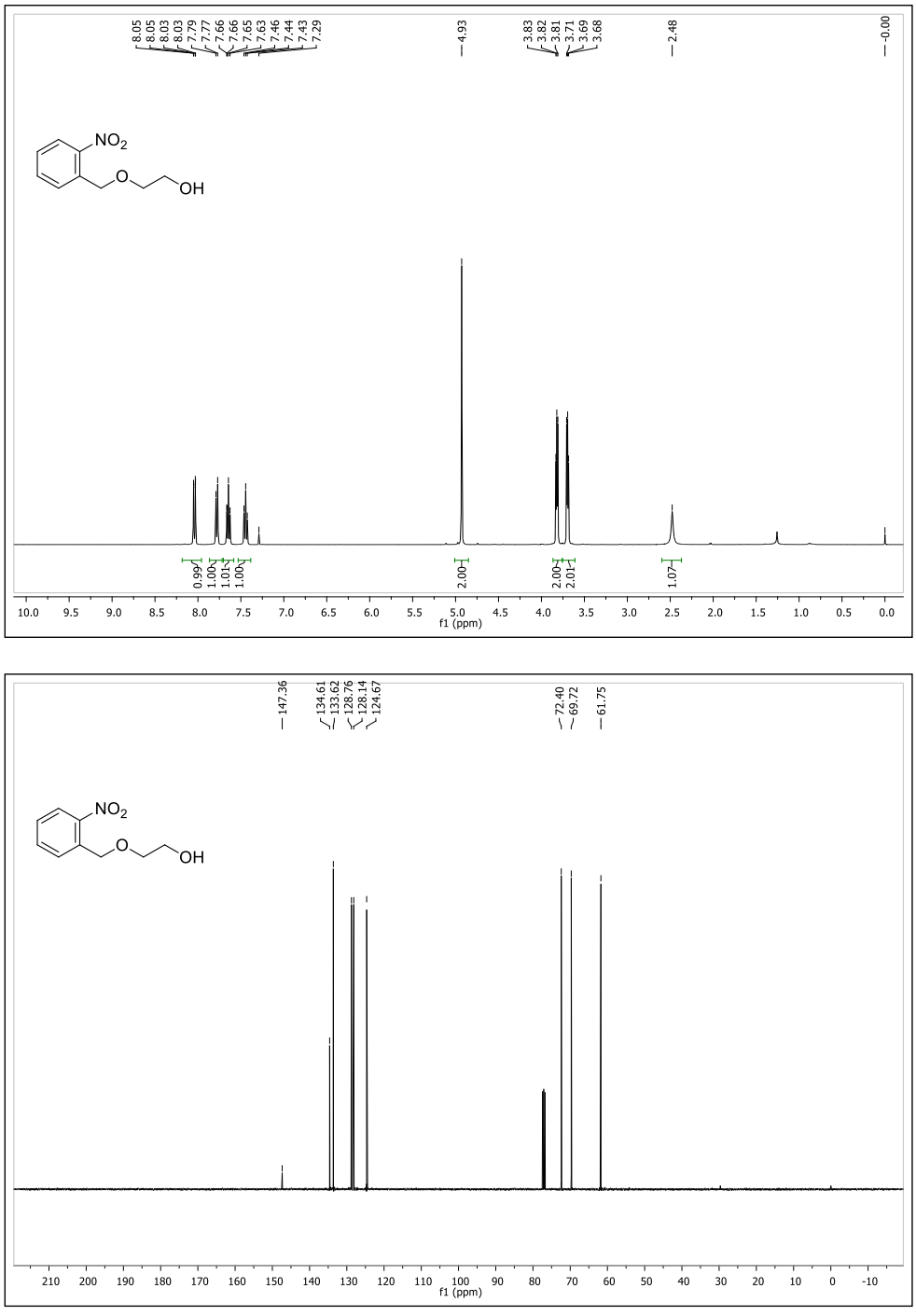


Figure S1. ^{1}H and ^{13}C NMR spectra of compound **2** in $CDCl_3$.

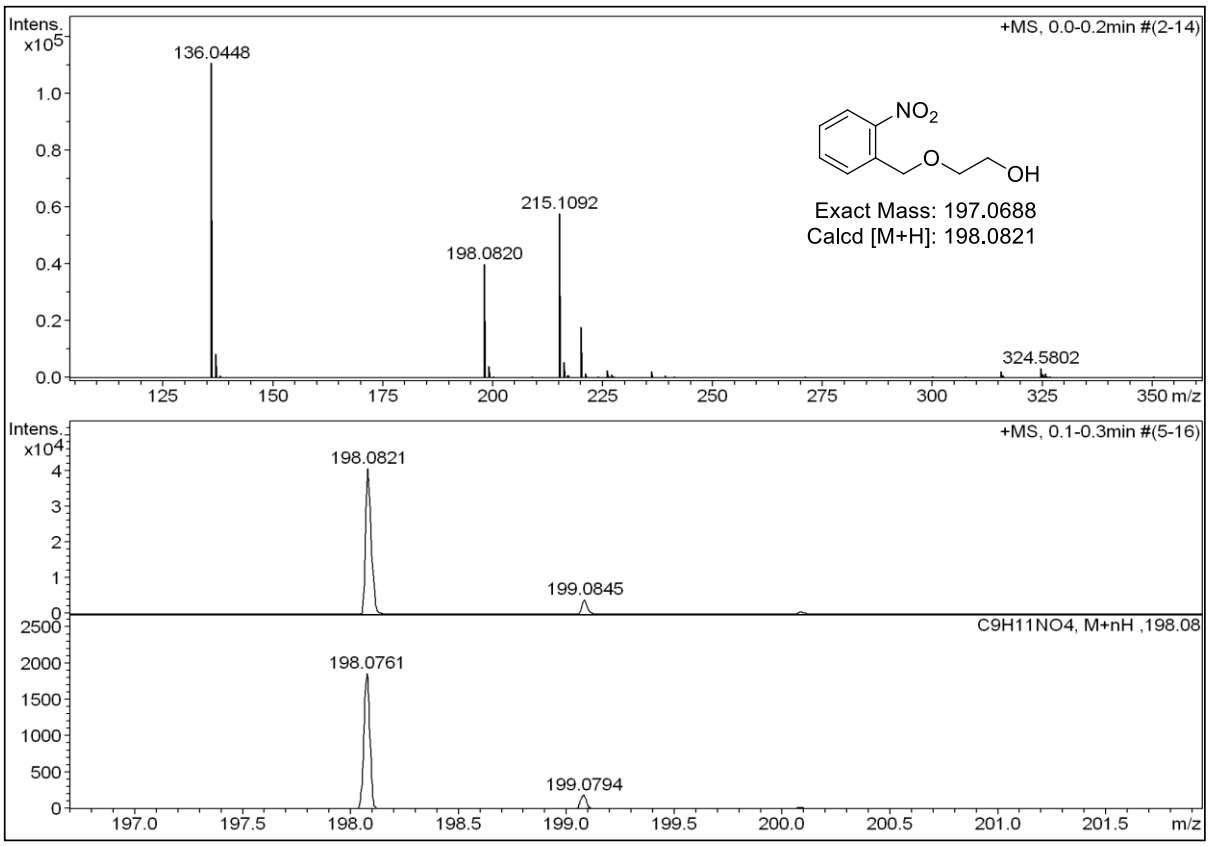


Figure S2. ESI-HRMS of compound **2**.

2. ^{1}H-/^{13}C-/ESI-MS/HRMS spectra of U-NBnOE nucleoside, 4

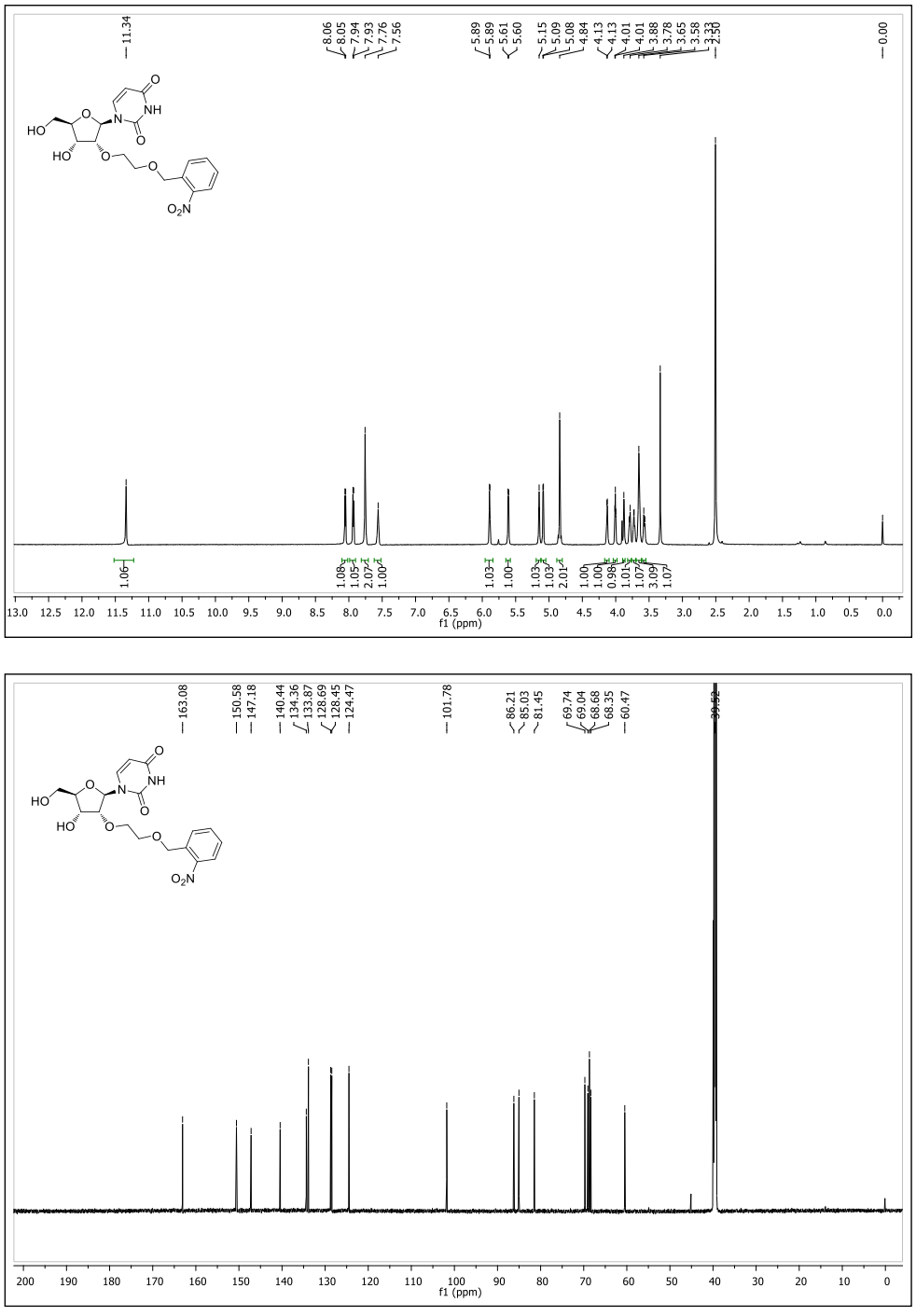


Figure S3. ^{1}H and ^{13}C NMR spectra of U-NBnOE nucleoside **4** in $CDCl_3$.

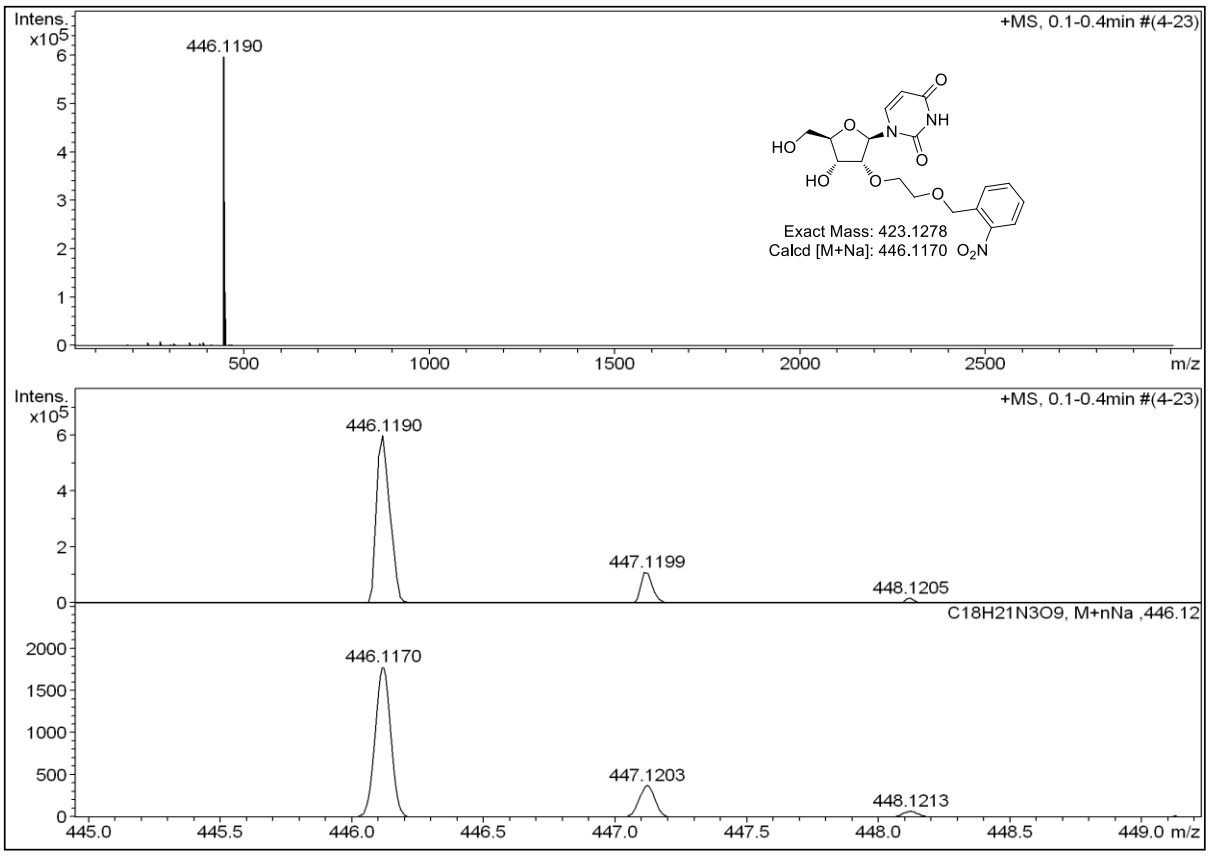


Figure S4. ESI-HRMS of U-NBnOE nucleoside **4**.

3.^{1}H-/^{13}C-ESI-MS/HRMS/D_2O exchange spectra of U-NBnOE nucleoside, **4a**

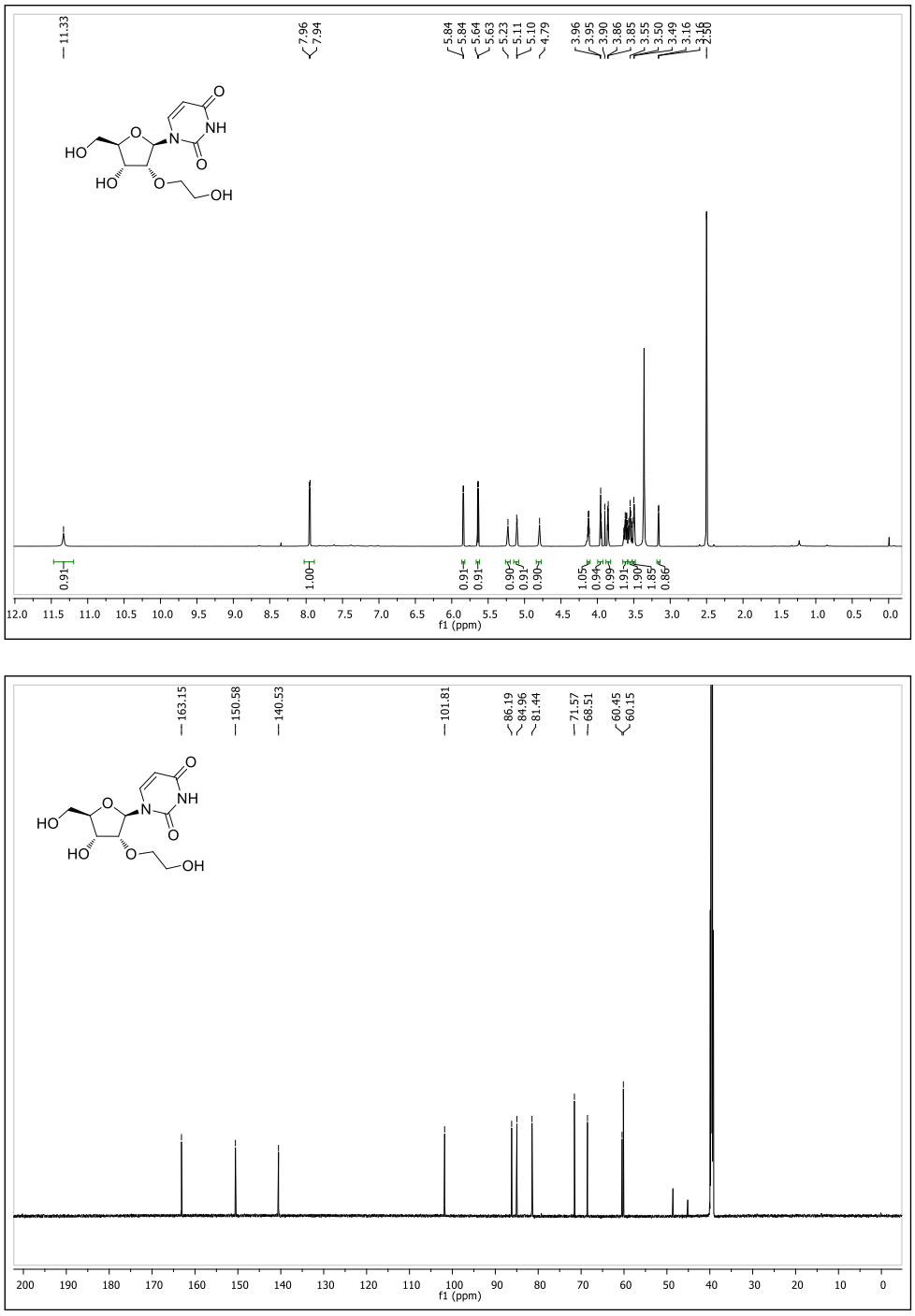


Figure S5. ^{1}H and ^{13}C NMR spectra of U-NBnOE nucleoside **4a** in DMSO-*d*$_6$.

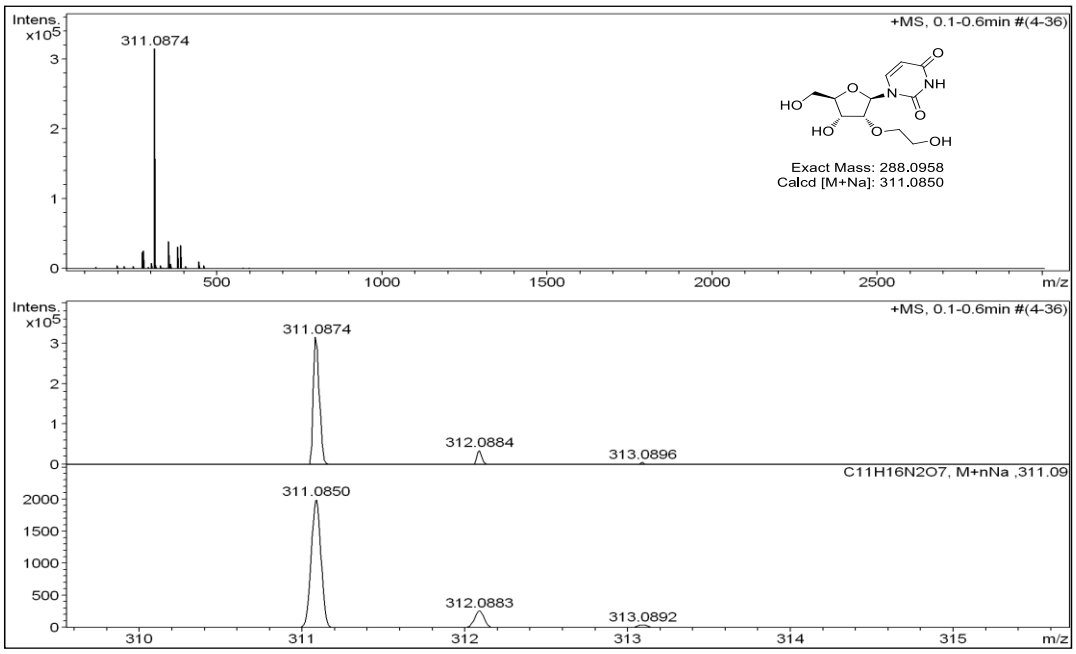


Figure S6. ESI-HRMS of U-NBnOE nucleoside **4a**.

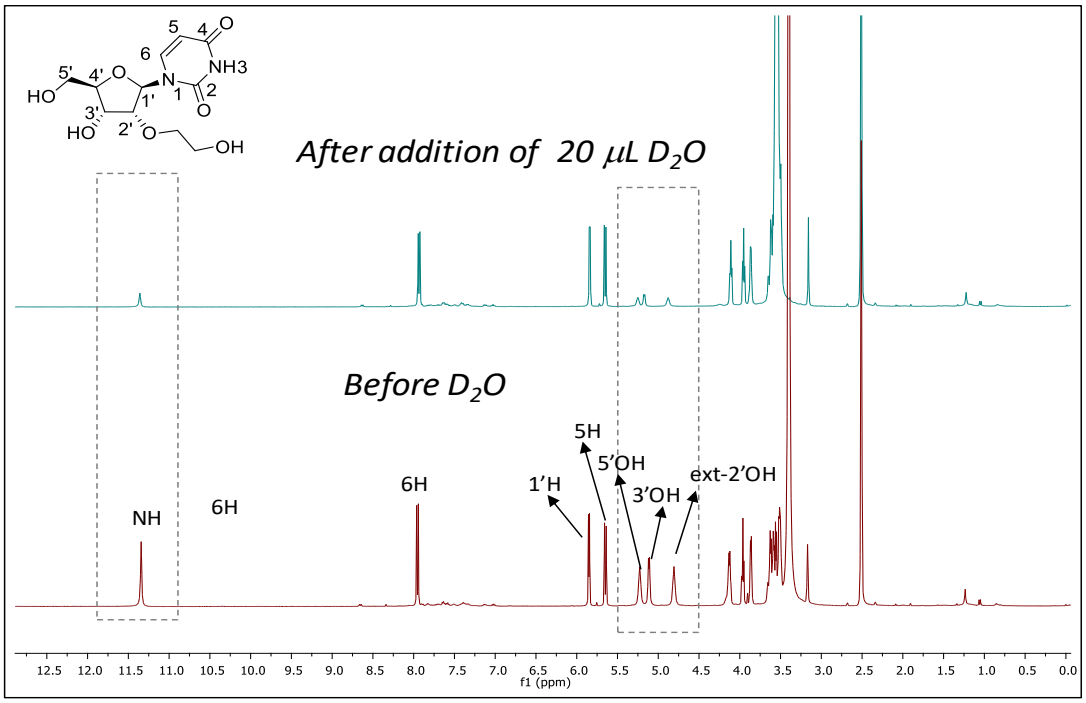


Figure S7. D_2O exchange ^1H NMR spectra of nucleoside **4a**, 2'-O extended OH, 3'-OH and 5'-OH are exchangeable. These changes are framed bfore and after D_2O shake.

4.^1H-/^{13}C-/ESI-MS/HRMS spectra of U-NBnOE nucleoside, **5**

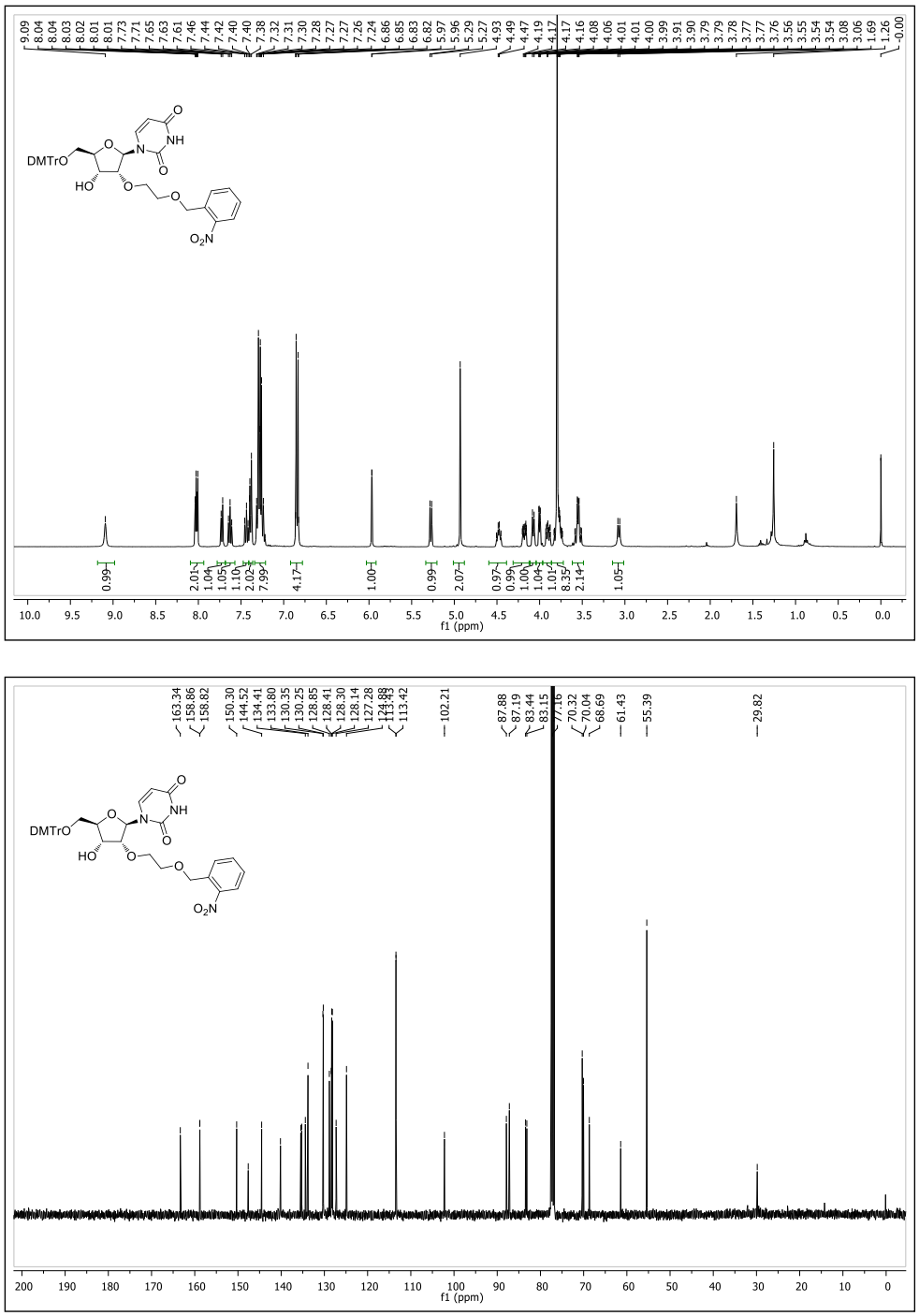


Figure S8. ^{1}H and ^{13}C NMR spectra of U-NBnOE derivative, **5** in $CDCl_3$.

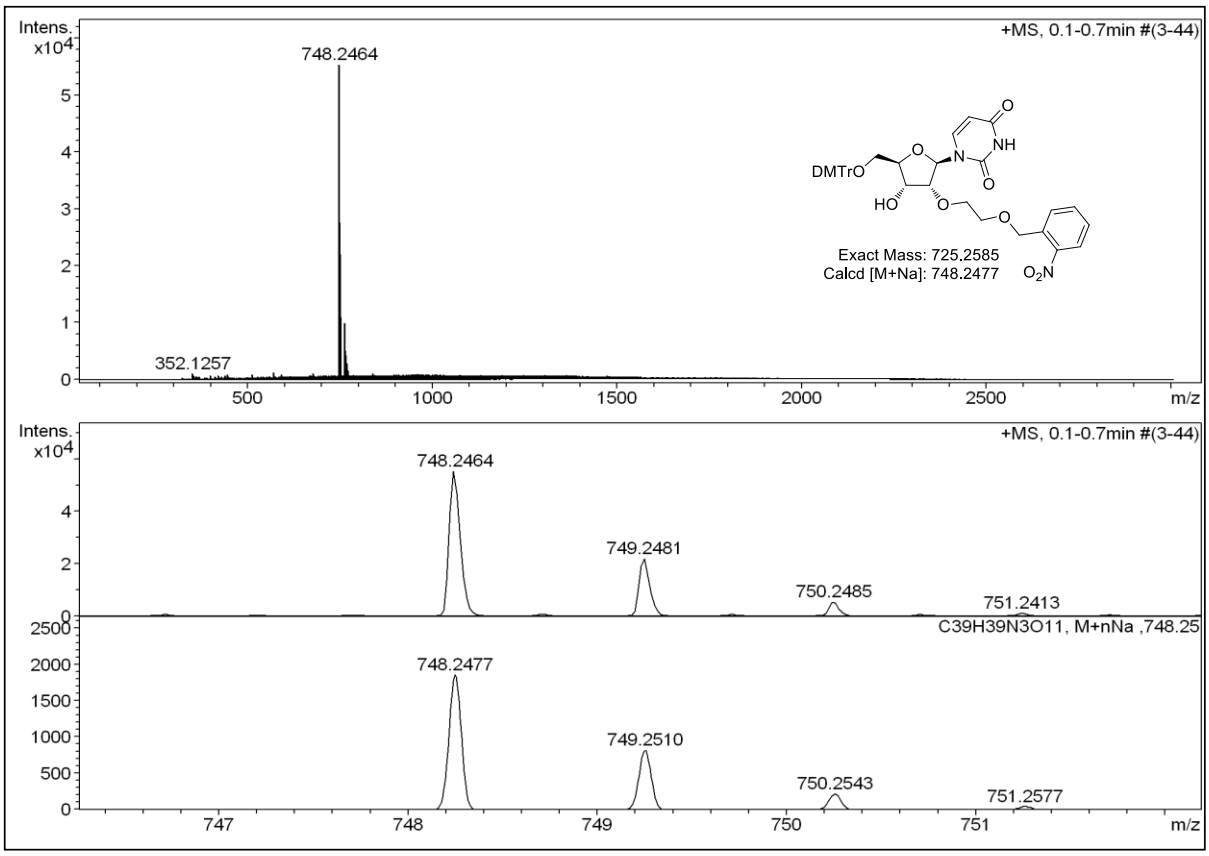


Figure S9. ESI-HRMS of U-NBnOE nucleoside **5**

5.^{1}H-/ ^{31}P-/ESI-MS/HRMS spectra of U-NBnOE phophoramidite, **6**

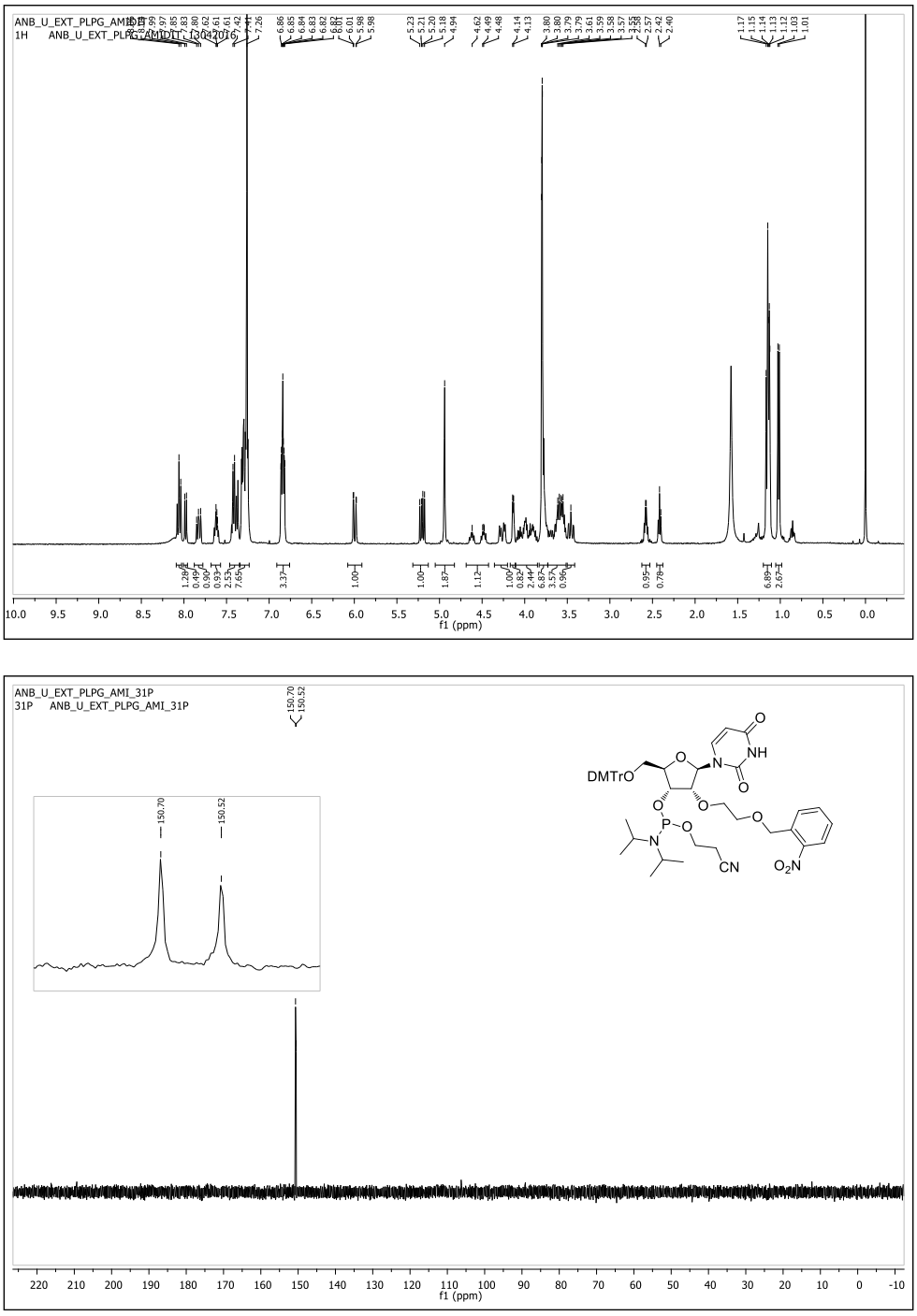


Figure S10. ^{1}H and ^{31}P NMR spectra of U-NBnOE phosphoramidite derivative **6** in $CDCl_3$

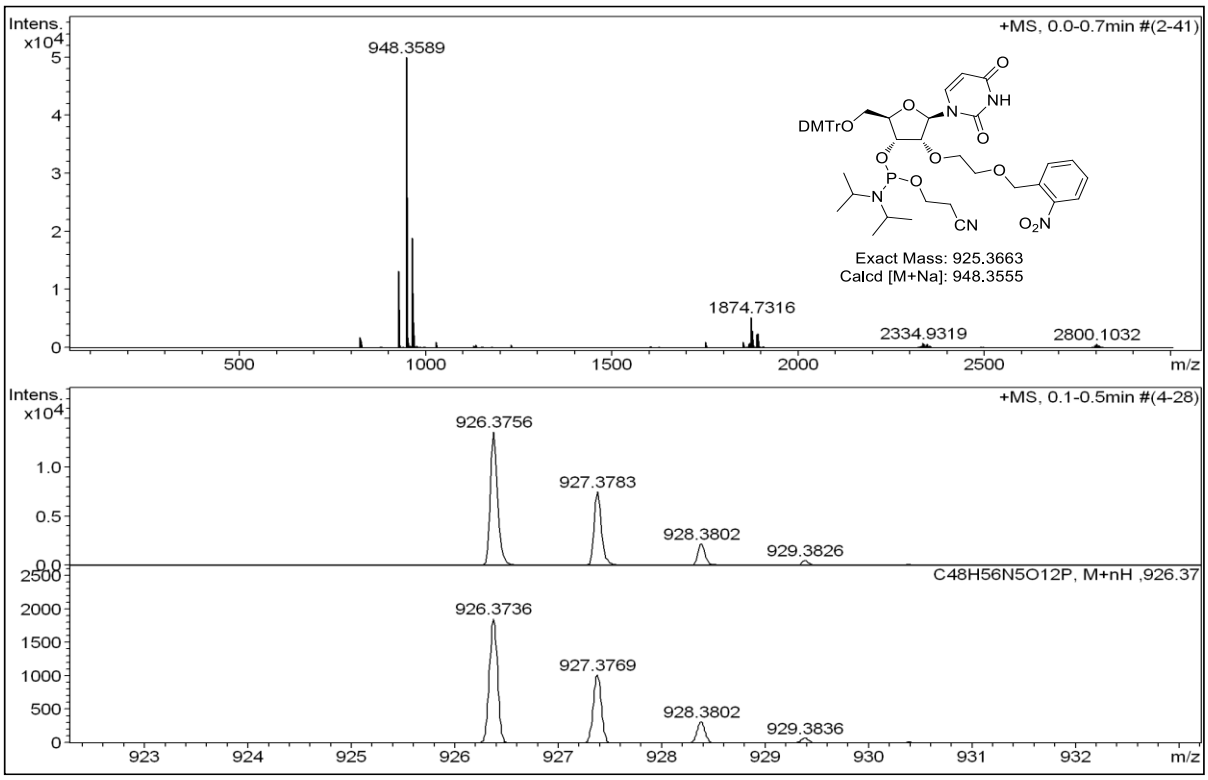


Figure S11. ESI-HRMS of U-NBnOE phosphoramidite derivative **6.**

6. ^{1}H-^{1}H COSY/NOESY NMR spectra of U-NBnOE nucleoside **4** and **4a**

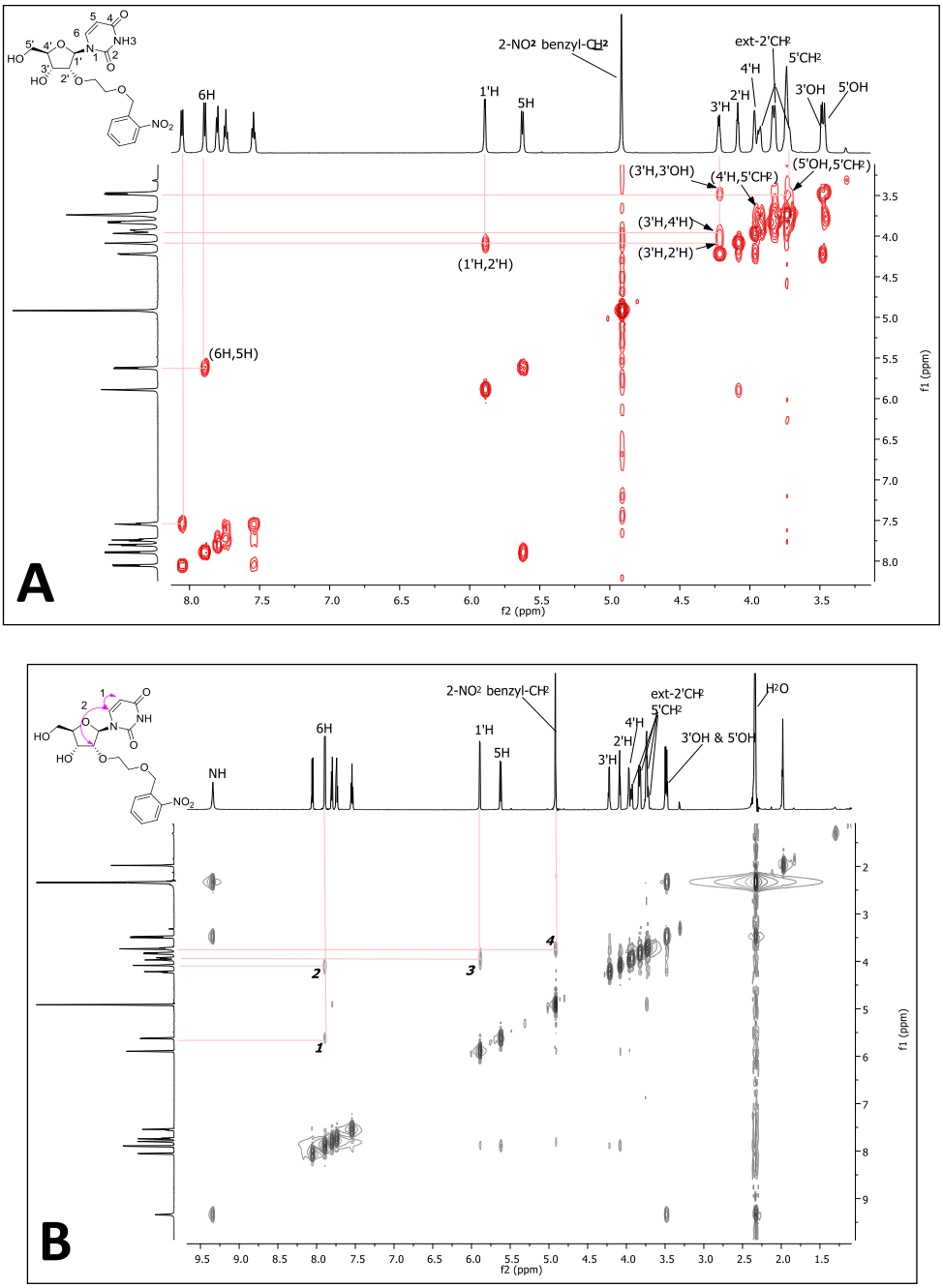


Figure S12. ^{1}H-^{1}H COSY (A) and ^{1}H-^{1}H NOESY (B) spectra of caged nucleoside **XX** in CD_3CN. (Cross peaks in ^{1}H-^{1}H NOESY assigned as 3 & 4 in spectra B are unable to assign due to very close/merged peaks).

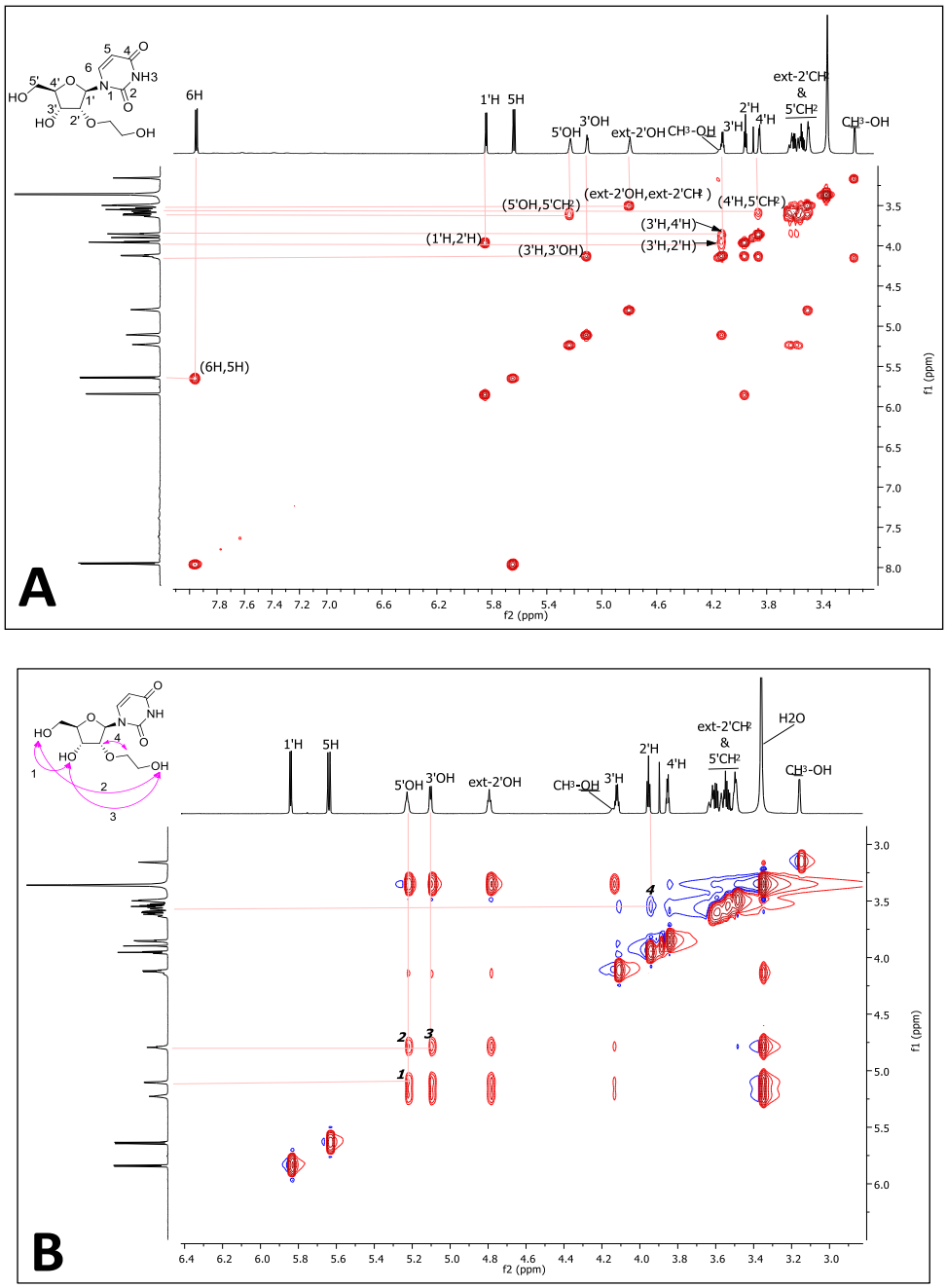


Figure 13. ^{1}H-^{1}H COSY (A) and ^{1}H-^{1}H NOESY (B) spectra of uncaged nucleoside **XX** in DMSO-*d*$_6$. Due to presence of water (δ3.4 ppm) in DMSO-*d*$_6$, we also observed some strong NOEs with water.

7.ESI-MS of RNA oligonucleotides and their photo cleavage study

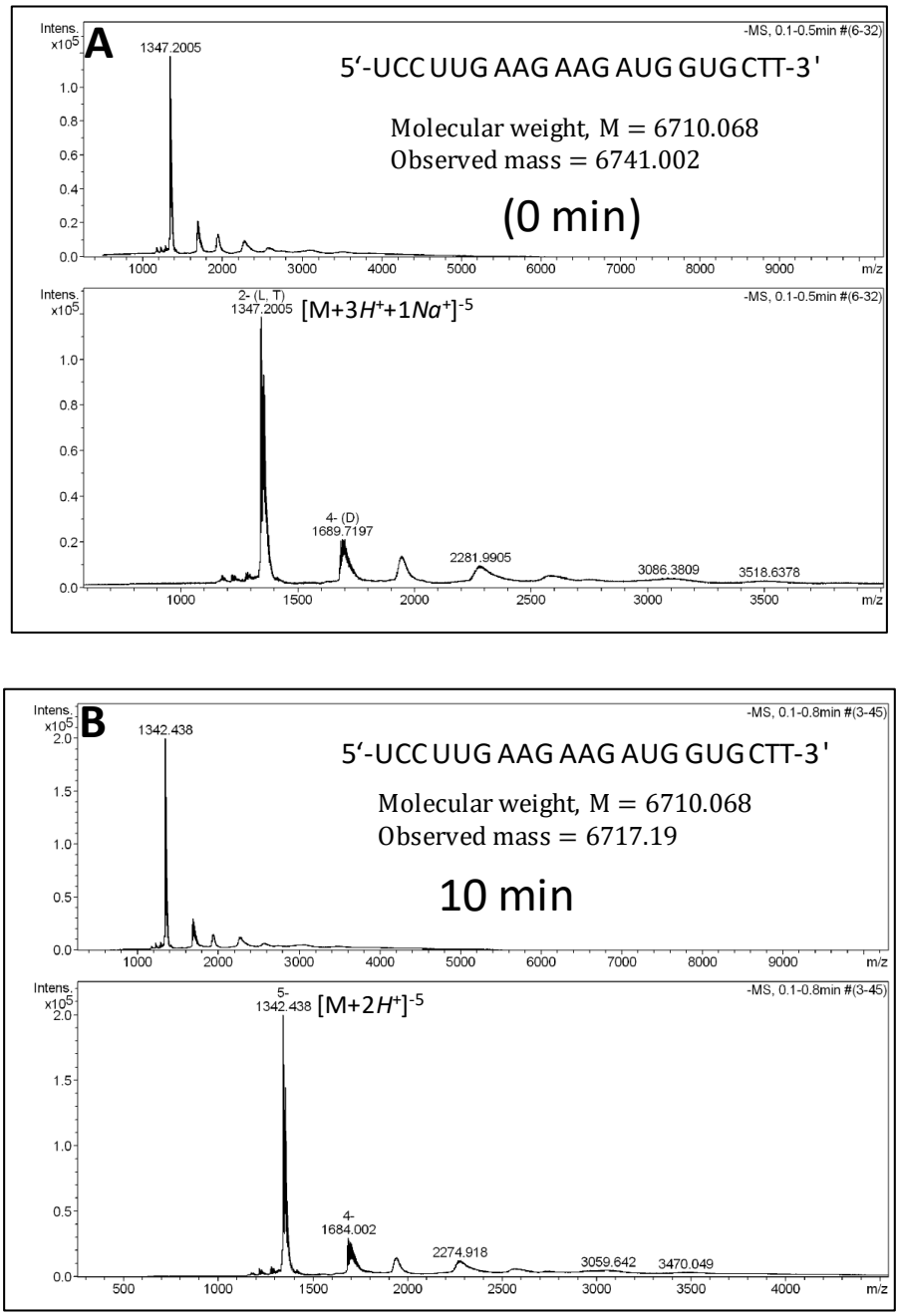


Figure S14. ESI-MS of unmodified control **RNA AS1**, before (A) and after (B) irradiating at 365 nm for 10 min. There are no significant changes in mass except loss of sodium ion after exposure to light.

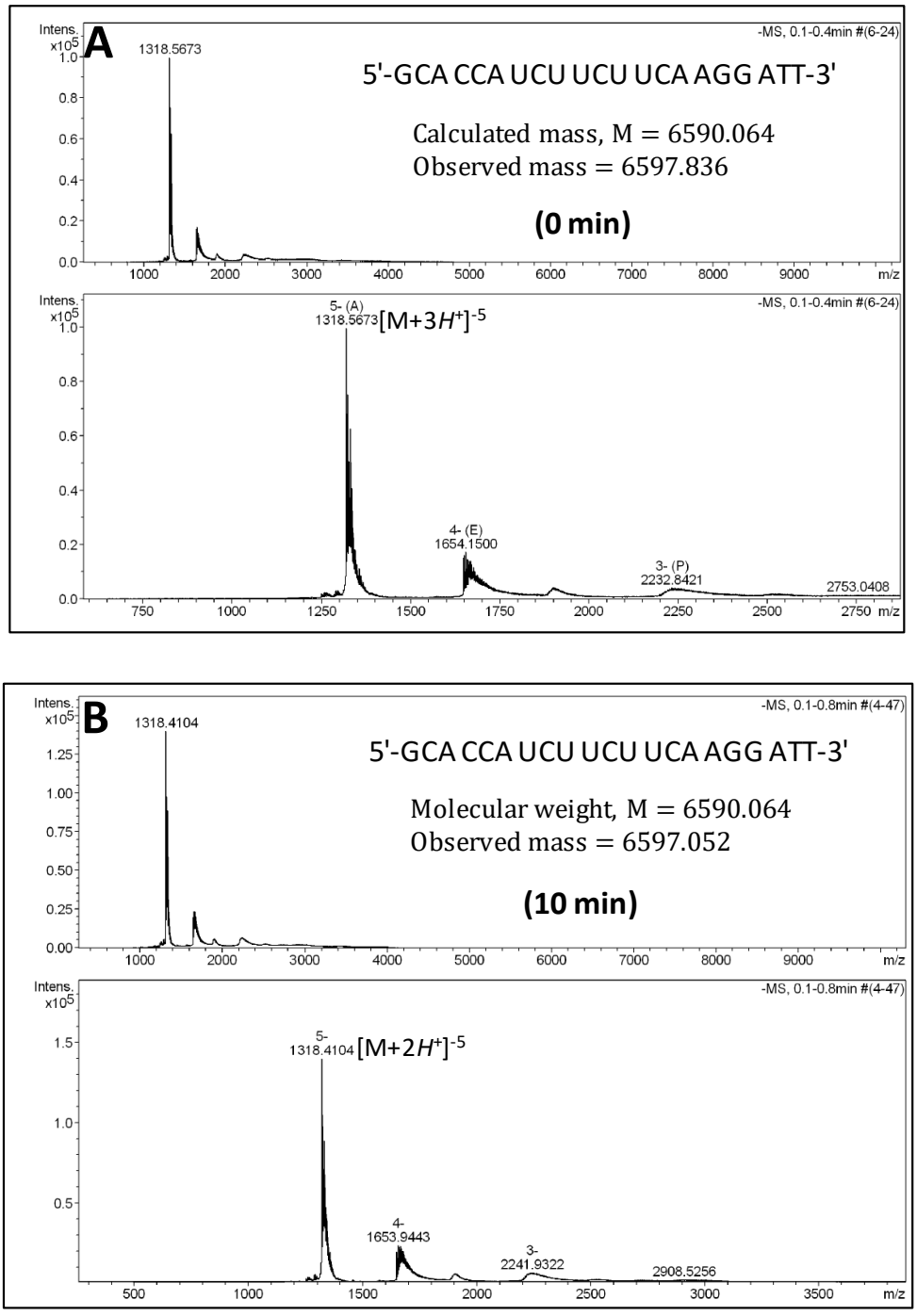


Figure S15. ESI-MS of unmodified control **RNA S1**, before (A) and after (B) irradiating at 365 nm for 10 min. There are no significant changes in mass after exposure of light.

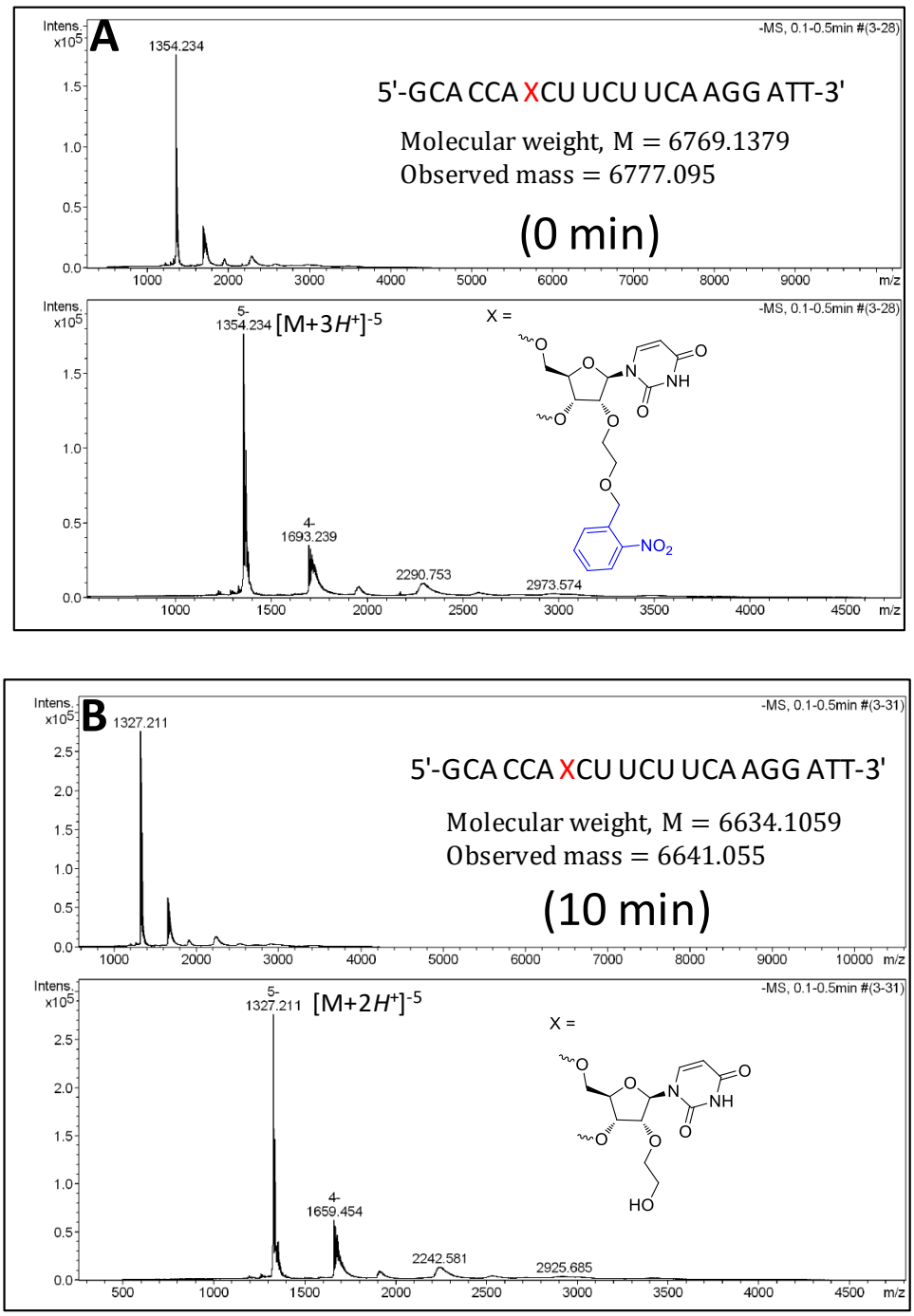


Figure S16. ESI-MS of U-NBnOE **RNA S2**, before (A) and after (B) irradiating at 365 nm for 10 min. (6777.095-6641.05 = 136.04 mass difference per one nitrobenzyl group) indicating the photo cleavage of 2-nitrobenzyl group.

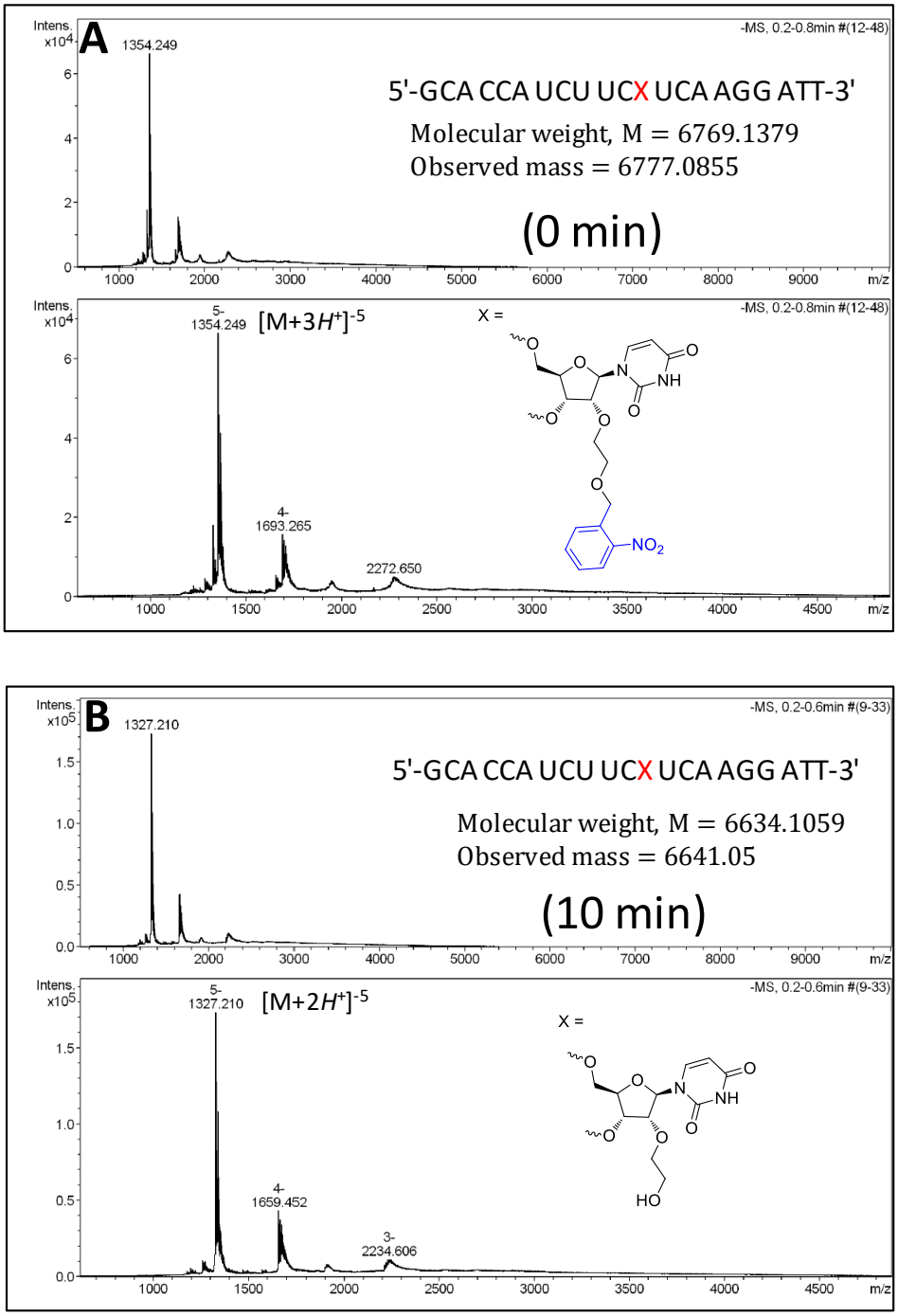


Figure S17. ESI-MS of U-NBnOE **RNA S3**, before (A) and after (B) irradiating at 365 nm for 10 min. (6777.08-6641.05 = 136.03 mass difference per one nitrobenzyl group) indicating the photo cleavage of 2-nitrobenzyl group.

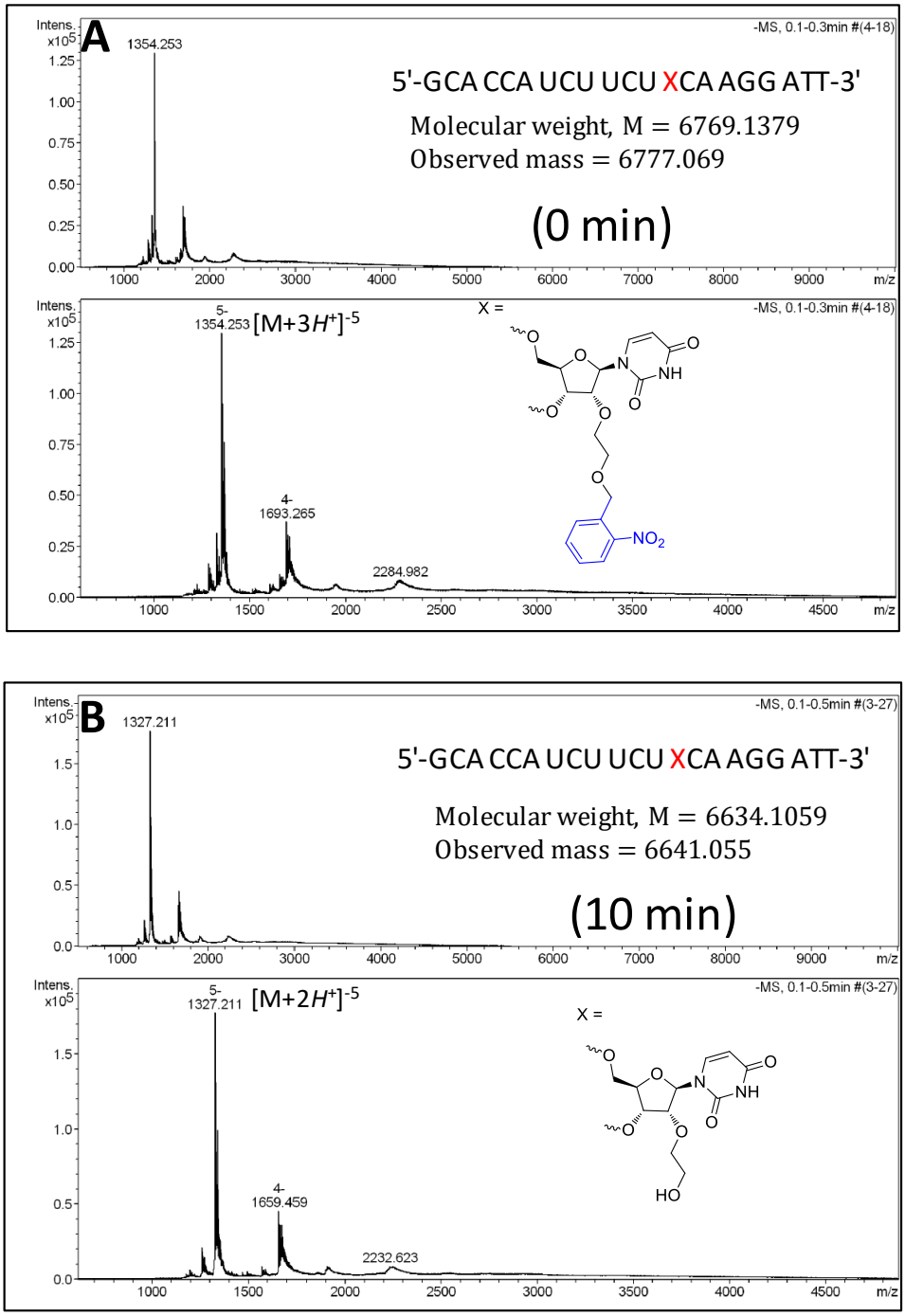


Figure S18. ESI-MS of U-NBnOE **RNA S4**, before (A) and after (B) irradiating at 365 nm for 10 min. (6777.069-6641.055 = 136.014 mass difference per one nitrobenzyl group) indicating the photo cleavage of 2-nitrobenzyl group.

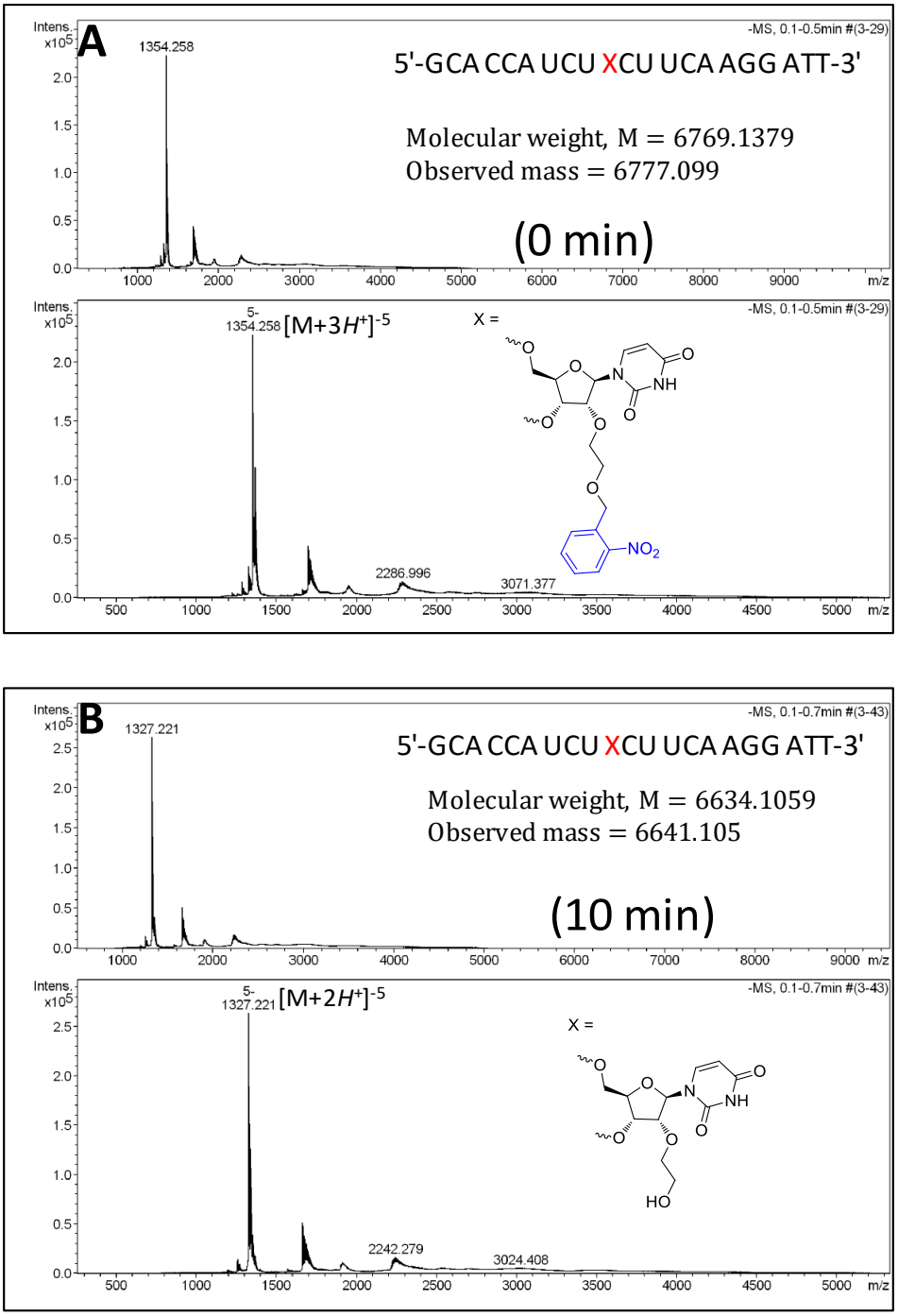


Figure S19. ESI-MS of U-NBnOE **RNA S5**, before (A) and after (B) irradiating at 365 nm for 10 min. (6777.099-6641.105 = 135.994 mass difference per one nitrobenzyl group) indicating the photo cleavage of 2-nitrobenzyl group.

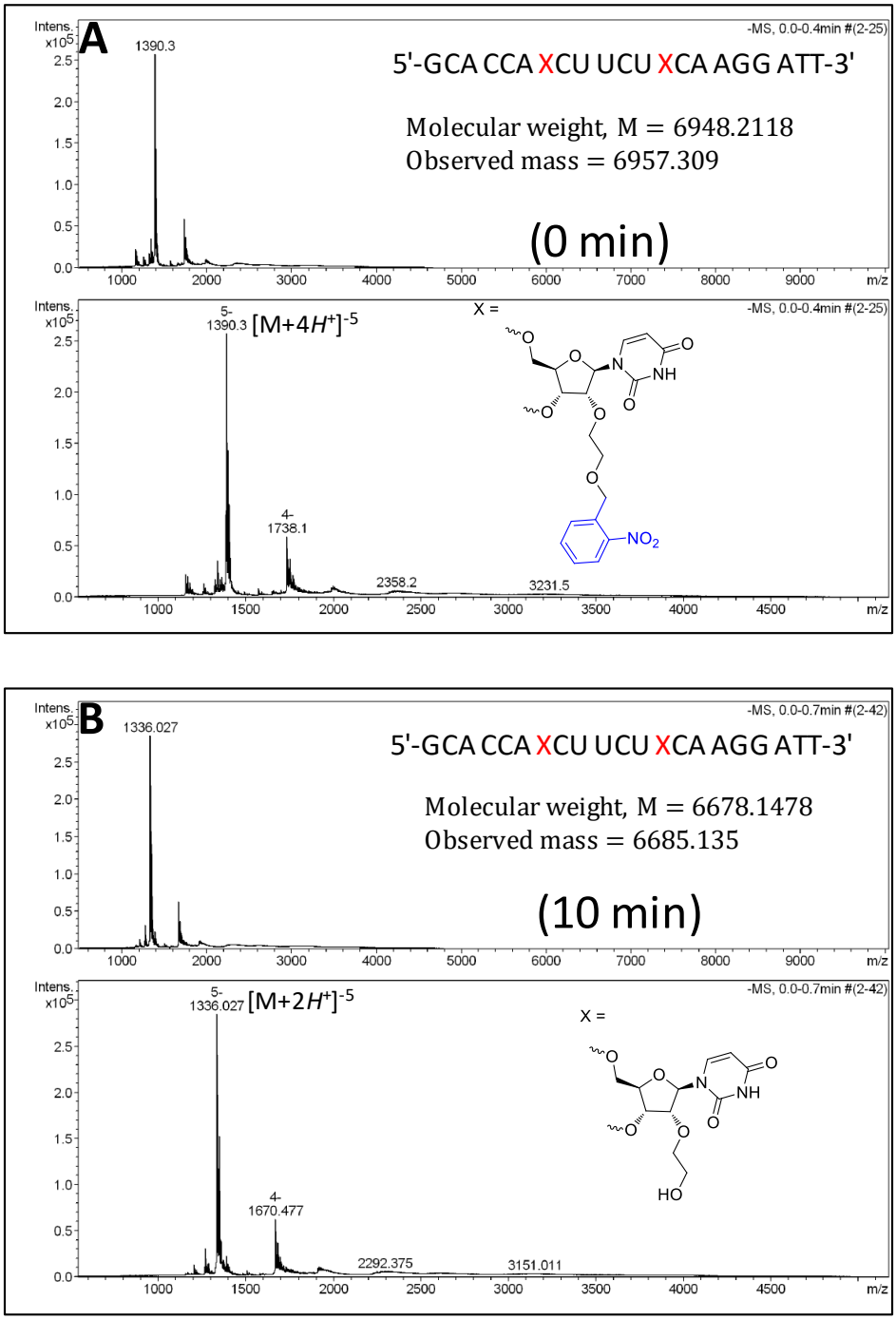


Figure S20. ESI-MS of U-NBnOE **RNA S6**, before (A) and after (B) irradiating at 365 nm for 10 min. (6957.309 -6685.135 = 272.174)/2 modifications = 136.087 mass difference per one 2-nitrobenzyl group) indicating the complete photo cleavage of two 2-nitrobenzyl groups.

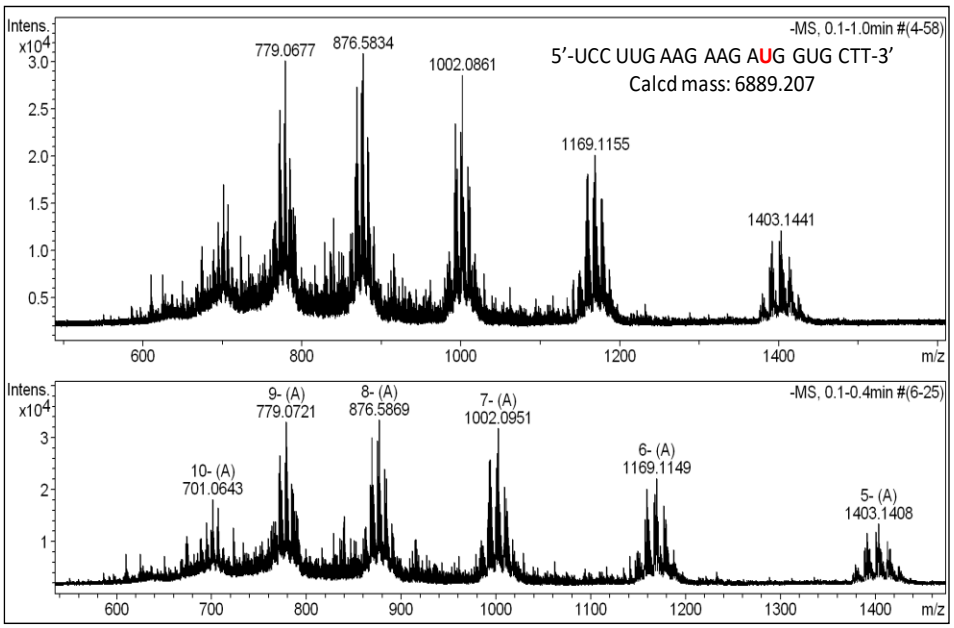


Figure 21. ESI-MS of U-NBnOE **RNA AS2,** mass is appeared in different charge state with adducts (Na^+/K^+).

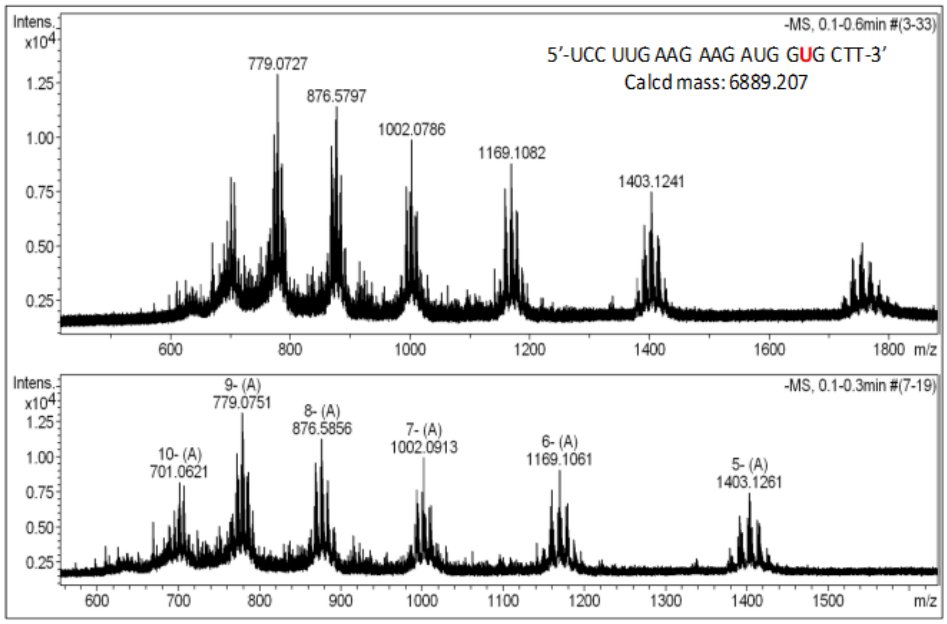


Figure 22. ESI-MS of U-NBnOE **RNA AS3**, mass is appeared in different charge state with adducts (Na^+/K^+).

8.CD spectra of U-NBnOE siRNAs before and after exposure of light (365 nm).

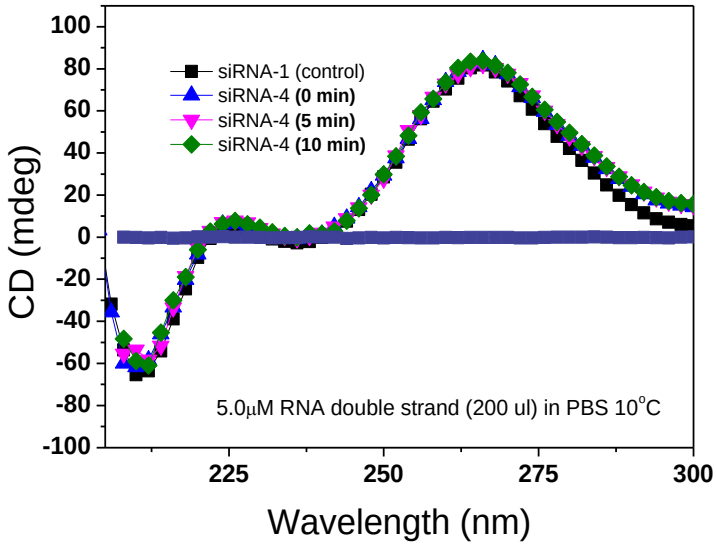


Figure S23. CD spectra of U-NBnOE **siRNA-4** in PBS buffer at different exposure times.

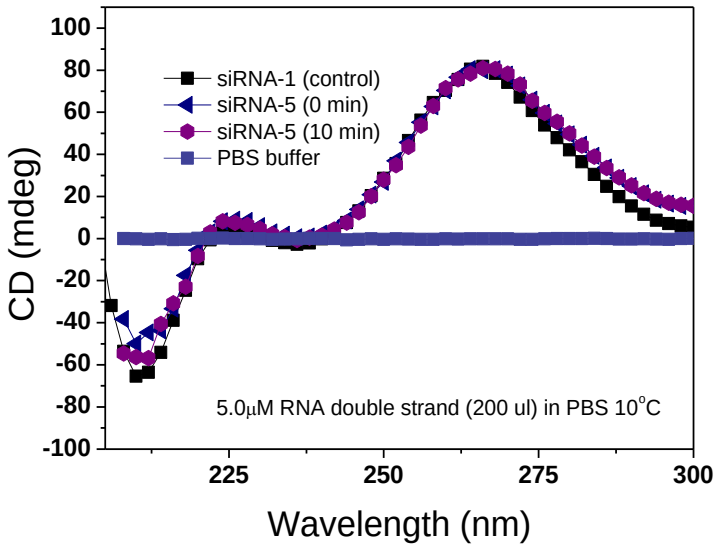


Figure S24. CD spectra of U-NBnOE **siRNA-5** in PBS buffer at different exposure times.

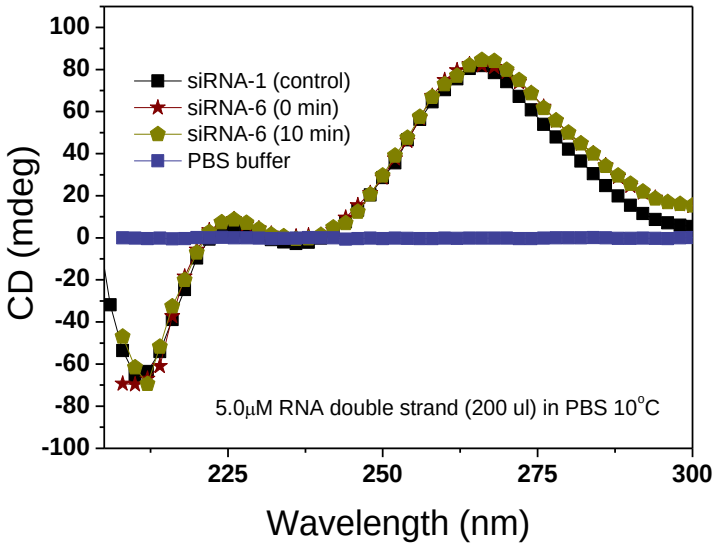


Figure S25. CD spectra of U-NBnOE **siRNA-6** in PBS buffer at different exposure times.

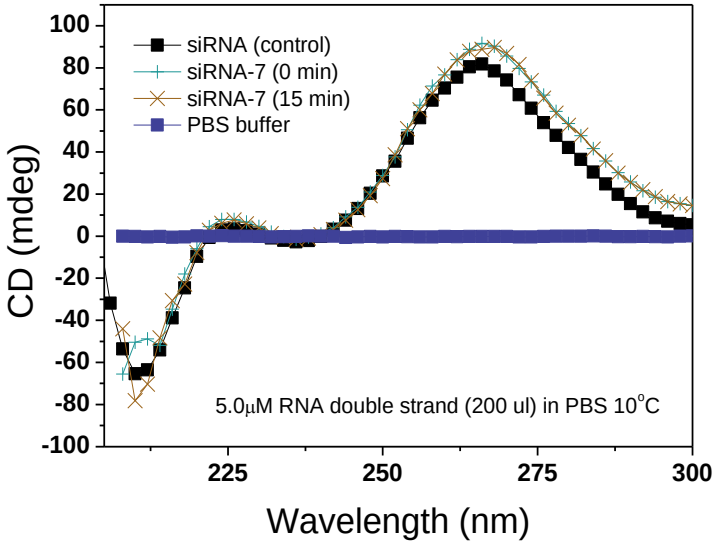


Figure S26. CD spectra of U-NBnOE **siRNA-7** in PBS buffer at different exposure times.

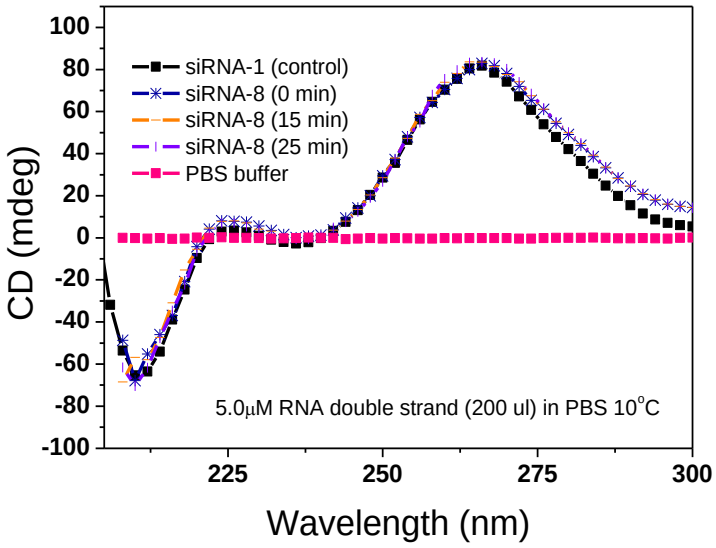


Figure S27. CD spectra of U-NBnOE **siRNA-8** in PBS buffer at different exposure times.

CHAPTER 2

Part-B

Synthesis of 2'-*O*-(*o*-nitrobenzyloxyethyl)adenosine, N^1-/N^6-nitrobenzyl-2'-*O*-hydroxyethyl-adenosine Analogues and their Incorporation into DNA/RNA

Chapter 2B. Synthesis of 2'-*O*-(*o*-nitrobenzyloxyethyl)adenosine, N^1-/N^6-nitrobenzyl-2'-*O*-hydroxyethyl-adenosine Analogues and their Incorporation into DNA/RNA

2B.1 Introduction

DNA methylation has been center of attention in cancer research for more than three decades.[1] In recent times, the identification and understanding the role of naturally occurring modified RNA in regulating the biochemical pathways such as gene regulation, translations and RNA splicing mechanisms are considerably increasing.[2-5] Apart of unmodified (normal) RNA functioning, the importance of naturally occurring modified RNA functioning is indisputable. More than 100 nucleoside modifications have been identified from different types of RNAs such as tRNA, sRNA and mRNA.[6] Most of these modifications are structurally diverse with simple modifications such as, alkylated, thiolated, isoforms, and complex modifications such as, having attached amino acid residue or heterocyclic ring.[7] Methylated nucleosides such as nucleobases methylated and 2'-*O*-methylated nucleosides are most common modifications found in different types of RNAs. The most abundant internal modification of mRNA is N^6-methyladenosine (m^6A), found in both internally and polyadenylated tails of mRNA and possess essential RNA regulatory mechanism.[8] They significantly accelerate protein translations.[9] Parisien and co-workers have found that m^6A facilitate the structural changes to mRNA and enhance binding interactions with heterogeneous nuclear ribonucleoprotein C (HNRNPC) in human cells. The absence of m^6A resulted in weak binding interactions and this regulatory role of m^6A is termed as 'm^6A-switch'. Okamura and co-workers have shown that specific inhibition of m^6A-RNA resulted in elongated circadian periods due to decrease in RNA processing efficiency; and considered m^6A as important pacemaker in

circadian gene expression and translations.[10, 11] The reversible site specific methylations of adenosine at N^6- and 2'-OH positions of RNA have been identified in mammalian cells and these reversible methylations are regulated by RNA methyltransferases (writers) and demethyltransferases (erasers) m^6A readers such as Methyltransferase Like 3 (METTL3)[12], Methyltransferase Like 14 (METTL14)[13] as 'writers', Fat mass and obesity-associated proteins (FTO)[14] and ALKBH5 as 'erasers' [15] and YTHDF-1/2, IGF2BP1 as 'readers' [16]. The recent studies reveal that the RNA methyltransferases (writers) and demethyltransferases (erasers) are upregulated in human cancer tissues.[17] The specific inhibitors of m^6A writers, readers, and erasers proteins are still at infancy and these inhibitors could be potential candidates for anti-cancer treatment. Mononucleotide, m^6A binding with YTH domain of YTHDF2 (m^6A reader protein) has been demonstrated.[8]

However, along with me^6A, N^1-methyl adenosine, me^1A is also found abundantly in noncoding RNAs and more prevalent in 5'untranslated regions (5'UTR) of *Homo sapiens* mRNA.[18, 19] The me^1A has a methyl group at Watson-Crick hydrogen bonding interface of complimentary binding site, which influences the regulatory functioning of RNA. Most importantly, the me^1A can rearrange to me^6A under neutral and basic pH, *via* Dimroth rearrangement.[20] This unusual rearrangement (me^1A to me^6A) of me^1A makes it difficult in handling and conducting biochemical and biophysical experiments.[18] Zhou and co-workers developed methods to detect me^1A-RNA with CRISPR-Cas13a/C2c2 system.[21]

Another N^6-alkylated adenosine is N^6-isopentenyladenosine (i^6A) which is also abundant. The i^6A nucleoside have an important role in cytokinin biosynthesis and also exert *in-vitro* and *in-vivo* anti-proliferative effects.[22-25] The i^6A nucleoside correlates with the expression and activity of farnesyl pyrophosphate synthase (FPPS), a key enzyme involved in the mevalonate (MVA) pathway, which

is aberrant in brain cancer. Recently, Bifulco and Co-workers (2017) designed and synthesized i^6A nucleosides by the introducing diverse chemical moieties at the N^6 position of adenosine. They tested the efficacy in U87 cells and in primary glioma cultures, derived from patients.[26] Importantly, N^6-benzyladenosine (Bn^6A) was found to be more potent than the naturally occurring i^6A. In the same article they propose that Bn^6A and its derivatives should consider as potential bioactive molecules in anti-glioma research.

Along with nucleobase methylated RNAs, many 2'-*O*-methylated (2'-*O*-me) RNAs also exist. Ribose 2′-*O*-methylation occurs in rRNA, tRNA, mRNA, and siRNA etc. at 2'-*O*H of adenosine (Am), guanosine (Gm), cytidine (Cm), and uridine (Um) nucleobases[27] and occurs in viruses, archaebacteria, eubacteria, yeasts, protists, fungi, and higher eukaryotes.[28] The presence of methyl group increases the hydrophobic character and also eliminates the canonical hydrogen bonding with nucleobases. Importantly, the presence of methyl group at 2'-OH comparatively resists the sugar ring puckering from dynamic 2'-endo and 3'-endo conformations and affects the global conformation of RNA. The 2'-OMe acts as restriction cap in 2'-OH involved in 3'-*O*-phosphodiester bond cleavages which are catalyzed *exo/endo* nucleases. The RNA duplexes having 2'-*O*-me nucleosides exhibit higher thermal melting temperatures.[29]

In development of RNA based therapeutic drugs, many modified RNA analogues are synthesized by the structural modification at their phosphate backbone, sugar ring, 2'-OH, and nucleobase residues.[30] The 2'-OH group of native RNAs are sensitive to *exo/endo* nucleases and prone to degrade in cellular environments which is major hurdles to use synthetic RNA as drugs.[31, 32] Thus 2'-OH capped RNA analogues are developed to counter with nuclease enzymes. The most common is 2'-*O*Me RNA which also exist in cellular RNA to affect secondary structure of RNAs and their interaction with proteins.[33] Further, 2'-OH modified RNA are prepared as 2'-tethered RNA

analogues which shows remarkable stability against *exo/endo* nuclease.[34] For examples, 2'-ethoxy, 2'-tethered hydroxyl ethyl, methoxy ethyl, aminoethyl, guanidinoethyl, allyl, and cynoethyl groups are introduced at 2'-OH of RNAs to enhance their half-life in serum.[35-37] Moreover, 2'-OH of RNA is caged with photo labile group (o-nitrobenzyl) for controlling the cleaving properties of hammerhead ribozyme with light.[38] The *o*-nitrobenzyl group has been introduced at N^6 position of adenine residue for caging (temporarily blocking) the Watson-Crick hydrogen bonding interface of adenine residue, and attempted to control the biological activity with light. Metzker and co-workers have synthesized N^6-(*o*-nitrobenzyl)-dATP which is sequence terminator in enzymatic DNA synthesis.[39] Recently Heckel and Co-workers have synthesized N^6-(*o*-nitrobenzyl)-adenosine analogues and incorporated into anti-miRNA oligonucleotides for controlling the light dependent angiogenesis.[40]

2B.2 Objective of our work

From brief review of RNA modifications, the development of caged nucleosides with 2'-tethered alcohol capped with photo caged group is certainly needed. These modifications are intended to exhibit conformational constraints without affecting the Watson-Crick base pairing and further, these constraints can be controlled with light. Additionally, there are only few literature reports of N^1 and N^6 alkylated adenosines, however, very few reports discuss *photo labile* N^6-adenosine nucleosides. There are no literature reports of N^1 and N^6 adenosine nucleosides having *photo labile* and *2'-tethered* groups. Since, N^1 and N^6 alkylated adenosines will directly interfere with the Watson-Crick base pairing; certainly, N^1-/N^6-*caged* adenosines with 2'-*tether* modifications needed to be examined. We took this opportunity to explore the synthesis of *2'-tether capped* with *photo labile* adenosine nucleoside (Figure 2B.1A) and *photo labile* N^1-/ N^6- alkylated adenosine

nucleosides (Figure 2B.1B/1C). Herein, we discuss the chemical synthesis of N^1 and N^6 (*o*-nitrobenzyl)-2'-tethered adenosine nucleosides, 2'-*O*-(*o*-nitrobenzyloxyethyl)adenosine nucleoside, their phosphoramidites, DNA/RNA oligonucleotides and their photochemical properties.

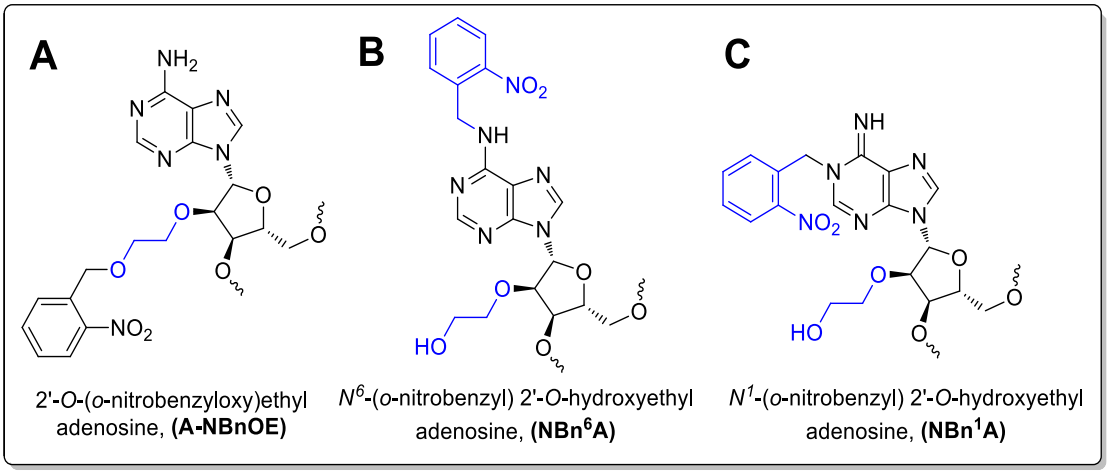


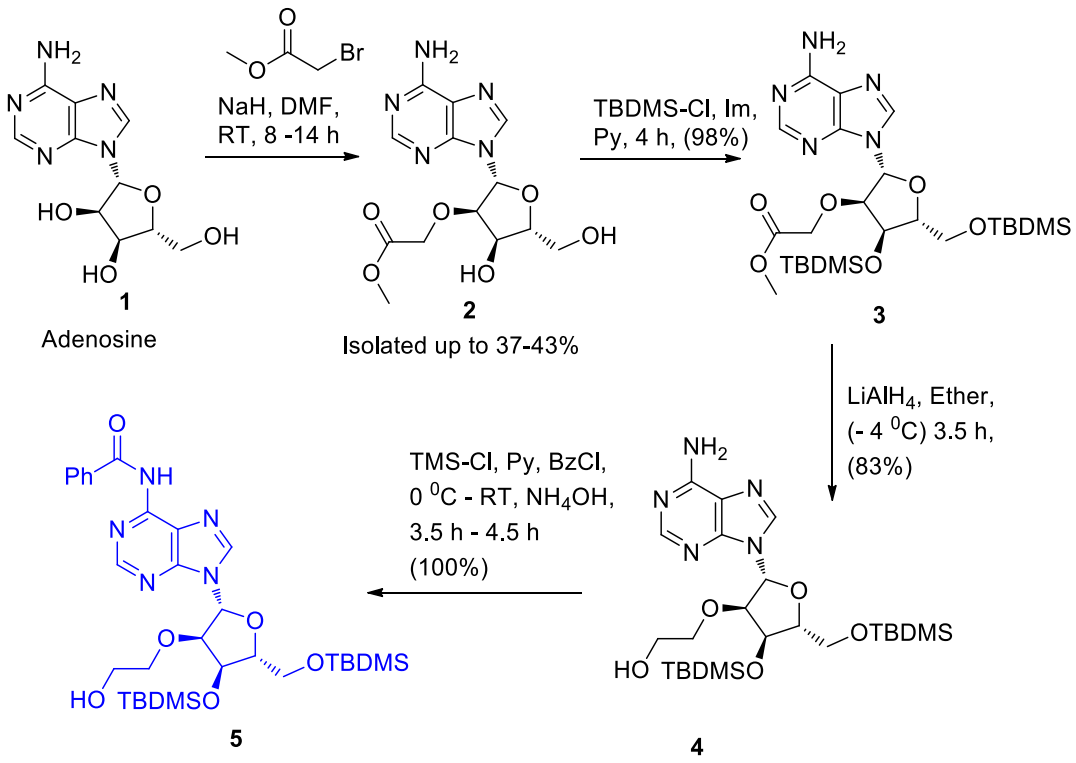
Figure 2B.1. Rationally designed caged adenosine modified nucleosides described in this section.

2B.3 Results and Discussion

Synthesis of photo labile N^1 and N^6 alkylated adenosine nucleosides commence with the preferential 2'-OH alkylation of adenosine with methyl bromoacetate/NaH provided nucleoside **2** in 37-43% isolated yields following reported procedure.[37] The 3'-OH and 5'-OH of nucleoside **2** were protected with TBDMS protecting group to produce 98% yield of nucleoside derivative 3. The methyl ester in nucleoside **3** was reduced to alcohol with lithium aluminum hydride to obtain 2'-*O*-(2-hydroxyethyl) nucleoside derivative **4**. Initially we designed to attach the photo labile group, *o*-nitrobenzyl group at 2'-tethered hydroxyl group under basic conditions with *o*-nitrobenzylbromide. So to avoid the alkylation at N^6 position of adenosine derivative **4** with

reactive *o*-nitrobenzylbromide, we protected the N^6 of adenosine derivative of **4** as its N^6 benzolyl adenosine (Bz^6A) derivative **5**, which is obtained by protection of 2'-tethered hydroxyl group with transient TMS protecting group followed by benzolylation of amine (Scheme 2B.1).

Scheme 2B.1. Synthesis of 2'-*O*-hydroxyethyl-Adenosine nucleoside analogues.



To facilitate the alkylation with *o*-nitrobenzylbromide at naked hydroxyl group in nucleoside **5**, all the functional groups were fully protected. The alkylation of nucleoside **5** under standard NaH/DMF conditions, which is widely used for the alkylation of hydroxyl groups in nucleic acid chemistry, was resulted in a mixture of compounds. These products were difficult to isolate and characterize in pure forms with very poor yields. In our hands, other reactions conditions such as, imidazole (Im)/pyridine or PPh_3/DIAD/*o*-nitrobenzylalcolol were ineffective even after varying the temperature. Instead of alkylating the *in-situ* ribonucleoside alkoxide of **5**, we converted

nucleoside **5** into its mesylate to perform electrophilic substitution reaction with *o*-nitrobenzyl alcohol. The resultant reaction mostly dominated with beta-elimination products along with partial deprotection of TBDMS groups.

However, in the reaction of nucleoside **5** with *o*-nitrobenzylbromide under K_2CO_3/AcCN conditions, two clean products were isolated, Spot-1(upper spot) and Spot-2 (lower spot) in ~ 1:1 ratio with 87% yield with different Rf values on TLC. The preliminary evaluation of proton NMR spectra Spot-1 and Spot-2 indicate the presence of nitrobenzyl group in both the isolated products. Surprisingly the mass of both these products are similar to that of anticipated target molecule, m/z 778.354. The mass spectra of both Spots-1/2 are depicted in Figure 2B.2A/2B which exhibit important peaks as 1580.8955 $(2M+Na)^+$/801.3426 $(M+Na)^+$/375.1209 for Spot-1, and 1558.7104 $(2M+H)^+$/779.3620 $(M+H)^+$/375.1231 for Spot-2. After a very careful evaluation of observed fragment ions at m/z 375.12 in both mass spectra, we analyzed this fragment ion as protonated nuclobase by cleavage of glycocidic bond.[41] This indicates that, none of the two Spots-1/2 fragment ions belongs to the target molecule's fragment ion, m/z 539.27. However, the exact position of nitrobenzyl group is not determined from these mass analyses.[41] These mass results confirm the alkylation with *o*-nitrobenzylbromide of compound 5 produced nucleobase nitrobenzylated products in 1:1 ratio instead of designed target molecule (Figure 2B.2C).

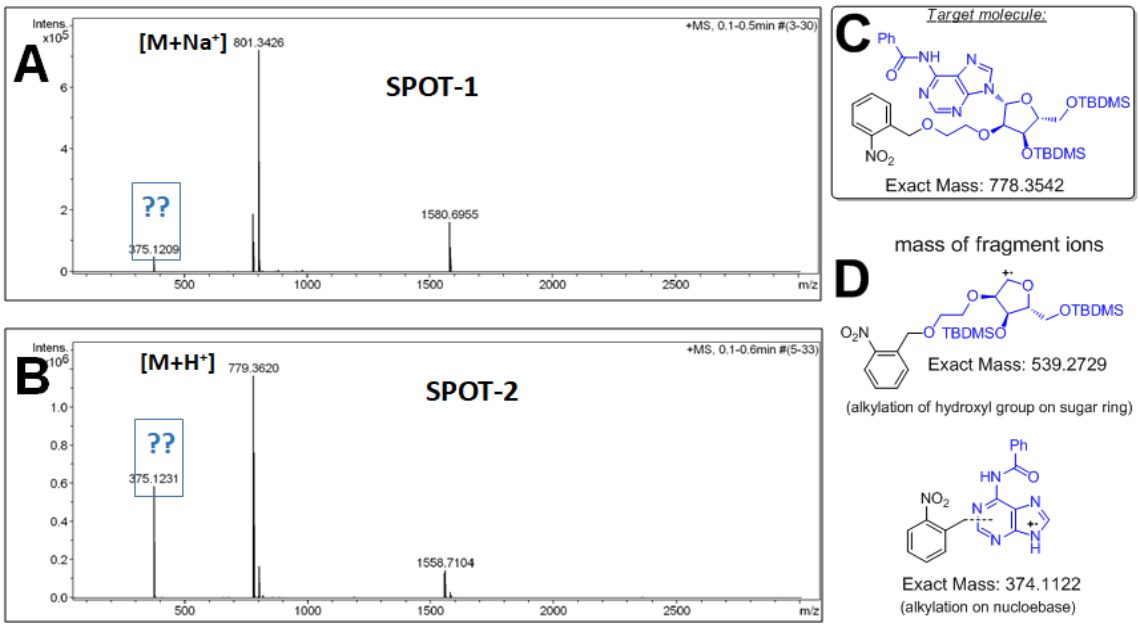


Figure 2B.2. ESI-Mass spectra of Spot-1(A) Spot-2 (B) and structural mass analysis of observed mass peaks (C/D).

To assign the structure of these products (Spot-1/2), possible alkylation sites in nucleoside **5** was examined from its structure. From the structure evaluation of nucleoside 5, there are three reasonable alkylation sites from their corresponding amide tautameric forms as shown in Figure 2B.3. The reaction of reactive aryl halide, *o*-nitrobenzylbromide, can alkylate all the three tautamers, interestingly we observed only two alkylated products.

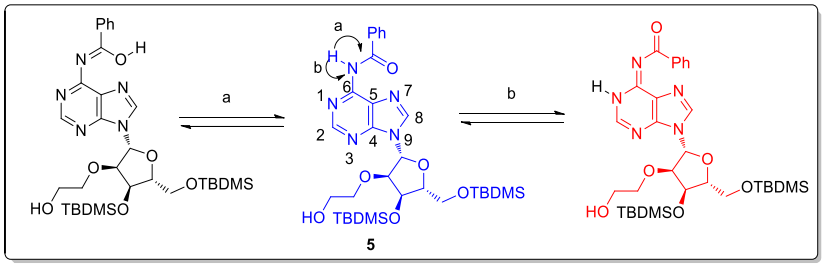
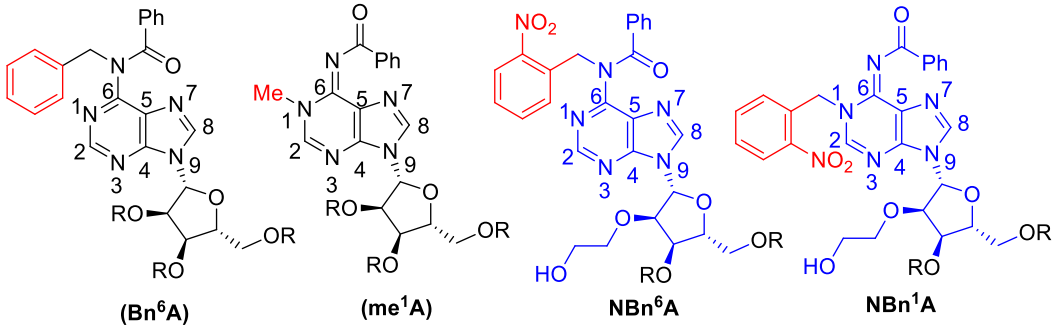


Figure 2B.3. Possible isomers of adenosine derivative 5.

Sekine and Co-workers have exclusively studied the alkylation products of N^6-acylated adenosine derivatives. They found, when benzoyl group was preferred as N-acyl group on adenosine derivatives, the exclusive N^1- and N^6- alkylated products were isolated from possible three amide tautomers. Importantly, the structure of these N^1-/N^6- alkylated derivatives were also differentiated and assigned by NMR spectroscopy. In general, the N^6-benzoyl adenosine derivatives (Bz^6-A) have a significant difference in ^{1}H and ^{13}C NMR spectra of adenosine residues in N^6- and N^1-alkylated products. The ^{1}H-NMR chemical shifts (δ) of 2-H and 8-H protons of N^6- and N^1-alkylated products are significantly different. The ^{1}H-NMR signals for 8-H in N^6-alkylated products is observed between δ8.19-8.45 ppm, while those for the N^1-alkylated products are observed at δ7.87 ppm. The ^{1}H-NMR signals for 2-H are shifted to lower ppm compared with those of N^6-alkylated products. Similarly, in the ^{13}C NMR spectra, the signals for 2-C, 4-C and 8-C in the N^1-alkylated products appear markedly lower ppm compared with those for the N^6-alkylated products.

Herein, the structure assessment of the two isolated spots (Spot-1/2) obtained by the alkylation of nucleoside **5** with *o*-nitrobenzyl bromide is readily assigned by comparing their NMR data (^{1}H and ^{13}C) with closely related N^1- and N^6-Bz-A derivatives (Table 2B.1).

Table 2B.1. $^{1}H/^{13}C$-NMR ($CDCl_3$) data of reported (Bn^6A and me^1A) and synthesized compounds (**6a/6b**).



R = TBDMS

Entry	δ (ppm)	**6a** (Spot-1)	Bn^6A*	**6b** (Spot-2)	me^1A*
1	2-H	8.54	8.51	8.06	7.95
2	2-C	152.28	152.59	146.75	146.52
3	4-C	152.08	151.8	144.99	145.44
4	5-C	126.91	127.08	122.76	122.42
5	6-C	153.81	154.03	148.14	147.65
6	8-H	8.25	8.19	7.95	7.87
7	8-C	142.44	142.46	138.45	138.58
8	C=O (Bz)	172.34	172.5	176.88	170.61

δ: Chemical shift; *Reported chemical shift values.[42]

There is a significant difference in the ^{1}H-NMR chemical shift of 2-H and 8-H protons for the Spot-1 and Spot-2. The signal for 8-H in Spot-1 is observed at δ8.25 ppm, while in Spot-2 it is appeared at δ7.95 ppm. The signals for 2-H in Spot-1 is observed at δ8.06 ppm while in Spot-2, appeared at δ8.51 ppm. Similarly, in the ^{13}C NMR spectra, the signals for 2-C, 4-C and 8-C in Spot-2 are shifted markedly to lower ppm compared with Spot-1. From the observed NMR data and careful analysis by comparing with closely related N^6- and N^1-alkylated products of N^6-benzoyl adenosine derivatives, we conclude that the Spot-1 is N^6-alkylated product **6a** (NBn^6A), where as Spot-2 is N^1-alkylated product (NBn^1A) (Figure 2B.4).

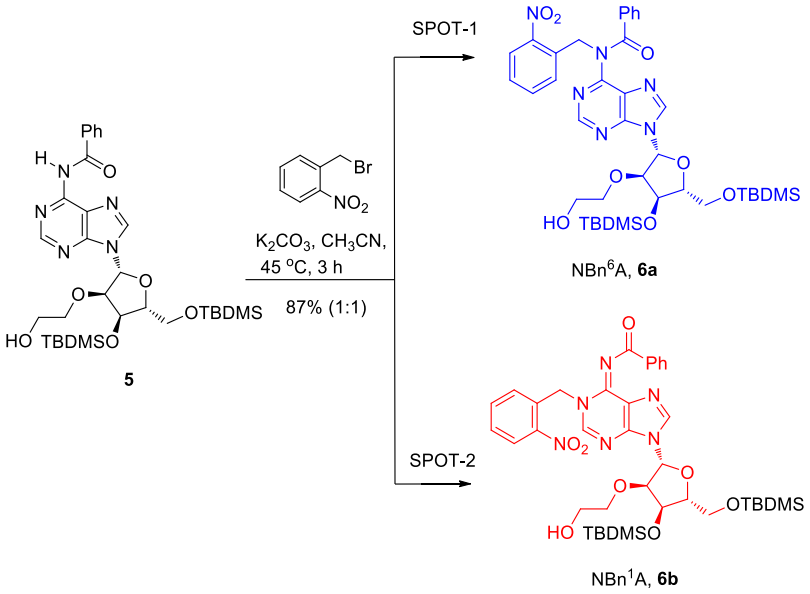
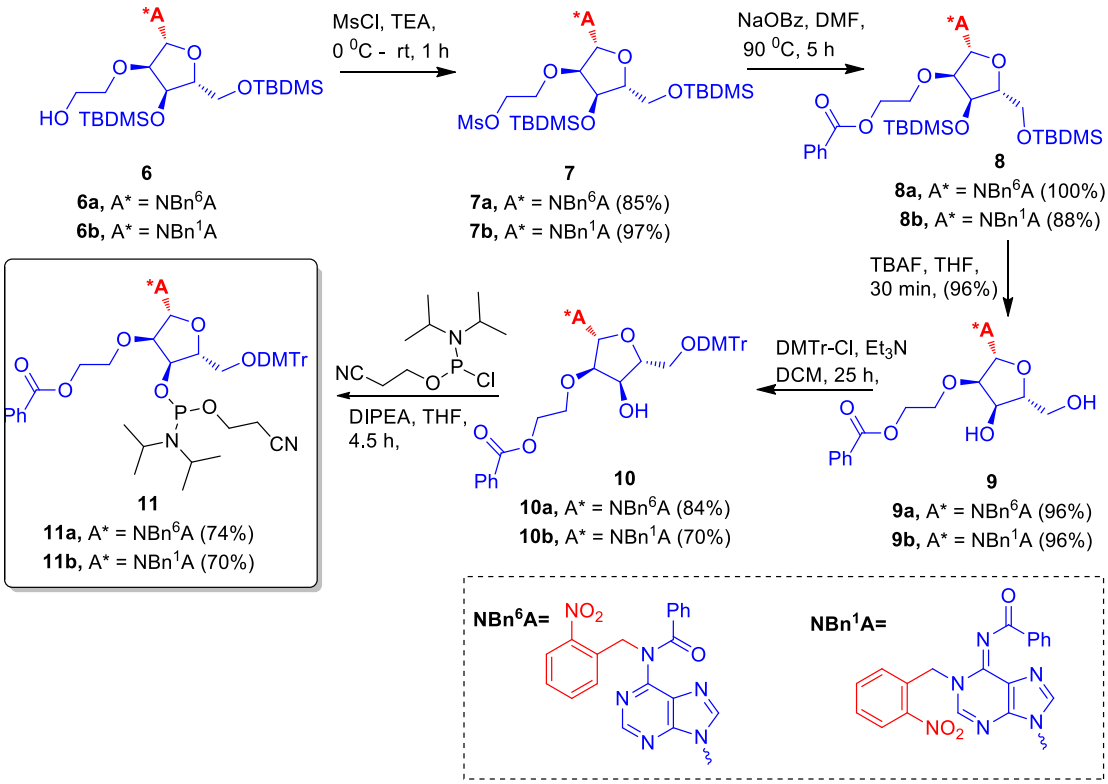


Figure 2B.4. Alkylation products of adenosine derivative **5**.

After the structure of **6a** (NBn^{6}A) and **6b** (NBn^{1}A) are confirmed, we diverted our efforts to explore the chemistry of these molecules. To incorporate into DNA/RNA oligonucleotides, we synthesized their corresponding phosphoramidites from **6a** (NBn^{6}A) and **6b** (NBn^{1}A) as shown in Scheme 2B.2. To avoid the intervention of naked hydroxyl group in **6a/6b** during DNA/RNA oligonucleotide synthesis, it is necessary to protect this functional group. For synthetic advantage we opted to protect this hydroxyl group in benzoate form but the direct benzolylation with benzoyl chloride of **6a/6b** were surprisingly not successful. Hence, we converted the hydroxyl group to their reactive mesylate derivatives **7a/7b** and then treated with sodium benzoate at 90 ^{o}C to obtain **8a/8b** derivatives in 90-96% yields. For 5'-*O*-DMTr and 3'-*O*-phosphoramidations, we removed the TBDMS protecting groups with TBAF to obtain **9a/9b** nucleosides in 95-96% yields. The 5'-OH was protected as DMTr derivatives (**10a/10b**) and 3'-OH is converted into their corresponding phosphoramidites **11a/11b** in 70-74% isolated yields. All the compounds are characterized by NMR and ESI-MS and their respective spectra are provided in Appendix.

Scheme 2B.2. Synthesis of *caged*-NBn^{6}A/NBn^{1}A phosphoramidities.



Cleavage of PLPG with light. Since *o*-nitrobenzyl is a known as photo labile protecting group, we planned to optimize the cleavage conditions for the removal of *o*-nitrobenzyl group with light (wavelength 365 nm) from nucleoside **9a/9b**. The photo cleavage reaction is monitored by NMR and Mass techniques. The photo cleavage of NBn^{6}A (**9a)** and NBn^{1}A (**9b)** derivatives are studied by 1H-NMR with varying exposure times at 365 nm in $CDCl_3$:D_2O (25:1). D_2O was added to interpret the 3'/5'-hydroxyl groups from sugar ring protons. 1H NMR spectra without/with D_2O (20 µl) are given in Appendix. The time dependent NMR spectra of caged nucleoside **9a** (NBn^{6}A) are provided in Figure 2B.5. In photo cleavage of NBn^{6}A (**9a**), 90-95% of *uncaging* occurred within 45-50 min of exposure at 365 nm. Before exposure, 1'-H of sugar ring appeared as doublet at δ5.83 ppm and benzylic -CH_2- as singlet at δ6.03 ppm. During the exposure time, the benzylic -C<u>H</u>$_2$- disappeared and a new peak at δ12.1 ppm appeared, which belongs to the aldehydic proton

of nitrsobenzaldehyde after photo cleavage. While the 1'-H of sugar ring is shifted from 5.83 ppm to 6.0 ppm after photo cleavage. In literature, the formation such aldehyde is also known from the photo cleavage of nitrobenzyl group.[43] Besides, 1'-H proton in sugar ring shifted from δ5.83 to δ6.0 after 60 min, this is possibly due to the conformational changes after *uncaging*.

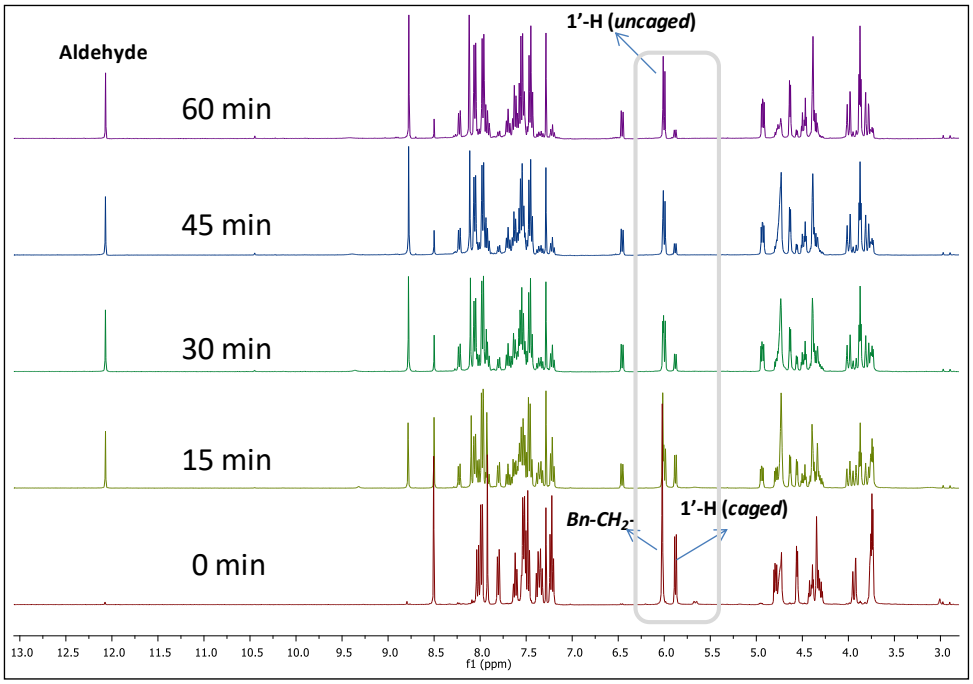


Figure 2B.5. ^{1}H-NMR spectra for photo cleavage of PLP group in *caged*-adenosine nucleoside **9a** with 365 nm UV light in $CDCl_3$.

Similarly, the photo cleavage study of NBn^{1}A (**9b**) derivative is provided in Figure 2B.6. The 1'-H of sugar ring appeared as doublet at δ5.83 ppm before exposure. Unlike NBn^{6}A (**9a**) derivative, the benzylic -C$\underline{H}_2$- appeared as doublet of doublet at δ5.75 ppm, this is possibly due to germinal coupling. In contrast to photo cleavage of NBn^{6}A (**9a**), there are no changes in NMR spectra even after exposure for 60 min at 365 nm and a similar experiments repeated in CD_3CN and found similar result, (1H NMR spectra provided in Appendix). These NMR studies strongly support the

photo cleavage of *o*-nitrobenzyl group of adenosine analogue **9a** (NBn^6A) while no photo cleavage of *o*-nitrobenyzl group in adenosine analogue **9b** (NBn^1A).

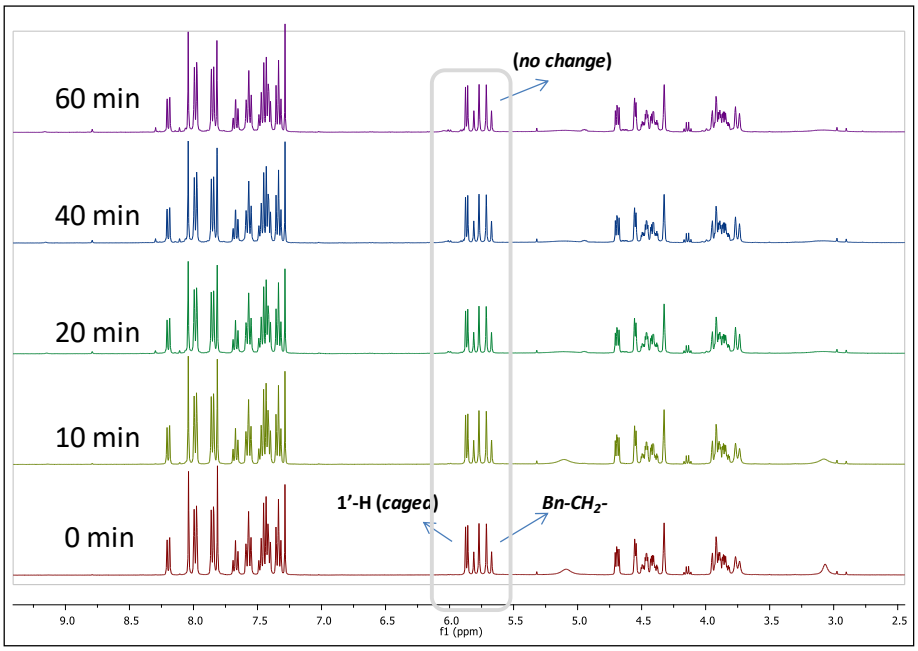


Figure 2B.6. ^{1}H-NMR spectra for photo cleavage of PLP group in *caged*-adenosine nucleoside **9b** with 365 nm UV light in $CDCl_3$.

UV-Vis and CD studies. Further we attempted to characterize nitrobenzyl protected adenosine derivatives (**6a**/**6b**/**9b**) with UV-Vis and CD techniques. The UV /CD spectra of NBn^6A and NBn^1A derivatives exhibit characteristic and significant variation in their respective spectra. In UV spectra, the NBn^6A derivative, **6a** exhibit absorption at λ280 nm while NBn^1A derivative **6b** has absorption at λ311 nm (Figure 2B.7A). Moreover, there are no significant changes in UV spectra of NBn^1A derivatives **6b**/**9b**. This observed shift in UV spectra of NBn^1A derivatives **6b**/**9b** (λ311 nm) compared to NBn^6A derivative **6a** (λ280 nm) is probably due to perturbation in aromatic property of adenosine nucleobase which is preserved in NBn^1A derivative **6a**. In CD spectra, both

NBn^{6}A and NBn^{1}A derivatives exhibit negative bands which are differ in absorption wavelengths. The NBn^{6}A derivative, **6a** has negative bands at λ280 and λ211 nm whereas NBn^{1}A derivative **6b** has broad negative band between λ260-320 nm, while in **9b** has distinguished sharp peaks at λ265 and λ300 nm (Figure 2B.7B). This distinguished absorption in UV /CD spectra, makes them useful tools to readily differentiate NBn^{6}A and NBn^{1}A alkylated nucleosides at monomer level.

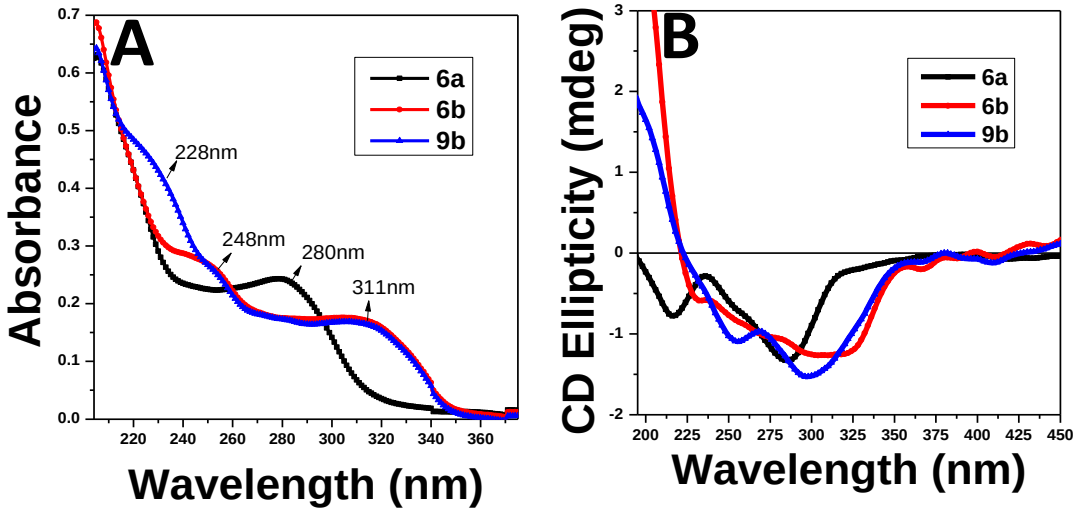


Figure 2B.7. UV spectra (A) and CD spectra (B) of NBn^{6}A **6a** and NBn^{1}A **6b/9b** derivatives in MeOH.

The photo cleavage of **6a/6b/9b** was also monitored by UV spectrophotometer with exposure of light (365 nm) at different time intervals. Before exposure (0 min) the UV spectra of **6a** exhibit absorption at λ280 nm (Figure 2B.8A). Upon irradiation of light (365 nm) the intensity of peak at 280 nm increases (hyperchromic shift) with respect to time up to 25 min., and further exposure of light up to 40 min have no significant changes. Then these **6a** samples (before light or after light exposure) were analyzed by ESI-Mass. The mass spectra of **6a** samples are provided in Figure 2B.9, which show the molecular ion peak of **6a**, at 779.4106 (m/z) $[M+H]^+$ before exposure with

light. After 40 min., exposure of light (365 nm), the mass spectra of **6a** exhibit a new peak at m/z 644.3857, this belongs to the molecular mass of *uncaged*-adenosine nucleoside **5**. These analyses confirm the photo cleavage of nitrobenzyl group of NBn^{6}A derivative **6a**.

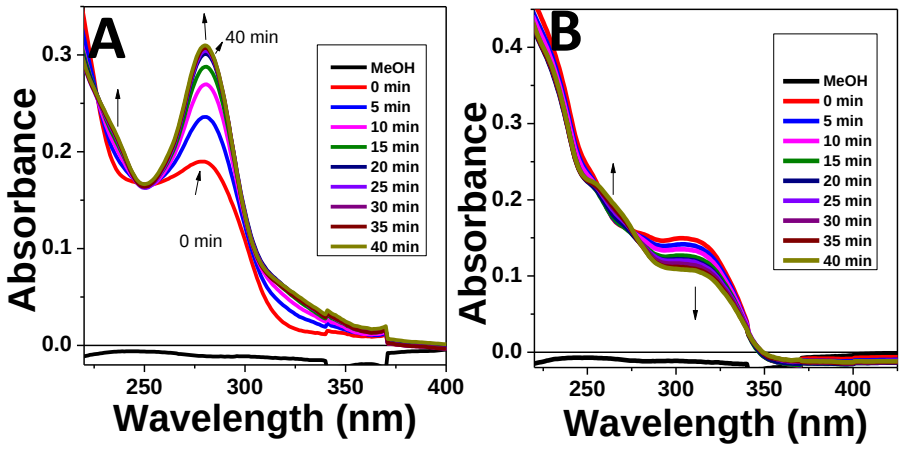


Figure 2B.8. UV spectra of NBn^{6}A, **6a** derivative (A) and NBn^{1}A, **9b** derivative (B) at different exposure times of light (365 nm)

The similar UV & Mass experiments were performed with NBn^{1}A derivatives **6b**/**9b**.The photo cleavage of NBn^{1}A derivatives, **6b/9b** has absorption at λ310 nm before exposure. Unlike **6a**, upon irradiation (365 nm) for 40 min, both the UV spectra of **6b**/**9b** exhibit small hypochromic shifts. The UV spectrum of **6b** is provided in Figure 2B.8B. The UV spectra of **6b** and mass data of **9b** are provided in the Appendix. Unlike **6a**, the mass spectrum of **9b** was remained same before and after exposure of the light. This result suggest that nitrobenzyl group of **6b/9b** are not photo cleavable. A similar photo cleavage result was also supported from our NMR studies (Figure 2B.5 and 2B.6).

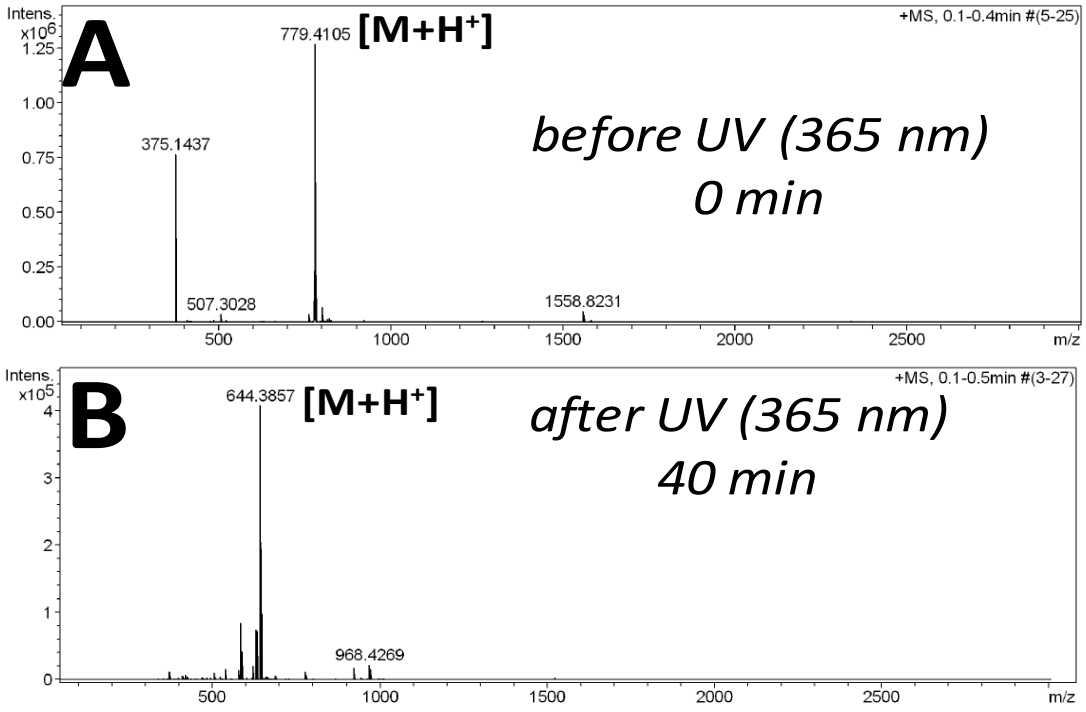


Figure 2B.9. ESI-Mass spectra of **6a** before (A) and after (B) exposure of light at 365 nm.

Before incorporation of NBn^6**A**/NBn^1**A** phosphoramidites into DNA/RNA, we examined chemical stability of nitrobenzyl group in modified adenosine nucleosides (**10a**/**10b**) under basic (NH_4OH solution) and also acidic (3% TCA in DCM) conditions by ESI-Mass studies. Their mass data are provided in the Appendix. In these studies, *o*-nitrobenzyl group was intact with adenosine residue with both compounds **10a**/**10b**. Thus their phosphoramidites could be incorporated into DNA/RNA oligonucleotides.

For further application of NBn^{6}A and NBn^{1}A, the synthesized adenosine modified N^6/N^1 alkylated phosphoramidites **11a/11b** were incorporated into DNA/RNA oligonucleotides and examined photochemical behavior by ESI-Mass studies. We incorporated newly synthesized phosphoramidites **11a/11b** into DNA pentamer 5'-TTATT-3' as 5'-TT(NBn^6**A)**TT-3' (NBn^{6}A-DNA) and 5'-TT(NBn^1**A)**TT-3' (NBn^{1}A-DNA) at oligosynthesizer under standard DNA synthesis

conditions, except the coupling times of **11a**/**11b** increased to 10 min. After the synthesis, these DNAs were cleaved from resin and incubated at 55 ^{o}C for 6 h in ammonium hydroxide for complete deprotection of nucleobase protecting groups. Further these were purified and desalted by passing through sephadex column and the fractions containing DNA were characterized by mass.

The mass spectra of 5'-TT(NBn^{6}**A)**TT-3', containing **11a** (NBn^{6}A) adenosine residue is depicted in Figure 2B.10. In its mass spectra, the mass of NBn^{6}A-DNA appeared as m/z 830.297 (Figure 2B.10A). This mass peak corresponds to the observed mass of m/z 1662.6 $[M\text{-}H^{+}]^{-1}$ in charge state z = 2, which is one unit less than the calculated mass (m/z 1663.18) of NBn^{6}A-DNA. This DNA was exposed with light (365 nm) for 5 min and again recorded their mass spectra. After exposure of light, the mass peak at m/z 830.297 is disappeared to trace amount and exhibit a new peak at m/z 762.7854 (Figure 2B.10B). This mass peak (m/z 762.7854) corresponds to the observed mass of m/z 1527.57 $[M\text{-}H^{+}]^{-1}$ in charge state z = 2, which is one unit less than the calculated mass (m/z 1528.06) of NBn^{6}A-DNA without *o*-nitrobenzyl group. The observed mass difference of 136 units is equal to mass of nitrobenzyl group. Thus the photo cleavage of nitrobenzyl group was occurred in NBn^{6}A DNA with light (365 nm) within 5 min.

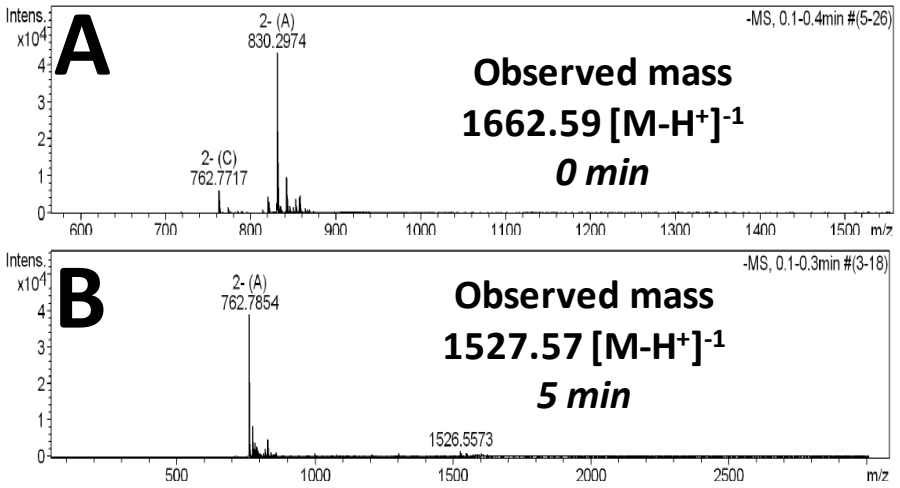


Figure 2B.10. Mass spectra of NBn^{6}A incorporated DNA pentamer before (A) and after (B) irradiation at 365 nm. Mass in observed in charge state z = 2. In Figure 8A, before exposure to light, small mass peak at m/z 762.77 is observed which corresponds mass of NBn^{6}A-DNA without nitrobenzyl group.

The incorporation of N^1-PLP phosphoramidite, **11b** in TT**A**TT is analyzed from its mass spectra. The mass spectra depict two mass peaks, m/z 762.78 and m/z 830.31 (Figure 2B.11A). Unambiguously, mass peak m/z 830.31 corresponds to the mass of NBn^{1}A DNA $(M-2H)^{2-}$ where M (1663.18). Another mass peak is analysed and found that peak at m/z 762.78 belong to $(M-2H)^{2-}$ where M (1528.57) is equal to the mass of DNA without containing nitrobenzyl unit. After exposure with light for 5 min, the mass peak at m/z 830.31 is disappeared (Figure 2B.11B). However, at monomer NBn^{1}A (**6b**), the cleavage of nitrobenzyl group was not observed with exposure of light even for 40 min. Therefore, we assume that the conversion of NBn^{1}A into NBn^{6}A is occurred during the synthesis of oligonucleotide via the Dimorth rearrangement. In literature, the conversion of N^1-alkyl adenosine to N^6-alkyl adenosine is reported through the Dimorth rearrangement reaction. For example i^{1}A (N^1-isopentyl adenosine) isomerize into i^{6}A (N^6-isopentyl adenosine) under basic conditions.[43-45]

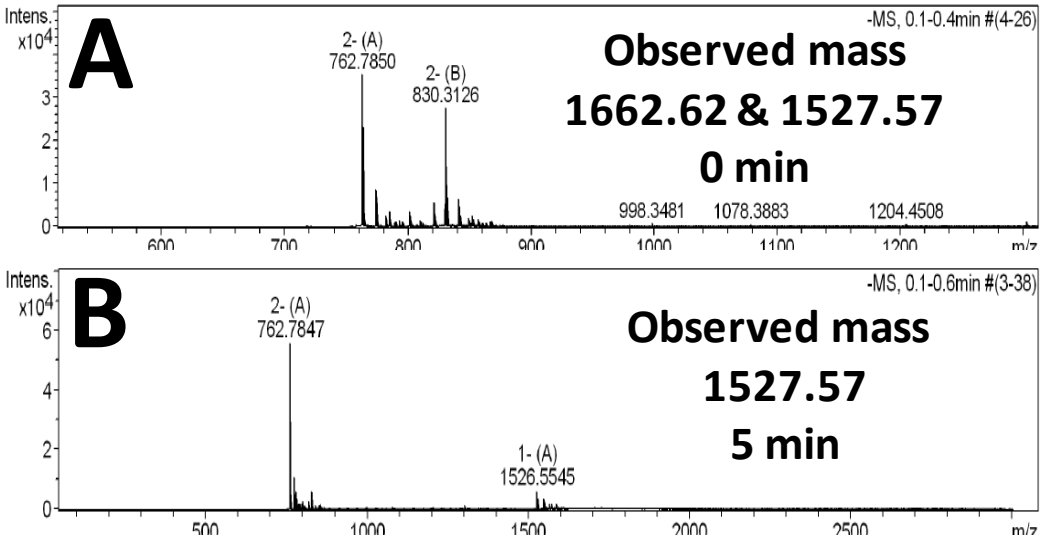


Figure 2B.11. Mass spectra of NBn^1A incorporated DNA before (A) and after (B) irradiation at 365 nm. Mass is observed in charge state z = 2.

Further, we incorporated NBn^6A and NBn^1A analogues into RNA oligonucleotides from respective phosphoramidites **11a** and **11b** at DNA synthesizer under standard RNA synthesis conditions, except the coupling times of **11a**/**11b** increased to 18 min. After the synthesis, these RNAs were cleaved and completely deprotected by incubating at 65 ^oC for 10 min in 1:1 mixture of methylamine (40% in water) and ammonium hydroxide. The 2'-*O*- Silyl (TBDMS/TOM) groups were deprotected with HF.TEA at 65 ^oC for 2 h. Further these were purified by 20% poly acrylamide gel electrophoresis (PAGE) and desalted by passing through sephadex-G25 columns. The fractions containing pure RNA were characterized by ESI-Mass.

The N^6 and N^1 -PLP phosphoramidites, **11a** and **11b** are incorporated in RNA (21mer) at different positions having one incorporation, two incorporations, and three incorporations. The NBn^6A and NBn^1A RNA sequence and position of modifications are provided in Table 2B.2 (Column 3). Their mass spectra are provided in the Appendix, while their deconvoluted mass values are given in Table 2B.2 (column 4). The mass of corresponding RNA oligonucleotides appears in their mass

spectra mostly in charge states z = 5 and z = 6 with sodium/pottasium adducts. The observed molecular mass of NBn^{6}A-RNA-1/2/3/4/5 (**11a** incorporated) are almost equal to the calculated ones (Entry 1-5, Table 2B.2) which confirm the syntheses of *o*-nitrobenzyl caged-RNAs (NBn^{6}A-RNAs). Unlike NBn^{6}A-RNAs, the observed molecular masses of NBn^{1}A-RNA-6/7/8 (**11b** incorporated) appeared to be significantly less than their calculated mass (Table 2B.2, Entry 6-8). The careful analyses of their mass peaks indicate the complete absence of photo labile (*o*-nitrobenzyl) groups. For example, the calculated mass of NBn^{1}A-RNA-6 (one **11b** incorporated) is 6792.19 (m/z), and its observed mass is 6746.17 (m/z) (Table 2B.2, Entry 6). Herein, we assign that this mass difference (-46.0 unit) occurs due to the chemical cleavage of nitrobenzyl group followed by adduct formation with Na^+/K^+. The similar mass differences are also observed with other NBn^{1}A-RNA-7 (one incorporated **11b**) and NBn^{1}A-RNA-8 (two incorporated **11b**), then we assigned successfully as the chemical cleavage of nitrobenzyl group followed by Na^+/K^+ adduct formation (Table 2B.2, Entry 7-8).

Photo cleavage of PLPG in RNAs. After successful synthesis of RNA oligonucleotides, we studied the photo cleavage of NBn^{6}A and NBn^{1}A RNA by ESI-Mass. Conceptually, the presence of photo labile, *o*-nitrobenzyl group add a mass of m/z 136 mass units to the total molecular mass of their respective RNA (for example mass of *caged*-RNA = M), when the *o*-nitrobenzyl group is photo cleaved with exposure of the light (365 nm), then there is a loss of mass by m/z 136 units for each unit of *o*-nitrobenzyl group from total molecular mass of that RNA (M-136). Herein, Each RNA oligonucleotide (500 pmol) was exposed at 365 nm using monochromatic laser light for 10 min and the mass spectrometric measurements were carried out before and after exposure of light. Their mass spectra are provided in Appendix. We also extracted their deconvoluted mass values which are given in (Table 2B.2, Column-5).

In N^6-photolabile modified RNA (NBn^6A-RNA-1/2/3/4/5), the photo labile groups were completely cleaved within 10 min of exposure. In NBn^6A-RNA-1 and NBn^6A-RNA-2, which possess single photo labile modifications which are differ in their position of modification, have a mass difference of m/z 135 before and after exposure. For example, the mass of NBn^6A-RNA-1 (one incorporated **11a**), (m/z 6866.21**),** was reduced to m/z 6731.11 after exposure of light (365 nm) for 10 min ((Entry 1, Table 2B.2). The reduced mass difference (~135.1 unit) is equivalent to the mass of *o*-nitrobenzyl, which confirms the photo cleavage of nitrobenzyl group from NBn^6A-RNA-1 (*Uncaging of* NBn^6A-RNA-1). Similarly, in NBn^6A-RNA-3 and NBn^6A-RNA-4, have two photo labile groups and after exposure of light (365 nm) for 10 min ((Entry 3-4, Table 2B.2) have a mass difference of m/z 271 (135.5 x 2), whereas in NBn^6A-RNA-5 have three photo labile groups and after exposure of light (365 nm) for 10 min (Entry 5, Table 2B.2) have a mass difference of m/z 405 (135.06 x 3) before and after exposure of the light. From these mass experiments, the anticipated mass difference of m/z 135/136 after irradiation with 365 nm is observed which indicate the presence of photo labile group and are also completely cleavable with exposure of light (365 nm).

To ensure the chemical cleavage of nitrobenzyl group from NBn^1A-RNA-6/7/8, we irradiated with light (365 nm) for 10 min and then analyzed by ESI-Mass. Their mass spectra are provided in the Appendix, while deconvoluted mass data are given in Table 2B.2 (column 5). In NBn^1A-RNA-6/7 with single modifications which are differed in their position, have no change in their mass before and after exposure at 365 nm for 10 min. For example, NBn^1A-RNA-6 have mass m/z 6746.17, had remained the same after exposure of the light (Entry 6, Table 2B.2). Similarly, in NBn^1A-RNA-8 which have two modifications have no changes in their mass before and after exposure (Entry 7-8, Table 2B.2). Even after prolonged exposure (up to 20 min) there is no change in their

mass spectra. The photo cleavage studies of NBn^{1}A-RNA-6/7/8 indicate the absence of photo labile (2-nitrobenzyl) groups. Thus the *o*-nitrobenzyl groups of NBn^{1}A-RNAs are chemically cleaved during RNA synthesis, unlike NBn^{6}A-RNA. Hence NBn^{6}A-RNAs are chemically stable and caged, but NBn^{1}A-RNAs are not caged with PLPG (*o*-nitrobenzyl group).

Table 2B.2. Mass data of NBn^{6}A-RNAs and NBn^{1}A-RNAs before and after irradiation of the light (365 nm).

Entry	Code	RNA sequence	Observed mass (0 min)	Observed mass (365 nm, 10 min)	Difference in mass
1	*NBn^{6}A*-RNA-1	5'-GCA CCA XCU UCU UCA AGG ATT-3'	6866.21 [$M + 3Na^+$] $^{-5}$	6731.11 [$M + 3Na^+$] $^{-5}$	135.1
2	*NBn^{6}A*-RNA-2	5'-GCA CCA UCU UCU XCA AGG ATT-3'	6888.78 [$M + 4Na^+$] $^{-5}$	6753.70 [$M + 4Na^+$] $^{-5}$	135.08
3	*NBn^{6}A*-RNA-3	5'-GCA CCA XCU XCU UCA AGG ATT-3'	7068.93 [$M + 3Na^+$] $^{-5}$	6797.65 [$M + 3Na^+$] $^{-5}$	271.28 (135.64 x 2)
4	*NBn^{6}A*-RNA-4	5'-GCA CCA XCU XCU XCA AGG ATT-3'	7209.37 [$M + 8H^+$] $^{-5}$	6804.18 [$M + 8H^+$] $^{-5}$	405.19 (135.06 x 3)
5	*NBn^{6}A*-RNA-5	5'-GCA CCA UCX XCU UCA AGG ATT-3'	7067.96 [$M + 3Na^+$] $^{-5}$	6797.86 [$M + 3Na^+$] $^{-5}$	270.1 (135.05 x 2)
6	*NBn^{1}A*-RNA-6	5'-GCA CCA UCU XCU UCA AGG ATT-3'	6746.17 [$M + 2Na^+ + 1K^+$] $^{-5}$	6746.17 [$M + 2Na^+ + 1K^+$] $^{-5}$	0
7	*NBn^{1}A*-RNA-7	5'-GCA CCA UCU UCU XCA AGG ATT-3'	6686.73 [$M + 1Na^+$] $^{-5}$	6686.7 [$M + 1Na^+$] $^{-5}$	0.3
8	*NBn^{1}A*-RNA-8	5'-GCA CCA UCX XCU UCA AGG ATT-3'	6781.25 [$M + 3(NH_4)^+$] $^{-5}$	6781.35 [$M + 3(NH_4)^+$] $^{-5}$	0.1

X = NBn^{6}A/NBn^{1}A adenosine modified nucleosides

Chemical cleavage of NBn^{1}A DNA/RNA.

The acidic conditions (3% TCA in DCM) and basic conditions (NH_4OH solution) are being used for DNA syntheses. Thus we examined the chemical stability NBn^{6}A (**10a**)/NBn^{1}A (**10b**) nucleoside derivatives under similar acidic and basic conditions. Pleasantly, we found that nitrobenzyl group of NBn^{6}A/NBn^{1}A nucleoside derivatives are quite stable under those conditions. However, in both N^1 photo labile modified DNA (*NBn^{1}A DNA)* and RNA (*NBn^{1}A RNA* 6-8), the photo labile (2-nitrobenzyl) group is either partially deprotected (in *NBn^{1}A DNA*) or completely deprotected (in *NBn^{1}A RNA* 6-8) before exposure of light (365 nm). Herein, we propose the chemical cleavage of nitrobenzyl group of NBn^{1}A-DNA/RNA in Figure 2B.12. During DNA/RNA

synthesis, the protonation of imine nitrogen of adenine ring in NBn^{1}A with trichloroacetic acid during DNA/RNA synthesis. In protonated of NBn^{1}A, adenine ring gets positively charged as resultant the nitrobenzyl group is prone to cleavage to retain the aromaticity in the influence of neighbouring nucleobases, which may enhance base stacking interactions.

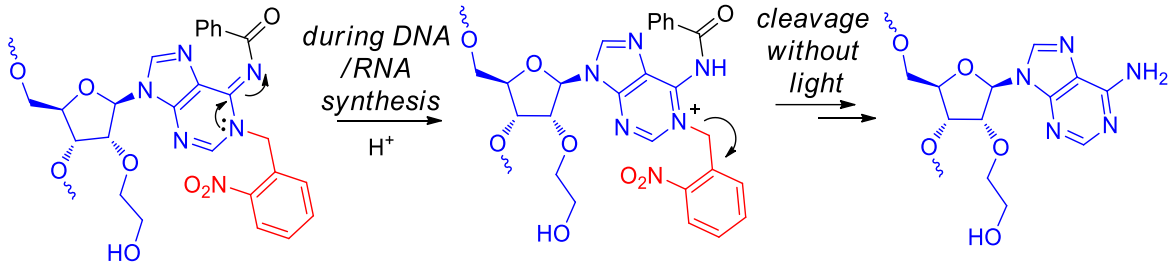


Figure 2B.12. Possible chemical cleavage of PLPG in NBn^{1}A DNA/RNA oligonucleotide.

Isomeraization of NBn^{1}A DNA into NBn^{6}A DNA. Herein we examined the stability of nitrobenzylated NBn^{1}A-DNA, formed partially and found that its nitrobenzyl group was photo cleavable with exposure of the light (365 nm), which is unusual as compared to NBn^{1}A nucleoside derivative. We firmly assume that NBn^{1}A residue has isomerized into NBn^{6}A.Thus we propose the photo cleavage mechanism of nitrobenzyl group in NBn^{1}A-DNA as the pyrimidine ring opening of nitrobenzylated adenine and then re-cyclization with N6-imine in water (Figure 2B.13). In literature, such ring opening of adenine and re-cyclization with N^6-imine are reported in N^1-alkylated adenosine nucleoside [44-46] and also observed in dynamic state in some eukaryotic mRNA.[47] Hence our propose mechanism supports the partial conversion of NBn^{1}A-DNA into *caged*-NBn^{6}A-DNA.

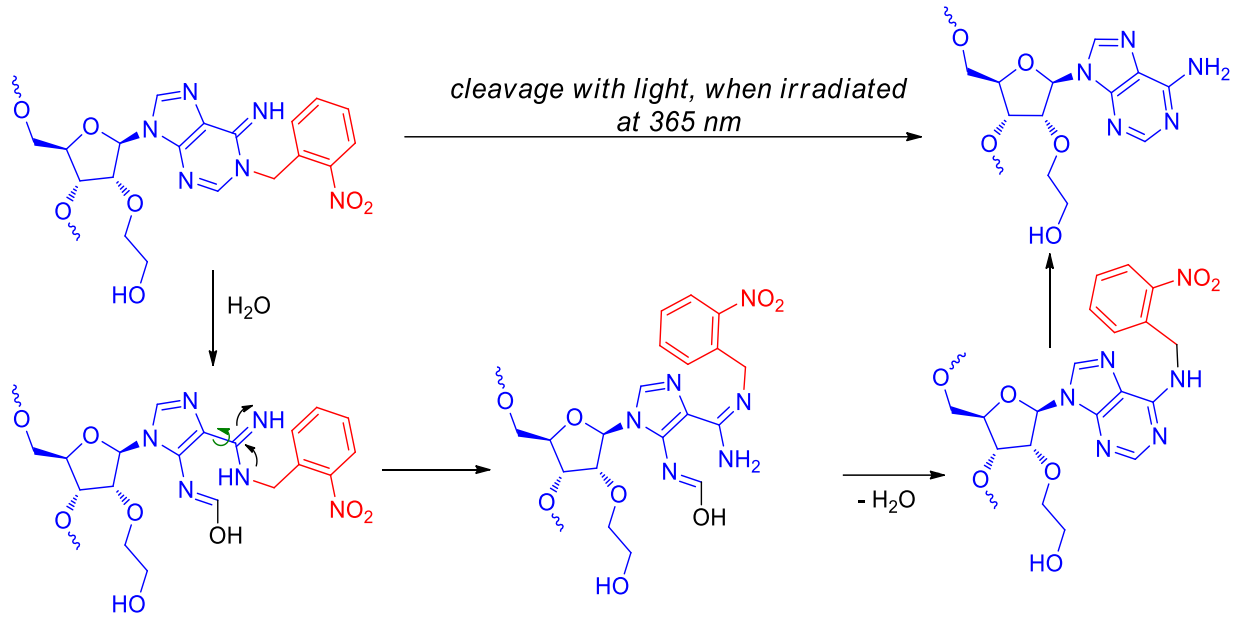


Figure 2B.13. Dimorth rearrangement of NBn[1]A to NBn[6]A and photo cleavage of rearranged NBn[1]A-DNA.

Synthesis of 2'-O-(o-nitrobenzyloxyethyl)adenosine analogue

In development of caged RNAs, we rationally designed adenosine nucleoside comprising 2'-tether capped with photo caged group, A-NBnOE (Figure 2B.14). For the synthesis of designed molecule (A-NBnOE), we initially opted to alkylate 2'-*O*-hydroxyethyl-adenosine derivative with nitrobenzyl bromide (path-a, Figure 2B.14). As discussed in above section, this reaction resulted in alkylation of adenosine nuclobase. After understanding the alkylated products of 2'-*O*-hydroxyethyl-adenosine derivative, we changed the synthetic route to path-b (Figure 2B.14). In literature, direct alkylation of adenosine at its 2'-OH with variety of alkyl halides is well reported.[48]

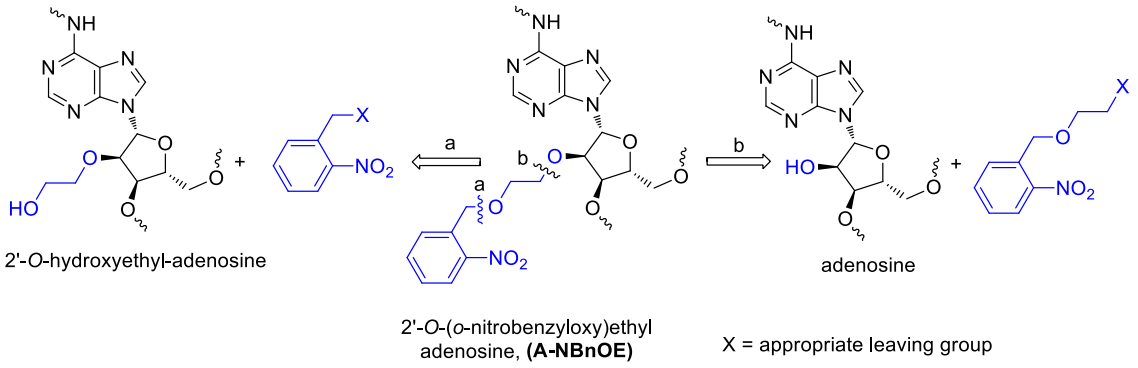
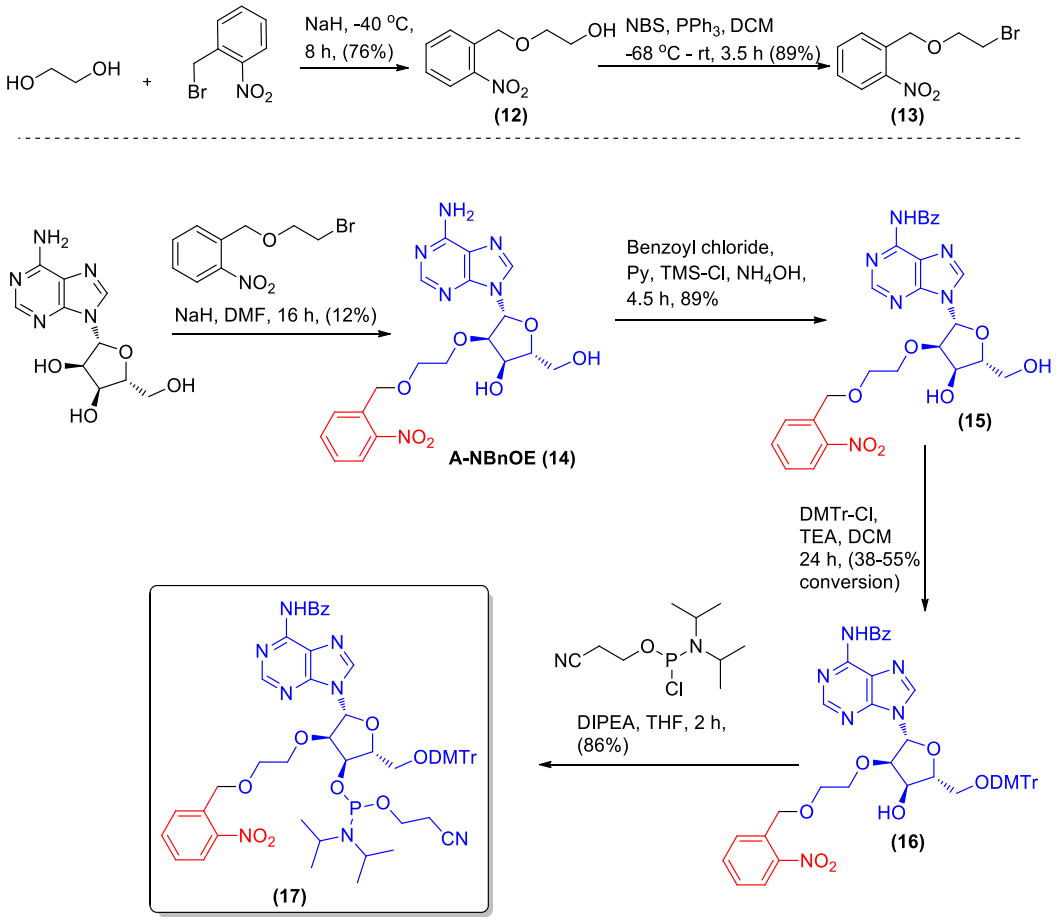


Figure 2B.14. Possible retro synthetic approach for the synthesis of A-NBnOE nucleoside.

Herein, the synthesis of designed caged adenosine phosphoramidite is provided in Scheme 2B.3. In this synthetic approach, first we prepared the alkyl bromide, 2-(*o*-nitrobenzyloxy)ethylbromide (**13**) and then alkylated the adenosine with 2-(*o*-nitrobenzyloxy)ethylbromide. The 2-(*o*-nitrobenzyloxy)ethylbromide (**13**) was prepared in two synthetic steps, starting from the mono alkylation of ethylene glycol with *o*-nitrobenzylbromide (76% yield) by following the reported procedure [49] the resulting 2-(*o*-nitrobenzyloxy)ethanol (**12**) is converted into 2-(*o*-nitrobenzyloxy)ethylbromide (**13**) under Mitsunobu reaction conditions. After the successful synthesis of alkyl halide **13**, the alkylation of adenosine with 2-(*o*-nitrobenzyloxy)ethylbromide (**13**) resulted isolation of 2'-*O*-alkylated adenosine, A-NBnOE (**14**) in 10-12% of isolated yields. For phosphoramidite synthesis, the N^6 amino functional group of adenosine derivative **14** is protected as its N^6-benzolyl adenosine (Bz^6A) derivative **15**, which is obtained by protection of 3'-/5'- hydroxyl groups with transient TMS protecting group followed by benzolylation of N^6 amine with benzoyl chloride. The 5'-OH of adenosine derivative **15** was protected as DMTr derivative **16** and the 3'-OH of DMTr derivative **16** is converted into corresponding phosphoramidite as **17** in 86% of isolated yield. All the compounds are characterized by NMR and ESI-MS and their respective spectra are provided in Appendix.

Scheme 2B.3. Synthesis of 2'-*O*-(*o*-nitrobenzyloxyethyl)adenosine nucleoside phosphoramidite.



To ensure the alkylation at 2'-OH, we analyzed the proton NMR and mass spectral data of A-NBnOE derivatives, **14**-**17**. For example, the proton NMR spectra of adenosine nucleoside **14** exhibit free $-NH_2$ at δ6.6 ppm, this indicate the alkylation did not occur at N^6 position of adenosine. This is further strongly supported from the mass spectra of phosphoramidite **17** (Figure 2B.15). The mass spectra of phosphoramidite **17** exhibit three mass peaks, these mass peaks are analyzed as molecular ion (m/z 1053.417) and fragment ions (m/z 970.306 and 814.341). The fragment ion at m/z 814.341 is analyzed as fragmentation product of glycosidic bond (C'1-N9 bond) in

phosphoramidite **17**. These analyses strongly indicate the presence of alkyl group, 2-(*o*-nitrobenzyloxy)ethyl group, at 2'-*O*- of adenosine in compounds, **14-15**.

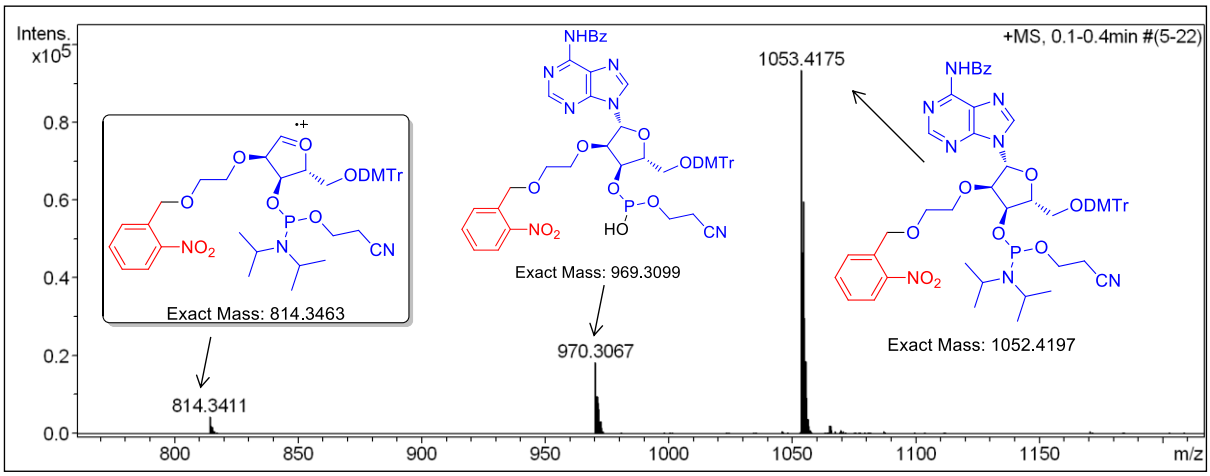


Figure 2B.15. ESI-Mass spectra of adenosine phosphoramidite **17**, molecular ion (m/z 1053), fragment ions (m/z 969 and m/z 814) and their respective structures are also indicated.

*Cleavage of PLPG of 2'-O-(o-nitrobenzyloxyethyl)adenosine analogue **15** with light.* The photochemical cleavage of PLPG, *o*-nitrobenzyl group, of 2'-*O*-(*o*-nitrobenzyloxyethyl)adenosine derivative **15** is monitored by ESI-MS before and after exposure of the light (365 nm) and their mass spectra are provided in Figure 2B.16. Before exposure of the light (0 min), the mass spectra of **15** exhibit molecular ion at m/z 551.194, which corresponds to the molecular ion of **15** (Figure 2B.16A). After 30 min exposure of light (365 nm), the mass spectra of **15** exhibit a new peak at m/z 437.565 ($M+Na^+$) and m/z 415.596 (M^+), these mass peaks belongs to the molecular mass of *uncaged*-adenosine nucleoside **15** (without *o*-nitrobenzyl group) (Figure 2B.16B). These analyses confirm the photo cleavage of nitrobenzyl group of 2'-*O*-(*o*-nitrobenzyloxyethyl)adenosine derivative **15**.

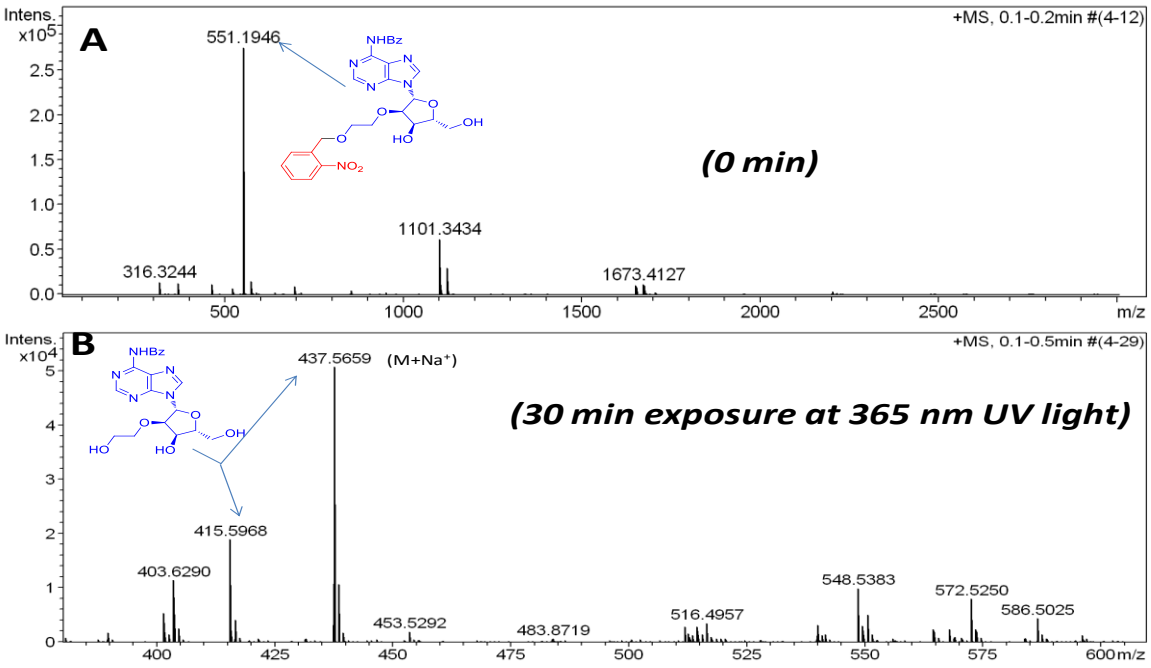


Figure 2B.16. ESI-Mass spectra of **15,** before (A) and after (B) exposure of light at 365 nm.

2'-O-(o-nitrobenzyloxyethyl)adenosine RNA oligonucleotides. Further, we incorporated *2'-O-(o*-nitrobenzyloxyethyl)adenosine phosphoramidite **17** into RNA oligonucleotides (21mer) at DNA synthesizer under standard RNA synthesis conditions, except the coupling times of phosphoramidite **17** is increased to 18 min. Other purification procedures are similar to NBn^6A/NBn^1A RNA oligonucleotides. The phosphoramidite **17** is incorporated in RNA (21mer) at different positions having one incorporation, two incorporations, and four incorporations. The *2'-O-(o*-nitrobenzyloxyethyl)adenosine RNA (A-NBnOE-RNA) sequence and position of modifications are provided in Table 2B.3 (Column 3). Their mass spectra are provided in the Appendix, while their deconvoluted mass values are given in Table 2B.3 (column 4). The mass of corresponding RNA oligonucleotides appear in their mass spectra mostly in charge states z = 5 and

z = 4 with Na^+/NH_4^+ adducts, however mostly we observed the major mass peak of completely desalted RNA. The observed molecular mass of A-NBnOE-RNA-1/2/3 are almost equal to the calculated molecular mass (Table 2B.3, column 4) of *2'-O-(o*-nitrobenzyloxyethyl)adenosine RNA, (A-NBnOE-RNA). For example, the calculated mass of A-NBnOE-RNA-1 is m/z 6769.15 and its observed mass is m/z 6770.24. Similarly, the calculated mass of A-NBnOE-RNA-2 is m/z 6948.33 and its observed mass is m/z 6950.60. In case of A-NBnOE-RNA-3 the calculated mass is m/z 7306.67 and its observed mass appeared as sodium adducts with m/z 7402.80, the observed mass difference (96 mass units) is attributed by four sodium ions. The mass spectral analyses of A-NBnOE-RNA-1/2/3 confirm the syntheses of *2'-O-(o*-nitrobenzyloxyethyl)adenosine RNA, (A-NBnOE-RNA).

Photo cleavage of PLPG in A-NBnOE-RNAs. After the successful synthesis of A-NBnOE-RNA oligonucleotides, we studied the photo cleavage of PLPG, *o*-nitrobenzyl group, by ESI-MS. The deconvoluted mass is provided in Table 2B.3 (Column 5). Herein, the mass of A-NBnOE-RNA-1 (one incorporation of **17**) before exposure to light exhibits observed mass of m/z 6770.24, after exposure to light (365 nm, 10 min)it was reduced to m/z 6635.16 (Entry 1, Table 2B.3). The reduced mass difference (~135.1 unit) is equivalent to the mass of *o*-nitrobenzyl, which confirms the photo cleavage of nitrobenzyl group from A-NBnOE-RNA-1 (*uncaging of* A-NBnOE-RNA-1). In case of A-NBnOE-RNA-2 (two incorporations of **17**), the observed mass is reduced from m/z 6950.60 to m/z 6679.43 after exposure to light (365 nm, 10 min). This mass difference of m/z 271.16 (135.07 x 2) is equivalent to the mass of two nitrobenzyl groups in A-NBnOE-RNA-2, which confirm the photo cleavage of both nitrobenzyl groups in A-NBnOE-RNA-2. In case of A-NBnOE-RNA-3 (four incorporations of **17**), have a mass difference of m/z 552.3 (138.07 x 4) after exposure of light (365 nm, 10 min), which is equivalent to mass of four *o*-nitrobenzyl groups and

confirms the photo cleavage of all four nitrobenzyl group from A-NBnOE-RNA-3. From these mass experiments, the anticipated mass difference of m/z 135/136 after exposure with 365 nm.

Table 2B.3. Mass data of A-NBnOE-RNAs before and after irradiation of the light (365 nm).

Entry	Code	RNA sequence	Observed mass (0 min)	Observed mass (365 nm, 10 min)	Difference in mass
1	*A-NBnOE*-RNA-1	5'-GCA CC**X** UCU UCU UCA AGG ATT-3	6770.24 $[M + 1H^+]^{-5}$	6635.16 $[M + 1H^+]^{-5}$	135.08
2	*A-NBnOE*-RNA-2	5'-GCA CC**X** UCU UCU UC**X** AGG ATT-3'	6950.60 $[M + 2H^+]^{-5}$	6679.43 $[M + 1H^+]^{-5}$	271.16 (135.58 x 2)
3	*A-NBnOE*-RNA-3	5'-GC**X** CC**X** UCU UCU UC**X** **X**GG ATT-3'	7402.80 $[[M + 4Na^+]_5 + 4H^+]^-$	6797.65 $[[M + 1NH_4 + 3Na^+]_5 - 4H^+]^-$	552.3 (138.07 x 4)

X = A-NBnOE-RNA modified nucleoside

2B.4 Conclusion

We have accomplished the nitrobenzylation of adenine residue at N^1/N^6 position of 2'-tethered-N^6-benzoyladenosine nucleosides derivatives (NBn^6A and NBn^1A) and also demonstrated the chemical synthesis of *2'-O-(o*-nitrobenzyloxyethyl)adenosine nucleoside derivative (A-NBnOE). The N^1/N^6 products exhibit characteristic UV and CD signal spectra, which can be used to distinguish N^1 and N^6 products just by UV and CD spectrophotometers. We have also achieved the 2'-thethered *caged*-adenosine and *2'-O-(o*-nitrobenzyloxyethyl)adenosine phosphoramidites and incorporated into DNA/RNA oligonucleotides. Pleasantly, we demonstrated that the nitrobenzyl groups are photo cleavable from NBn^6A and A-NBnOE containing nucleosides and oligonucleotides, but not cleavable from NBn^1A nucleosides. The nitrobenzyl group of NBn^1A-DNA has partially retained, which is photo cleavable. The partial retention of nitrobenzyl group in DNA, also photo labile, is presumably due to the isomerization into NBn^6A DNA. However, the nitrobenzyl group of NBn^1A containing RNA has not retained because of its chemical cleavage.

There is no report about photo labile protection at N^1 of adenosine and *2'-O-(o*-nitrobenzyloxyethyl)adenosine, though a very few literature reports are available about N^6-photolabile protected adenosine analogues. This study reveals the unusual chemical synthesis of N^1 and N^6 photo labile analogues of 2'-*O*-hydroxyethyl-N^6Bz-adenosine derivative and unusual photochemical behaviour of NBn^1A in nucleoside and their oligonucleotide analogues. The NBn^6A and NBn^1A nucleosides and oligonucleotides with photo labile groups and 2'-*O*-tethered have a potential impact with altered base pairing properties on biochemical pathways and exert rational photo-controlled activities. While, the presence of bulkier photo cleavable group in A-NBnOE-RNA oligonucleotides can exhibit the spatial interactions without interfering Watson-Crick base pairing with enzymes and these interactions can controll by light. Hence 2'-tethered NBn^6A, NBn^1A, A-NBnOE nucleosides and their DNA/RNA oligonucleotides are potential candidates in the development of *caged*-ASO analogues.

2B.5 Experimental section

Materials. All the required chemicals were obtained from commercial suppliers and used without any further purification. Pyridine and methylene chloride solvents were distilled over calcium hydride and stored under 4Å molecular sieves prior using for reactions. Tetrahydrofuran was freshly distilled over sodium and benzophenone for reactions. All the reactions were monitored by thin layer chromatography, visualized by UV and Ninhydrin. Column chromatography was performed with 230-400 mesh silica. MS and HRMS were obtained from Bruker micrOTOF-Q II Spectrometer. $^1H/^{13}C/^{31}P$ NMR spectra were recorded on Bruker AV-400 or 700 MHz at 298 K. ^1H and ^{13}C NMR chemical shifts were recorded in ppm downfield from tetramethyl silane or

residual solvent peak. Splitting patterns are abbreviated as: s, singlet; d, doublet; dd, doublet of doublet; t, triplet; q, quartet; dq, doublet of quartet; m, multiplet.

Experimental procedures

Experimental procedures for compound **2-5** are provided in literature reports

methyl- 2-(((2R,3R,4R,5R)-2-(6-amino-9H-purin-9-yl)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-3-yl)oxy)acetate (**2).** ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.30 (s, 1H), 7.85 (s, 1H), 6.78 (d, *J* = 11.2 Hz, 1H), 6.04 (d, *J* = 32.4 Hz, 2H), 5.94 (d, *J* = 7.7 Hz, 1H), 4.75 (dd, *J* = 7.7, 4.1 Hz, 1H), 4.53 – 4.20 (m, 4H), 3.94 (dd, *J* = 15.1, 4.9 Hz, 2H), 3.85 – 3.57 (m, 4H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 172.04 (s), 156.32 (s), 152.42 (s), 148.43 (s), 141.04 (s), 120.99 (s), 89.02 (s), 87.77 (s), 84.40 (s), 70.83 (s), 68.95 (s), 63.17 (s), 52.44 (s). ESI-HRMS $[M+H]^+$ calcd for $C_{13}H_{17}N_5O_6$ 340.1252, found 340.1259.

methyl 2-(((2R,3R,4R,5R)-2-(6-amino-9H-purin-9-yl)-4-((tert-butyldimethylsilyl)oxy)-5-(((tert-butyldimethylsilyl)oxy)methyl)tetrahydrofuran-3-yl)oxy)acetate (**3**). ^{1}H NMR (400 MHz, DMSO) δ 8.30 (s, 1H), 8.12 (s, 1H), 7.30 (s, 2H), 6.07 (d, *J* = 5.4 Hz, 1H), 4.77 (t, *J* = 5.0 Hz, 1H), 4.65 (t, *J* = 4.1 Hz, 1H), 4.21 (s, 2H), 4.00 – 3.92 (m, 1H), 3.85 (dd, *J* = 11.4, 4.9 Hz, 1H), 3.69 (dd, *J* = 11.4, 3.9 Hz, 1H), 3.51 (s, 3H), 0.87 (d, *J* = 30.3 Hz, 18H), 0.14 (d, *J* = 1.8 Hz, 6H), 0.09 – -0.05 (m, 6H). ESI-HRMS $[M+H]^+$ calcd for $C_{25}H_{45}N_5O_6Si_2$ 568.2981, found 568.2933.

2-(((2R,3R,4R,5R)-2-(6-amino-9H-purin-9-yl)-4-((tert-butyldimethylsilyl)oxy)-5-(((tert-butyldimethylsilyl)oxy)methyl)tetrahydrofuran-3-yl)oxy)ethanol (**4**). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.32 (s, 1H), 8.20 (s, 1H), 6.18 (d, *J* = 3.1 Hz, 1H), 5.84 (s, 2H), 4.50 (t, *J* = 5.2 Hz, 1H), 4.32 (t, *J* = 3.9 Hz, 1H), 4.18 – 4.08 (m, 1H), 4.01 (dd, *J* = 11.6, 2.9 Hz, 1H), 3.93 – 3.65 (m, 5H), 0.93 (d, *J* = 2.6 Hz, 18H), 0.11 (s, 12H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 155.45 (s), 153.03 (s), 149.24 (s),

139.03 (s), 120.07 (s), 87.67 (s), 84.69 (s), 83.01 (s), 72.65 (s), 69.71 (s), 61.91 (s), 61.56 (s), 26.03 (s), 25.72 (s), 18.47 (s), 18.09 (s), -4.57 (s), -4.81 (s), -5.38 (s), -5.40 (s). ESI-HRMS $[M+H]^+$ calcd for $C_{24}H_{45}N_5O_5Si_2$ 540.3032, found 540.3017.

N-(9-((2R,3R,4R,5R)-4-((tert-butyldimethylsilyl)oxy)-5-(((tert-butyldimethylsilyl)oxy)methyl)-3-(2-hydroxyethoxy)tetrahydrofuran-2-yl)-9H-purin-6-yl)benzamide (**5**). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.47 (s, 1H), 8.74 (s, 1H), 8.38 (s, 1H), 8.02 (d, *J* = 7.3 Hz, 2H), 7.56 (t, *J* = 7.4 Hz, 1H), 7.48 (t, *J* = 7.5 Hz, 2H), 6.22 (d, *J* = 3.6 Hz, 1H), 4.50 (t, *J* = 5.1 Hz, 1H), 4.42 – 4.28 (m, 1H), 4.13 (dt, *J* = 5.6, 2.9 Hz, 1H), 3.99 (dd, *J* = 11.6, 3.3 Hz, 1H), 3.84 – 3.64 (m, 5H), 1.02 – 0.82 (m, 18H), 0.09 (d, *J* = 5.1 Hz, 12H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 165.01 (s), 152.65 (s), 151.31 (s), 149.77 (s), 141.72 (s), 133.80 (s), 132.77 (s), 128.82 (s), 128.12 (s), 123.49 (s), 87.62 (s), 85.13 (s), 82.85 (s), 72.50 (s), 69.85 (s), 61.86 (s), 61.76 (s), 29.76 (s), 26.11 (s), 25.79 (s), 18.55 (s), 18.16 (s), -4.49 (s), -4.72 (s), -5.28 (s), -5.31 (s). ESI-HRMS $[M+H]^+$ calcd for $C_{31}H_{50}N_5O_6Si_2$ 644.3300, found 644.3326.

To the solution of Compound **5** in dry AcCN was added anhydrous K_2CO_3 (1.43 g, 10.3 mmol) and stirred for 45 min. Then 2-Nitrobenzyl bromide (0.393 ml, 4.15 mmol) was added in one portion and stirred at 55 ^{o}C for 6 h. The crude reaction mixture was filtered and concentrated to dryness under vacuum. The residual yellow oil was purified via chromatography with 60% EtOAc/ n-Hexane to yield 87% of approximately 1:1 regioisomers, **6a** and **6b**. Isolated yield are 0.7 g of **6a** and 0.65 g of **6b** after multiple column purifications. R_f 0.6 and 0.37 for **6a** and **6b** respectively in 60% EtOAc/ n-Hexane.

N-(9-((2R,3R,4R,5R)-4-((tert-butyldimethylsilyl)oxy)-5-(((tert-butyldimethylsilyl)oxy)methyl)-3-(2-hydroxyethoxy)tetrahydrofuran-2-yl)-9H-purin-6-yl)-N-(2-nitrobenzyl)benzamide **(6a).** ^{1}H

NMR (400 MHz, $CDCl_3$) δ 8.54 (s, 1H), 8.25 (s, 1H), 8.01 (d, *J* = 8.1 Hz, 1H), 7.83 (d, *J* = 7.8 Hz, 1H), 7.51 (dd, *J* = 17.1, 7.7 Hz, 3H), 7.39 – 7.27 (m, 2H), 7.17 (t, *J* = 7.6 Hz, 2H), 6.11 (d, *J* = 3.7 Hz, 1H), 6.02 (s, 2H), 4.44 (t, *J* = 5.0 Hz, 1H), 4.21 (t, *J* = 4.2 Hz, 1H), 4.15 – 4.07 (m, 1H), 3.93 (d, *J* = 3.0 Hz, 1H), 3.80 – 3.56 (m, 5H), 2.27 (s, 1H), 0.90 (d, *J* = 5.9 Hz, 18H), 0.13 – 0.03 (m, 12H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 172.34, 153.81, 152.28, 152.08, 148.38, 142.44, 135.75, 133.99, 133.84, 131.11, 128.94, 128.05, 127.84, 126.91, 125.05, 87.39, 85.24, 83.06, 72.42, 69.92, 62.04, 61.76, 49.26, 26.13, 25.83, 18.57, 18.22, -4.48, -4.68, -5.27, -5.31. HRMS (ESI-TOF) m/z: $[M + Na]^+$ Calcd for $C_{38}H_{54}N_6O_8Si_2Na$, 801.3434; Found 801.3426.

N-(9-((2R,3R,4R,5R)-4-((tert-butyldimethylsilyl)oxy)-5-(((tert-butyldimethylsilyl)oxy)methyl)-3-(2-hydroxyethoxy)tetrahydrofuran-2-yl)-1-(2-nitrobenzyl)-1H-purin-6(9H)-ylidene)benzamide **(6b).** ^1H NMR (400 MHz, $CDCl_3$) δ 8.16 (d, *J* = 7.7 Hz, 1H), 8.06 (s, 1H), 7.95 (s, 1H), 7.83 (d, *J* = 7.3 Hz, 2H), 7.62 (d, *J* = 7.5 Hz, 1H), 7.54 (d, *J* = 7.5 Hz, 1H), 7.41 (d, *J* = 7.4 Hz, 1H), 7.35 (d, *J* = 7.7 Hz, 1H), 7.29 (t, *J* = 7.6 Hz, 2H), 6.06 (d, *J* = 4.7 Hz, 1H), 5.76 (d, *J* = 2.3 Hz, 2H), 4.43 (t, *J* = 4.6 Hz, 1H), 4.22 (t, *J* = 4.7 Hz, 1H), 4.10 (d, *J* = 3.9 Hz, 1H), 3.85 (d, *J* = 3.6 Hz, 1H), 3.81 – 3.66 (m, 4H), 3.61 (d, *J* = 6.5 Hz, 1H), 2.27 (s, 1H), 0.92 (s, 9H), 0.88 (s, 9H), 0.11 (d, *J* = 2.0 Hz, 6H), 0.05 (d, *J* = 5.9 Hz, 6H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 176.88, 148.14, 147.39, 146.75, 144.99, 138.45, 135.59, 134.21, 132.08, 131.36, 129.76, 129.38, 129.12, 128.74, 128.05, 127.50, 125.69, 122.76, 86.92, 85.73, 83.12, 72.31, 70.32, 62.31, 61.94, 49.88, 26.13, 25.84, 18.57, 18.23, -4.49, -4.62, -5.18, -5.33. HRMS (ESI-TOF) m/z: $[M + H]^+$ Calcd for $C_{38}H_{55}N_6O_8Si_2$, 779.3614; Found 779.3620.

2-(((2R,3R,4R,5R)-4-((tert-butyldimethylsilyl)oxy)-5-(((tert-butyldimethylsilyl)oxy)methyl)-2-(6-(N-(2-nitrobenzyl)benzamido)-9H-purin-9-yl)tetrahydrofuran-3-yl)oxy)ethyl methanesulfonate (**7a**). Compound **6a** (0.33 g, 0.43 mmol) was dissolved in freshly distilled dry

DCM (15 ml) and cooled to 0 ^{o}C, then triethylamine (0.3 ml, 2.1 mmol) followed by methanesulfonyl chloride (0.07 ml, 0.86 mmol) was added and stirred for 1 h at ambient temperature. The crude reaction mixture was evaporated under reduced pressure. The resultant residual oil was purified by column chromatography. The silica gel was prewashed with 0.5% triethylamine/1.5% MeOH/CH_2Cl_2 and eluted with 2% MeOH/CH_2Cl_2 to yield 0.31 g (85%) of compound **7a** as white solid. R_f 0.7 (3% MeOH/CH_2Cl_2). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.52 (s, 1H), 8.25 (s, 1H), 8.01 (dd, *J* = 8.2, 1.0 Hz, 1H), 7.83 (d, *J* = 7.8 Hz, 1H), 7.60 – 7.45 (m, 3H), 7.39 – 7.27 (m, 2H), 7.18 (t, *J* = 7.7 Hz, 2H), 6.08 (d, *J* = 3.5 Hz, 1H), 6.02 (s, 2H), 4.54 – 4.45 (m, 1H), 4.38 – 4.27 (m, 2H), 4.24 – 4.17 (m, 1H), 4.12 – 4.01 (m, 1H), 3.95 (dd, *J* = 11.7, 3.1 Hz, 1H), 3.90 – 3.72 (m, 3H), 2.96 (s, 3H), 0.90 (d, *J* = 4.5 Hz, 180H), 0.14 – 0.02 (m, 12H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 172.22, 153.69, 152.15, 151.96, 148.29, 142.32, 135.61, 133.84, 133.7, 131.03, 128.84, 128.81, 127.95, 127.74, 126.71, 124.91, 86.83, 84.81, 82.84, 69.70, 68.64, 68.46, 61.50, 49.10, 37.59, 26.01, 25.70, 18.46, 18.08, -4.60, -4.85, -5.40, -5.42. HRMS (ESI-TOF) m/z: $[M + H]^+$ Calcd for $C_{39}H_{57}N_6O_{10}Si_2S$, 857.3390; Found 857.3377.

2-(((2R,3R,4R,5R)-2-(6-(benzoylimino)-1-(2-nitrobenzyl)-1H-purin-9(6H)-yl)-4-((tert-butyldimethylsilyl)oxy)-5-(((tert-butyldimethylsilyl)oxy)methyl)tetrahydrofuran-3-yl)oxy)ethyl methanesulfonate (**7b**). Under similar conditions compound **6b** (0.62 g, 0.77 mmol) yielded 0.66 g (97%) of compound **7b** as white solid. R_f 0.37 (2% MeOH/CH_2Cl_2). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.16 (d, *J* = 8.2 Hz, 1H), 8.06 (s, 1H), 7.92 (s, 1H), 7.82 (d, *J* = 7.6 Hz, 2H), 7.64 (s, 1H), 7.53 (s, 1H), 7.47 – 7.35 (m, 2H), 7.29 (t, *J* = 7.1 Hz, 2H), 6.03 (d, *J* = 5.0 Hz, 1H), 5.76 (s, 2H), 4.43 (t, *J* = 4.3 Hz, 1H), 4.35 – 4.22 (m, 3H), 4.07 (d, *J* = 3.5 Hz, 1H), 3.98 – 3.84 (m, 2H), 3.74 (dd, *J* = 9.4, 4.6 Hz, 2H), 2.96 (s, 3H), 0.90 (d, *J* = 15.9 Hz, 18H), 0.11 (s, 6H), 0.06 (d, *J* = 3.4 Hz, 6H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 176.88, 148.14, 147.50, 146.63, 145.08, 138.50, 135.58, 134.24,

132.08, 131.36, 129.76, 129.49, 129.11, 128.06, 125.66, 122.75, 86.49, 85.86, 83.13, 70.36, 68.76, 68.44, 62.37, 49.94, 37.65, 26.13, 25.84, 18.58, 18.24, -4.48, -4.67, -5.17, -5.32. HRMS (ESI-TOF) m/z: $[M + H]^+$ Calcd for $C_{39}H_{57}N_6O_{10}Si_2S$, 857.3390; Found 857.3375.

2-(((2R,3R,4R,5R)-4-((tert-butyldimethylsilyl)oxy)-5-(((tert-butyldimethylsilyl)oxy)methyl)-2-(6-(N-(2-nitrobenzyl)benzamido)-9H-purin-9-yl)tetrahydrofuran-3-yl)oxy)ethyl benzoate (**8a**). Compound **7a** (0.3 g, 0.35 mmol) was dissolved and stirred in dry DMF (5 ml) at room temperature. Sodium benzoate (0.12 g, 0.77 mmol) was added to the resulting solution and heated to 90 ^oC. After stirring for 4 h, the reaction mixture was dried under vacuum. The residual viscous oil was purified via chromatography with 50% EtOAc/ n-Hexane to yield 0.31 g (100%) of compound **8a** as a white solid. R_f 0.4 (50% EtOAc/ n-Hexane). ^1H NMR (400 MHz, $CDCl_3$) δ 8.53 (s, 1H), 8.27 (s, 1H), 8.00 (dd, *J* = 15.1, 7.8 Hz, 3H), 7.86 (d, *J* = 7.7 Hz, 1H), 7.60 – 7.46 (m, 4H), 7.41 (t, *J* = 7.2 Hz, 2H), 7.35 (s, 1H), 7.28 (s, 1H), 7.17 (d, *J* = 7.2 Hz, 2H), 6.14 (s, 1H), 6.03 (s, 2H), 4.49 (s, 1H), 4.43 (s, 2H), 4.31 (s, 1H), 4.13 (s, 1H), 4.04 – 3.88 (m, 2H), 3.88 – 3.70 (m, 2H), 0.91 (s, 18H), 0.08 (d, *J* = 11.9 Hz, 12H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 172.25, 166.45, 153.74, 152.31, 152.10, 148.36, 142.59, 135.73, 133.96, 133.77, 133.24, 131.05, 129.92, 129.72, 128.89, 128.45, 127.97, 127.80, 126.87, 124.98, 87.09, 85.04, 82.59, 70.09, 68.91, 63.90, 61.76, 49.20, 26.10, 25.95, 18.53, 18.19, -4.50, -4.79, -5.30 , -5.33. HRMS (ESI-TOF) m/z: $[M + Na]^+$ Calcd for $C_{45}H_{58}N_6O_9Si_2Na$, 905.3696; Found 905.3508.

2-(((2R,3R,4R,5R)-2-(6-(benzoylimino)-1-(2-nitrobenzyl)-1H-purin-9(6H)-yl)-4-((tert-butyldimethylsilyl)oxy)-5-(((tert-butyldimethylsilyl)oxy)methyl)tetrahydrofuran-3-yl)oxy)ethyl benzoate (**8b**). Under similar conditions compound **7b** (0.64 g, 0.75 mmol) yielded 0.58 g (88%) of compound **8b** as white solid. R_f 0.25 (40% EtOAc/ n-Hexane). ^1H NMR (400 MHz, $CDCl_3$) δ 8.16 (d, *J* = 8.1 Hz, 1H), 7.93 (dd, *J* = 6.3, 2.1 Hz, 3H), 7.90 (s, 1H), 7.83 (d, *J* = 7.3 Hz, 2H), 7.59

(s, 1H), 7.52 (dd, *J* = 7.4, 3.7 Hz, 2H), 7.40 (dd, *J* = 14.7, 7.0 Hz, 3H), 7.29 (dd, *J* = 10.4, 4.7 Hz, 3H), 6.07 (d, *J* = 4.7 Hz, 1H), 5.69 (q, *J* = 16.7 Hz, 2H), 4.44 (q, *J* = 4.7 Hz, 3H), 4.32 (t, *J* = 4.7 Hz , 1H), 4.10 (d, *J* = 4.0 Hz, 1H), 3.86 (dd, *J* = 18.1, 14.2 Hz, 3H), 3.75 (d, *J* = 3.2 Hz, 1H), 0.89 (d, *J* = 9.2 Hz, 18H), 0.10 (s, 6H), 0.05 (d, *J* = 5.3 Hz, 6H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 176.77, 166.52, 148.09, 147.22, 146.75, 145.05, 138.67, 135.67, 134.20, 133.26, 132.04, 131.42, 129.95, 129.77, 129.71, 129.24, 129.05, 128.53, 128.04, 125.66, 122.79, 86.84, 85.58, 82.67, 70.46, 68.94, 63.99, 62.31, 49.82, 26.14, 25.83, 18.59, 18.24, -4.50, -4.69, -5.17, -5.31. HRMS (ESI-TOF) m/z: $[M + Na]^+$ Calcd for $C_{45}H_{59}N_6O_9Si_2$, 883.3837; Found 883.3875.

2-(((2R,3R,4R,5R)-4-hydroxy-5-(hydroxymethyl)-2-(6-(N-(2-nitrobenzyl)benzamido)-9H-purin-9-yl)tetrahydrofuran-3-yl)oxy)ethyl benzoate (**9a**). Compound **8a** (0.31 g, 0.35 mmol) was dissolved in freshly distilled dry THF (5 ml) and cooled to 0 ^{o}C. Tetrabutylammonium fluoride 1M solution (0.78 ml, 0.78 mmol) was added drop wise and allowed to stir for 25 min at the same temperature. The solvents were removed under vacuum and the resulting pale orange residue was partitioned between EtOAc (25 mL) and saturated $NaHCO_3$ (10 mL). The organic layer separated and the aqueous layer was extracted again with EtOAc (2 × 25 mL). The combined organic extracts were washed with saturated aqueous NaCl (10 mL), dried over anhydrous Na_2SO_4, and concentrated under vacuum. The residual viscous oil was purified via chromatography with 3% MeOH/CH_2Cl_2 to produce 0.22 g (96%) of compound **9a** as a white solid. R_f 0.3 (3% MeOH/CH_2Cl_2). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.48 (s, 1H), 8.01 (dd, *J* = 8.2, 1.1 Hz, 1H), 7.99 – 7.94 (m, 2H), 7.89 (s, 1H), 7.78 (d, *J* = 7.4 Hz, 1H), 7.59 (d, *J* = 7.4 Hz, 1H), 7.54 – 7.43 (m, 5H), 7.39 – 7.29 (m, 2H), 7.20 (t, *J* = 7.7 Hz, 2H), 6.00 (s, 2H), 5.85 (d, *J* = 7.3 Hz, 1H), 5.68 (dd, *J* = 11.5, 1.9 Hz, 1H), 4.78 (dd, *J* = 7.3, 4.6 Hz, 1H), 4.55 (d, *J* = 4.5 Hz, 1H), 4.34 (ddd, *J* = 24.5, 12.8, 7.9 Hz, 3H), 3.92 (d, *J* = 12.9 Hz, 1H), 3.81 – 3.67 (m, 3H), 2.95 (s, 1H), 1.66 (s, 1H). ^{13}C

NMR (101 MHz, $CDCl_3$) δ 172.26, 166.30, 154.64, 151.43, 151.33, 148.35, 143.64, 135.35, 133.69, 133.66, 133.50, 131.33, 129.61, 129.35, 128.83, 128.71 , 128.62 , 128.10, 127.87, 127.69, 124.95, 89.43, 87.93, 81.50, 70.78, 69.38, 63.54, 63.08, 49.13. HRMS (ESI-TOF) m/z: $[M + Na]^+$ Calcd for $C_{33}H_{30}N_6O_9Na$, 677.1966; Found 677.1986.

2-(((2R,3R,4R,5R)-2-(6-(benzoylimino)-1-(2-nitrobenzyl)-1H-purin-9(6H)-yl)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-3-yl)oxy)ethyl benzoate (**9b**). Under similar conditions compound **8b** (0.556 g, 0.63 mmol) yielded 0.39 g (96%) of compound **9b** as white solid. R_f 0.25 (4% MeOH/ CH_2Cl_2). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.17 (d, *J* = 8.1 Hz, 1H), 8.02 (s, 1H), 7.96 (d, *J* = 7.3 Hz, 2H), 7.87 – 7.75 (m, 3H), 7.65 (s, 1H), 7.54 (t, *J* = 7.8 Hz, 2H), 7.41 (dt, *J* = 14.4, 7.3 Hz, 4H), 7.31 (t, *J* = 7.7 Hz, 2H), 5.85 (d, *J* = 6.9 Hz, 1H), 5.71 (dd, *J* = 38.4, 16.5 Hz, 2H), 4.66 (dd, *J* = 6.8, 4.8 Hz, 1H), 4.52 (d, *J* = 3.9 Hz, 1H), 4.43 (dd, *J* = 5.0, 3.0 Hz, 1H), 4.39 (dd, *J* = 6.7, 2.9 Hz, 1H), 4.30 (s, 1H), 3.94 – 3.79 (m, 3H), 3.72 (d, *J* = 12.4 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 176.92, 166.48, 148.19, 147.48, 145.97, 143.86, 140.34, 135.13, 134.29, 133.51, 132.38, 130.80, 129.73, 129.69, 129.52, 129.37, 128.69, 128.20, 125.78, 124.01, 89.13, 87.56, 82.36, 70.58, 69.48, 63.70, 62.97, 50.04. HRMS (ESI-TOF) m/z: $[M + Na]^+$ Calcd for $C_{33}H_{30}N_6O_9Na$, 677.1966; Found 677.1963.

2-(((2R,3R,4R,5R)-5-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-4-hydroxy-2-(6-(N-(2-nitrobenzyl)benzamido)-9H-purin-9-yl)tetrahydrofuran-3-yl)oxy)ethyl benzoate (**10a**). Compound **9a** (0.2 g, 0.31 mmol) was co-evaporated with dry pyridine for 3 times and then dissolved in dry CH_2Cl_2 (4 ml). Dry triethylamine (0.22 ml, 1.56 mmol) followed by DMTr-Cl (0.17 gm, 0.5 mmol) was added to the resulting solution and stirred at room temperature for 24 h. The crude reaction mixture concentrated to dryness and then subjected to column chromatography (0.5% triethylamine/2.5% MeOH/CH_2Cl_2) to yield 0.25 g (84%) of compound **10a** as white solid.

R_f 0.3 (2.5% MeOH/ CH_2Cl_2). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.48 (s, 1H), 8.10 (s, 1H), 8.03 (dd, *J* = 8.2, 1.0 Hz, 1H), 8.01 – 7.96 (m, 2H), 7.82 (d, *J* = 7.8 Hz, 1H), 7.57 (d, *J* = 7.4 Hz, 1H), 7.54 – 7.47 (m, 3H), 7.46 – 7.38 (m, 4H), 7.38 – 7.22 (m, 11H), 7.16 (t, *J* = 7.7 Hz, 2H), 6.81 (d, *J* = 8.9 Hz, 4H), 6.13 (d, *J* = 3.4 Hz, 1H), 6.02 (s, 2H), 4.61 – 4.38 (m, 4H), 4.20 (dd, *J* = 8.8, 4.1 Hz, 1H), 4.17 – 4.06 (m, 1H), 4.02 – 3.92 (m, 1H), 3.80 (d, *J* = 8.1 Hz, 6H), 3.50 (dd, *J* = 10.7, 2.9 Hz, 1H), 3.38 (dd, *J* = 10.7, 4.4 Hz, 1H), 2.77 (s, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 172.18, 166.45, 158.6, 153.76, 152.16, 152.01, 149.86, 148.21, 144.45, 142.18, 135.56, 135.52, 135.48, 133.85, 133.72, 133.37, 131.13, 130.07, 129.59, 129.54, 128.77, 128.71, 128.55, 128.09, 128.01, 127.93, 127.74, 126.99 , 126.74, 124.95, 123.76, 113.23, 87.14, 86.66, 83.82, 82.13, 69.76, 69.56, 63.41, 62.80, 55.27, 49.25. HRMS (ESI-TOF) m/z: $[M + H]^+$ Calcd for $C_{54}H_{49}N_6O_{11}$, 957.3454; Found 957.3500.

2-(((2R,3R,4R,5R)-2-(6-(benzoylimino)-1-(2-nitrobenzyl)-1H-purin-9(6H)-yl)-5-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-4-hydroxytetrahydrofuran-3-yl)oxy)ethyl benzoate (**10b**). Under similar conditions compound **9b** (0.335 g, 0.51 mmol) yielded 0.34 g (70%) of compound **10b** as white solid. R_f 0.45 (5% MeOH/ CH_2Cl_2). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.15 (dd, *J* = 7.9, 1.5 Hz, 1H), 8.01 – 7.94 (m, 2H), 7.89 – 7.80 (m, 4H), 7.60 – 7.48 (m, 3H), 7.47 – 7.36 (m, 5H), 7.34 – 7.14 (m, 11H), 6.77 (d, *J* = 8.9 Hz, 4H), 6.04 (d, *J* = 4.3 Hz, 1H), 5.67 (s, 2H), 4.63 – 4.35 (m, 4H), 4.20 (d, *J* = 3.5 Hz, 1H), 4.09 – 3.91 (m, 2H), 3.74 (s, 6H), 3.49 – 3.27 (m, 2H), 2.77 (d, *J* = 5.8 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 176.78, 166.57, 158.69, 148.12, 147.30, 146.64, 144.89, 144.69, 138.79, 135.71, 135.68, 135.50, 134.18, 133.49, 132.14, 131.27, 130.22, 130.20, 129.77, 129.69, 129.34, 129.07, 128.70, 128.21, 128.09, 128.00, 127.03, 125.65, 123.00, 113.31, 87.08, 86.65, 84.16, 82.36, 70.00, 69.69, 63.62, 63.29, 60.54, 55.36, 49.82, 14.33. HRMS (ESI-TOF) m/z: $[M + Na]^+$ Calcd for $C_{54}H_{48}N_6O_{11}Na$, 979.3273; Found 979.3256.

2-(((2R,3R,4R,5R)-5-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-4-(((2-cyanoethoxy)(diisopropylamino)phosphino)oxy)-2-(6-(N-(2-nitrobenzyl)benzamido)-9H-purin-9-yl)tetrahydrofuran-3-yl)oxy)ethyl benzoate (**11a**). 2-Cyanoethyldiisopropylchlorophosphoramidite (0.11 ml, 0.5 mmol) was added at room temperature to a stirred solution of compound **10a** (0.23 g, 0.24 mmol) in freshly distilled dry THF (2 ml) and N,N-Diisopropyl ethylamine (0.21 ml, 1.2 mmol). After stirred for 4.5 h, the crude reaction mixture was directly subjected to column chromatography (silicagel was prewashed with 1% triethylamine/70% EtOAc/ n-Hexane and eluted with 70% EtOAc/ n-Hexane) to yield 0.21 gm (74%) of compound **11a** as white foam. R_f 0.2 (70% EtOAc/ n-Hexane). ^{31}P NMR (162 MHz, $CDCl_3$) δ 150.22 (s), 150.00 (s). HRMS (ESI-TOF) m/z: $[M + H]^+$ Calcd for $C_{63}H_{66}N_8O_{12}P$, 1157.4532; Found 1157.4558.

2-(((2R,3R,4R,5R)-2-(6-(benzoylimino)-1-(2-nitrobenzyl)-1H-purin-9(6H)-yl)-5-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-4-(((2-cyanoethoxy)(diisopropylamino)phosphino)oxy)tetrahydrofuran-3-yl)oxy)ethyl benzoate (**11b**). Under similar conditions compound **10b** (0.32 g, 0.23 mmol) yielded 0.39 g (70%) of compound **11b** as white foam. R_f 0.22 (75% EtOAc/ n-Hexane). ^{31}P NMR (162 MHz, $CDCl_3$) δ 150.53 (s), 150.09 (s). HRMS (ESI-TOF) m/z: $[M + Na]^+$ Calcd for $C_{63}H_{65}N_8O_{12}PNa$, 1179.4352; Found 1179.4306.

1-((2-bromoethoxy)methyl)-2-nitrobenzene (**13**). N-bromosuccinamide (0.75 g, 4.2 mmol) is dissolved in CH_2Cl_2 (9 ml) and the resulting solution is cooled to -68 ^oC . Then triphenylphosphine (1.11 g, 4.2 mmol) dissolved in CH_2Cl_2 (6 ml) and added drop wise to the reaction mixture, the resulting solution turned into clear pale yellow color. Subsequently, Compound **12** (0.6 g, 3.04 mmol) is dissolved in CH_2Cl_2 (30 ml) and added to the reaction mixture drop wise, after complete

addition, the temperature is warmed to 0 ^{o}C. The reaction is monitored by TLC, after complete consumption of starting material (3.5 h); the crude reaction mixture is concentrated under reduced pressure. The residual yellow oil is purified via column chromatography with 30% EtOAc/ n-Hexane to yield 0.75 g (92%) of compound **13** as yellow liquid. R_f 0.57 (30% EtOAc/ n-Hexane). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.06 (dd, *J* = 8.2, 1.1 Hz, 1H), 7.85 (dd, *J* = 7.8, 0.5 Hz, 1H), 7.66 (td, *J* = 7.7, 1.1 Hz, 1H), 7.44 (ddd, *J* = 8.1, 1.2, 0.6 Hz, 1H), 4.95 (s, 2H), 3.91 (t, *J* = 5.9 Hz, 2H), 3.55 (t, *J* = 5.9 Hz, 2H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 147.19, 134.70, 133.88, 128.76, 128.20, 124.79, 77.48, 77.16, 76.84, 71.01, 69.61, 30.39.

(2R,3R,4R,5R)-5-(6-amino-9H-purin-9-yl)-2-(hydroxymethyl)-4-(2-((2-nitrobenzyl)oxy)ethoxy)tetrahydrofuran-3-ol (**14**). Adenosine (2 g, 7.5 mmol)is dissolved in anhydrous DMF (50 ml) and warmed until a clear solution. The resulting solution is cooled in ice-bath and then added sodium hydride (60% mineral oil) (0.39 g, 9.7 mmol) and lithium bromide (1.9 g, 22.5 mmol). After stirred for 40 min, compound **13** is added drop wise with syringe and allowed to stir at ambient temperature for 12 h. The reaction mixture is quenched with acetic acid (0.1 ml) and concentrated under reduced pressure. The crude reaction mixture dissolved in EtOAc (25 ml) and water (15 ml). The layers were separated and the aqueous layer extracted with EtOAc (25 ml, 3 times). The combined EtOAc extracts is washed with brine (15 ml) and dried over anhydrous sodium sulphate. The EtOAc layer separated and evaporated to dryness. The resulting thick syrup directly loaded on to silica-gel column and purified with 5% MeOH/ CH_2Cl_2 to produce 0.26 g of compound **14** as white foam with 9% isolated yield. R_f 0.25 (4% MeOH/ CH_2Cl_2). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.20 (s, 1H), 7.98 (d, *J* = 8.0 Hz, 1H), 7.88 (s, 1H), 7.64 – 7.54 (m, 2H), 7.44 – 7.34 (m, 1H), 6.72 (b, 1H), 6.62 (s, 2H), 5.91 (d, *J* = 7.4 Hz, 1H), 4.84 (dd, *J* = 7.3, 4.5 Hz, 1H), 4.77 (t, *J* = 9.6 Hz, 2H), 4.57 (d, *J* = 4.5 Hz, 1H), 4.32 (s, 1H), 3.92 (d, *J* = 12.0 Hz,

1H), 3.81 – 3.54 (m, 5H), 2.89 (dd, *J* = 53.5, 23.8 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 156.26, 152.46, 148.44, 147.32, 140.91, 133.91, 133.78, 128.60, 128.39, 124.85, 120.96, 89.50, 88.12, 81.69, 70.79, 70.26, 69.92, 69.82, 63.23. ESI-HRMS $[M+Na]^+$ calcd for $C_{19}H_{22}N_6O_7Na$ 469.1442, found 469.1412.

N-(9-((2R,3R,4R,5R)-4-hydroxy-5-(hydroxymethyl)-3-(2-((2-nitrobenzyl)oxy)ethoxy)tetrahydrofuran-2-yl)-9H-purin-6-yl)benzamide (**15**). Compund **14** (0.2 g, 0.45 mmol) was dissolved in anhydrous pyridine (3 ml) and cooled to 0 ºC. Trimethylsilyl chloride (0.5 ml, 3.6 mmol) was added to the solution and resulting cloudy mixture was stirred for 40 min. Then benzoylchloride (0.3 ml, 2.5 mmol) was added drop wise. After stirred for 5 h, the reaction mixture was cooled to 0 ºC and added water (0.5 ml). After 30 min, concentrated NH_4OH (1 ml) was added, and allowed to stir at 0 ºC for additional 60 min. The crude reaction mixture concentrated under reduced pressure. The resulting oil partitioned between EtOAc (20 ml) and water (20 ml).The aqueous layer was extracted with EtOAc (3 × 20 ml). The combined organic extracts were washed with brine (20 ml), and dried over $NaSO_4$, and concentrated under vacuum. The residual colorless viscous oil was purified via chromatography eluted with 3% MeOH/ CH_2Cl_2 to yield 0.22 g (89%) of **15** as white foam. R_f 0.2 (4% MeOH/ CH_2Cl_2). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.16 (s, 1H), 8.76 (s, 1H), 8.08 (s, 1H), 8.05 (t, *J* = 7.4 Hz, 3H), 7.69 – 7.60 (m, 3H), 7.54 (t, *J* = 7.6 Hz, 2H), 7.51 – 7.44 (m, 1H), 6.10 (dd, *J* = 11.6, 1.7 Hz, 1H), 5.97 (d, *J* = 7.5 Hz, 1H), 4.93 – 4.81 (m, 3H), 4.59 (d, *J* = 4.4 Hz, 1H), 4.38 (s, 1H), 3.99 (d, *J* = 12.9 Hz, 1H), 3.71 (dd, *J* = 14.0, 12.8, 8.5 Hz, 5H), 3.56 (s, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 164.63, 152.27, 150.68, 150.49, 147.60, 143.46, 133.88, 133.75, 133.54, 133.12, 129.07, 128.78, 128.66, 128.05, 125.03, 124.75, 89.84, 88.24, 81.75, 70.85, 70.33, 70.13, 69.91, 63.43. ESI-LCMS $[M+H]^+$ calcd for $C_{26}H_{26}N_6O_8$ 551.1885, found 551.1946.

N-(9-((2R,3R,4R,5R)-5-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-4-hydroxy-3-(2-((2-nitrobenzyl)oxy)ethoxy)tetrahydrofuran-2-yl)-9H-purin-6-yl)benzamide (**16**). Compound **15** (0.82 g, 1.5 mmol) was co-evaporated with dry pyridine for 3 times and then dissolved in dry CH_2Cl_2 (10 ml). Dry triethylamine (1 ml, 7.5 mmol) followed by DMTr-Cl (0.81 gm, 2.4 mmol) was added to the resulting solution and stirred at room temperature for 24 h. The crude reaction mixture concentrated to dryness and then subjected to column chromatography (silica-gel prewashed with 0.5% triethylamine/2.5% MeOH/CH_2Cl_2 and eluted with 2.5% MeOH/CH_2Cl_2 to yield 0.74 g (58%) of compound **10a** as white solid and 0.3 g of starting material was recovered. R_f 0.3 (3% MeOH/ CH_2Cl_2). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.10 (s, 1H), 8.73 (s, 1H), 8.23 (s, 1H), 8.03 (d, *J* = 7.3 Hz, 3H), 7.67 (d, *J* = 7.3 Hz, 1H), 7.61 (dd, *J* = 10.6, 4.2 Hz, 2H), 7.52 (t, *J* = 7.5 Hz, 2H), 7.47 – 7.40 (m, 3H), 7.37 – 7.17 (m, 8H), 6.80 (t, *J* = 5.9 Hz, 4H), 6.22 (d, *J* = 4.3 Hz, 1H), 4.90 (s, 2H), 4.74 (t, *J* = 4.7 Hz, 1H), 4.53 (q, *J* = 5.1 Hz, 1H), 4.28 (dd, *J* = 8.0, 4.1 Hz, 1H), 4.03 (ddd, *J* = 11.1, 4.8, 2.3 Hz, 1H), 3.92 – 3.83 (m, 1H), 3.83 – 3.68 (m, 8H), 3.53 (dd, *J* = 10.6, 3.2 Hz, 1H), 3.42 (dd, *J* = 10.7, 4.3 Hz, 1H), 3.27 (d, *J* = 5.6 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 164.70, 158.70, 152.85, 151.60, 149.61, 147.57, 144.63, 141.92, 135.82, 135.76, 134.11, 133.83, 133.79, 132.92, 130.22, 130.20, 129.00, 128.74, 128.49, 128.31, 128.03, 128.00, 127.09, 124.93, 123.60, 113.33, 87.36, 86.75, 84.32, 82.26, 70.46, 70.33, 70.09, 70.05, 63.19, 55.36. ESI-LCMS $[M+H]^+$ calcd for $C_{47}H_{44}N_6O_{10}$ 853.3192, found 853.3285.

(2R,3R,4R,5R)-5-(6-benzamido-9H-purin-9-yl)-2-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-4-(2-((2-nitrobenzyl)oxy)ethoxy)tetrahydrofuran-3-yl (2-cyanoethyl) diisopropylphosphoramidite (**17**). 2-Cyanoethyldiisopropylchlorophosphoramidite (0.12 ml, 0.47 mmol) was added at room temperature to a stirred solution of compound **16** (0.2 g, 0.23 mmol) in freshly distilled dry THF (1 ml) and N,N-Diisopropyl ethylamine (0.24 ml, 1.4

mmol). After stirred for 4.5 h, the crude reaction mixture was directly subjected to column chromatography (silica-gel was prewashed with 1% triethylamine/60% EtOAc/ n-Hexane and eluted with 60% EtOAc/ n-Hexane) to yield 0.21 gm (86%) of compound **17** as white foam. R_f 0.45 (60% EtOAc/ n-Hexane). ^{31}P NMR (162 MHz, $CDCl_3$) δ 151.09 (s), 150.57 (s). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.01 (d, *J* = 7.3 Hz, 1H), 8.70 (d, *J* = 10.3 Hz, 1H), 8.23 (d, *J* = 20.0 Hz, 1H), 8.03 (d, *J* = 7.6 Hz, 3H), 7.61 (dd, *J* = 8.2, 6.3 Hz, 2H), 7.53 (t, *J* = 7.6 Hz, 3H), 7.48 – 7.41 (m, 2H), 7.35 – 7.28 (m, 4H), 7.28 – 7.20 (m, 4H), 6.80 (dd, *J* = 8.7, 7.3 Hz, 4H), 6.26 – 6.19 (m, 1H), 4.91 (dt, *J* = 17.8, 4.9 Hz, 1H), 4.83 (d, *J* = 8.3 Hz, 2H), 4.73 – 4.59 (m, 1H), 4.41 (dd, *J* = 22.5, 3.9 Hz, 1H), 3.97 (ddt, *J* = 23.7, 17.8, 5.7 Hz, 2H), 3.88 – 3.82 (m, 1H), 3.82 – 3.74 (m, 6H), 3.71 (dd, *J* = 8.7, 4.7 Hz, 3H), 3.67 – 3.51 (m, 4H), 3.42 – 3.26 (m, 1H), 2.69 – 2.56 (m, 1H), 2.37 (t, *J* = 6.4 Hz, 1H), 1.25 – 1.12 (m, 9H), 1.05 (t, *J* = 13.3 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 164.64, 158.70, 152.80, 151.70, 149.60, 147.33, 144.56, 142.33, 142.21, 135.80, 135.74, 135.72, 135.04, 134.81, 133.78, 133.75, 132.89, 130.24, 129.00, 128.62, 128.57, 128.42, 128.34, 128.15, 127.99, 127.09, 124.77, 124.73, 123.56, 117.88, 117.54, 113.28, 87.54, 86.78, 86.70, 83.85, 81.39, 81.02, 71.43, 71.29 , 71.07 , 70.91, 70.53, 70.44, 69.87, 63.06, 62.66, 58.98.76, 58.17, 55.35, 43.53, 43.37, 43.25, 24.77, 24.70. ^{31}P NMR (162 MHz, $CDCl_3$) δ 151.09, 150.57. ESI-LCMS $[M+H]^+$ calcd for $C_{56}H_{62}N_8O_{11}P$ 1053.4270, found 1053.4175.

Solid-phase RNA/DNA Oligonucleotide synthesis:

All the RNA/DNA reagents were purchased from Sigma-Aldrich. The standard 2'-*O*-TOM/TBDMS-phosphoramidites were purchased from ChemGenes Inc.. All the RNA oligonucleotides were synthesized using ABI 394 DNA/RNA Synthesizer on multiple 0.2 µmol scale. Coupling time for modified phosphormaidites was increased to 18 min. 0.10 M concentrations were used for standard amidites and 0.13 M concentrations were used for modified

amidite in dry CH_3CN. After the synthesis, solid support (polystyrene resin beads) was transferred in to screw cap vial for cleavage and deprotection using ammonium hydroxide and 40% aqueous methyl amine (1:1) at 60 ^{0}C for 20 min. After brief cooling at 0 ^{0}C, the samples were centrifuged at 12000 rpm for 10 min and supernatant was transferred in to 2.0 ml eppendorf tube and concentrated using speed-vac. The resultant residue was dissolved in anhy. DMSO (100 ul), triethylamine-trifluoride (150 µl) and heated at 65 ^{0}C for 2.5 h. The crude RNA oligonucleotides were cooled to room temperature and precipitated with sodium acetate (3M) and 90% ethanol. The precipitated oligonucleotides were purified by 20% denaturating PAGE using 1xTBE (pH 8.3). The purified RNA was desalted using NAP-G25 columns using sterile water and quantified (OD260) using UV-Vis spectroscopy, the fractions containing RNA oligonucleotides were confirmed by mass spectroscopy.

Sample preparation for ESI-MS:

Purified single strand RNA's (appox. 2 µm) are diluted with 100 ul of Acetonitrile:Water:Formic acid, 0.1% (1:1) mixture and recorded mass in negative ion high mode.[2] The mass spectra were deconvoluted to obtain the multiple charge states of the observed mass fragment peaks. The following parameters were used to record the mass spectra.

Acquisition Parameter

Source Type	ESI	Ion Polarity	Negative	Set Nebulizer	1.0 Bar
Focus	Not active	Set Capillary	3500 V	Set Dry Heater	120 °C
Scan Begin	500 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	10000 m/z	Set Collision Cell RF	2700.0 Vpp	Set Divert Valve	Waste

2B.6 Reference

1. Ahuja, N.; Sharma, A. R.; Baylin, S. B., Epigenetic therapeutics: a new weapon in the war against cancer. *Annu. Rev. Med.* **2016,** *67*, 73-89.
2. Bjork, G., *Hatfield, DL, Lee, BJ and Pirtle, RM, Transfer RNA in Protein Synthesis*. CRC Press1992.
3. Roundtree, I. A.; Evans, M. E.; Pan, T.; He, C., Dynamic RNA modifications in gene expression regulation. *Cell* **2017,** *169*, 1187-1200.
4. HELSER, T. L.; DAVIES, J. E.; DAHLBERG, J. E., Change in methylation of 16S ribosomal RNA associated with mutation to kasugamycin resistance in Escherichia coli. *Nat. New Biol.* **1971,** *233*, 12-14.
5. Bokar, J. A., The biosynthesis and functional roles of methylated nucleosides in eukaryotic mRNA. In *Fine-tuning of RNA functions by modification and editing*, Springer2005; pp 141-177.
6. Limbach, P. A.; Crain, P. F.; McCloskey, J. A., Summary: the modified nucleosides of RNA. *Nucleic Acids Res.* **1994,** *22*, 2183-2196.
7. Agris, P. F., The importance of being modified: roles of modified nucleosides and Mg2+ in RNA structure and function. *Prog. Nucleic Acid Res. Mol. Biol.* **1996,** *53*, 79-129.
8. Li, F.; Zhao, D.; Wu, J.; Shi, Y., Structure of the YTH domain of human YTHDF2 in complex with an m 6 A mononucleotide reveals an aromatic cage for m 6 A recognition. *Cell Res.* **2014,** *24*, 1490-1492.
9. Pan, T., N6-methyl-adenosine modification in messenger and long non-coding RNA. *Trends Biochem. Sci.* **2013,** *38*, 204-209.

10. Fustin, J.-M.; Doi, M.; Yamaguchi, Y.; Hida, H.; Nishimura, S.; Yoshida, M.; Isagawa, T.; Morioka, M. S.; Kakeya, H.; Manabe, I., RNA-methylation-dependent RNA processing controls the speed of the circadian clock. *Cell* **2013,** *155*, 793-806.

11. Hastings, M. H., m6A mRNA methylation: a new circadian pacesetter. *Cell* **2013,** *155*, 740-741.

12. Bokar, J. A.; Rath-Shambaugh, M. E.; Ludwiczak, R.; Narayan, P.; Rottman, F., Characterization and partial purification of mRNA N6-adenosine methyltransferase from HeLa cell nuclei. Internal mRNA methylation requires a multisubunit complex. *J. Biol. Chem.* **1994,** *269*, 17697-17704.

13. Liu, J.; Yue, Y.; Han, D.; Wang, X.; Fu, Y.; Zhang, L.; Jia, G.; Yu, M.; Lu, Z.; Deng, X., A METTL3–METTL14 complex mediates mammalian nuclear RNA N 6-adenosine methylation. *Nat. Chem. Biol.* **2014,** *10*, 93-95.

14. Jia, G.; Fu, Y.; Zhao, X.; Dai, Q.; Zheng, G.; Yang, Y.; Yi, C.; Lindahl, T.; Pan, T.; Yang, Y.-G., N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat. Chem. Biol.* **2011,** *7*, 885-887.

15. Zheng, G.; Dahl, J. A.; Niu, Y.; Fedorcsak, P.; Huang, C.-M.; Li, C. J.; Vågbø, C. B.; Shi, Y.; Wang, W.-L.; Song, S.-H., ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol. Cell* **2013,** *49*, 18-29.

16. Wang, X.; Lu, Z.; Gomez, A.; Hon, G. C.; Yue, Y.; Han, D.; Fu, Y.; Parisien, M.; Dai, Q.; Jia, G., N 6-methyladenosine-dependent regulation of messenger RNA stability. *Nature* **2014,** *505*, 117-120.

17. Lan, Q.; Liu, P. Y.; Haase, J.; Bell, J. L.; Hüttelmaier, S.; Liu, T., The critical role of RNA m6A methylation in cancer. *Cancer Res.* **2019,** *79*, 1285-1292.

18. Li, X.; Xiong, X.; Wang, K.; Wang, L.; Shu, X.; Ma, S.; Yi, C., Transcriptome-wide mapping reveals reversible and dynamic N 1-methyladenosine methylome. *Nat. Chem. Biol.* **2016,** *12*, 311-316.

19. Song, M.-G.; Bail, S.; Kiledjian, M., Multiple Nudix family proteins possess mRNA decapping activity. *RNA* **2013,** *19*, 390-399.

20. Macon, J. B.; Wolfenden, R., 1-Methyladenosine. Dimroth rearrangement and reversible reduction. *Biochemistry* **1968,** *7*, 3453-3458.

21. Chen, Y.; Yang, S.; Peng, S.; Li, W.; Wu, F.; Yao, Q.; Wang, F.; Weng, X.; Zhou, X., N 1-Methyladenosine detection with CRISPR-Cas13a/C2c2. *Chem. Sci.* **2019,** *10*, 2975-2979.

22. Eun, S. Y.; Kim, H. J.; Kang, E. S.; Kim, H. J.; Lee, J. H.; Chang, K. C.; Kim, J.-H.; Hong, S.-C.; Seo, H. G., Farnesyl diphosphate synthase attenuates paclitaxel-induced apoptotic cell death in human glioblastoma U87MG cells. *Neurosci. Lett.* **2010,** *474*, 115-120.

23. Castiglioni, S.; Casati, S.; Ottria, R.; Ciuffreda, P.; AM Maier, J., N6-isopentenyladenosine and its analogue N6-benzyladenosine induce cell cycle arrest and apoptosis in bladder carcinoma T24 cells. *Anticancer Agents Med. Chem.* **2013,** *13*, 672-678.

24. Laezza, C.; Caruso, M. G.; Gentile, T.; Notarnicola, M.; Malfitano, A. M.; Di Matola, T.; Messa, C.; Gazzerro, P.; Bifulco, M., N6-isopentenyladenosine inhibits cell proliferation and induces apoptosis in a human colon cancer cell line DLD1. *Int. J. Cancer* **2009,** *124*, 1322-1329.

25. Ciaglia, E.; Abate, M.; Laezza, C.; Pisanti, S.; Vitale, M.; Seneca, V.; Torelli, G.; Franceschelli, S.; Catapano, G.; Gazzerro, P., Antiglioma effects of N 6-

isopentenyladenosine, an endogenous isoprenoid end product, through the downregulation of epidermal growth factor receptor. *Int. J. Cancer* **2017,** *140*, 959-972.

26. Ciaglia, E.; Grimaldi, M.; Abate, M.; Scrima, M.; Rodriquez, M.; Laezza, C.; Ranieri, R.; Pisanti, S.; Ciuffreda, P.; Manera, C., The isoprenoid derivative N6-benzyladenosine CM223 exerts antitumor effects in glioma patient-derived primary cells through the mevalonate pathway. *Br. J. Pharmacol.* **2017,** *174*, 2287-2301.

27. Al-Arif, A.; Sporn, M. B., 2′-O-Methylation of adenosine, guanosine, uridine, and cytidine in RNA of isolated rat liver nuclei. *Proc. Natl. Acad. Sci. USA* **1972,** *69*, 1716-1719.

28. Feder, M.; Pas, J.; Wyrwicz, L. S.; Bujnicki, J. M., Molecular phylogenetics of the RrmJ/fibrillarin superfamily of ribose 2′-O-methyltransferases. *Gene* **2003,** *302*, 129-138.

29. Inoue, H.; Hayase, Y.; Imura, A.; Iwai, S.; Miura, K.; Ohtsuka, E., Synthesis and hybridization studies on two complementary nona (2'-O-methyl) ribonucleotides. *Nucleic Acids Res.* **1987,** *15*, 6131-6148.

30. Freier, S. M.; Altmann, K.-H., The ups and downs of nucleic acid duplex stability: structure-stability studies on chemically-modified DNA: RNA duplexes. *Nucleic Acids Res.* **1997,** *25*, 4429-4443.

31. Levin, A. A., Treating disease at the RNA level with oligonucleotides. *N. Engl. J. Med.* **2019,** *380*, 57-70.

32. Levin, A. A., Targeting therapeutic oligonucleotides. *N Engl J Med.* **2017,** *376*, 86-88.

33. Marchand, V.; Blanloeil-Oillo, F.; Helm, M.; Motorin, Y., Illumina-based RiboMethSeq approach for mapping of 2′-O-Me residues in RNA. *Nucleic Acids Res.* **2016,** *44*, e135-e135.

34. Prakash, T. P., An overview of sugar-modified oligonucleotides for antisense therapeutics. *Chem. Biodivers.* **2011,** *8*, 1616-1641.

35. Odadzic, D.; Bramsen, J. B.; Smicius, R.; Bus, C.; Kjems, J.; Engels, J. W., Synthesis of 2′-O-modified adenosine building blocks and application for RNA interference. *Bioorg. Med. Chem.* **2008,** *16*, 518-529.

36. Griffey, R. H.; Monia, B. P.; Cummins, L. L.; Freier, S.; Greig, M. J.; Guinosso, C. J.; Lesnik, E.; Manalili, S. M.; Mohan, V.; Owens, S., 2 '-O-Aminopropyl ribonucleotides: a zwitterionic modification that enhances the exonuclease resistance and biological activity of antisense oligonucleotides. *J. Med. Chem.* **1996,** *39*, 5100-5109.

37. Jin, S.; Miduturu, C. V.; McKinney, D. C.; Silverman, S. K., Synthesis of amine-and thiol-modified nucleoside phosphoramidites for site-specific introduction of biophysical probes into RNA. *J. Org. Chem.* **2005,** *70*, 4284-4299.

38. Chaulk, S. G.; MacMillan, A. M., Synthesis of oligo-RNAs with photocaged adenosine 2′-hydroxyls. *Nat. Protoc.* **2007,** *2*, 1052-1058.

39. Wu, W.; Stupi, B. P.; Litosh, V. A.; Mansouri, D.; Farley, D.; Morris, S.; Metzker, S.; Metzker, M. L., Termination of DNA synthesis by N 6-alkylated, not 3′-O-alkylated, photocleavable 2′-deoxyadenosine triphosphates. *Nucleic Acids Res.* **2007,** *35*, 6339-6349.

40. Schäfer, F.; Wagner, J.; Knau, A.; Dimmeler, S.; Heckel, A., Regulating Angiogenesis with Light-Inducible AntimiRs. *Angew. Chem., Int. Ed.* **2013,** *52*, 13558-13561.

41. Edmonds, C. G.; Gupta, R.; McCloskey, J. A.; Crain, P. F., Ribonucleic acid modification in microorganisms. ACS Symposium Series1994; Vol. 541, pp 147-158.

42. Aritomo, K.; Wada, T.; Sekine, M., Alkylation of 6-N-acylated adenosine derivatives by the use of phase transfer catalysis. *J. Chem. Soc., Perkin Trans. 1* **1995**, 1837-1844.

43. Tang, X.; Zhang, J.; Sun, J.; Wang, Y.; Wu, J.; Zhang, L., Caged nucleotides/nucleosides and their photochemical biology. *Org. Biomol. Chem.* **2013,** *11*, 7814-7824.

44. Leonard, N. J.; Hecht, S. M.; Skoog, F.; Schmitz, R. Y., Cytokinins: Synthesis of 6-(3-methyl-3-butenylamino)-9-β-D-Ribofuranosylpurine (3iPA), and the Effect of Side-Chain Unsaturation on the Biological Activity of Isopentylaminopurines and their Ribosides. *Proc. Natl. Acad. Sci. USA* **1968,** *59*, 15-21.

45. Ottria, R.; Casati, S.; Manzocchi, A.; Baldoli, E.; Mariotti, M.; Maier, J. A.; Ciuffreda, P., Synthesis and evaluation of in vitro anticancer activity of some novel isopentenyladenosine derivatives. *Bioorg. Med. Chem.* **2010,** *18*, 4249-4254.

46. Leonard, N. J.; Achmatowicz, S.; Loeppky, R. N.; Carraway, K. L.; Grimm, W.; Szweykowska, A.; Hamzi, Q.; Skoog, F., Development of cytokinin activity by rearrangement of 1-substituted adenines to 6-substituted aminopurines: inactivation by N6, 1-cyclization. *Proc. Natl. Acad. Sci. USA* **1966,** *56*, 709-716.

47. Dominissini, D.; Nachtergaele, S.; Moshitch-Moshkovitz, S.; Peer, E.; Kol, N.; Ben-Haim, M. S.; Dai, Q.; Di Segni, A.; Salmon-Divon, M.; Clark, W. C., The dynamic N 1-methyladenosine methylome in eukaryotic messenger RNA. *Nature* **2016,** *530*, 441-446.

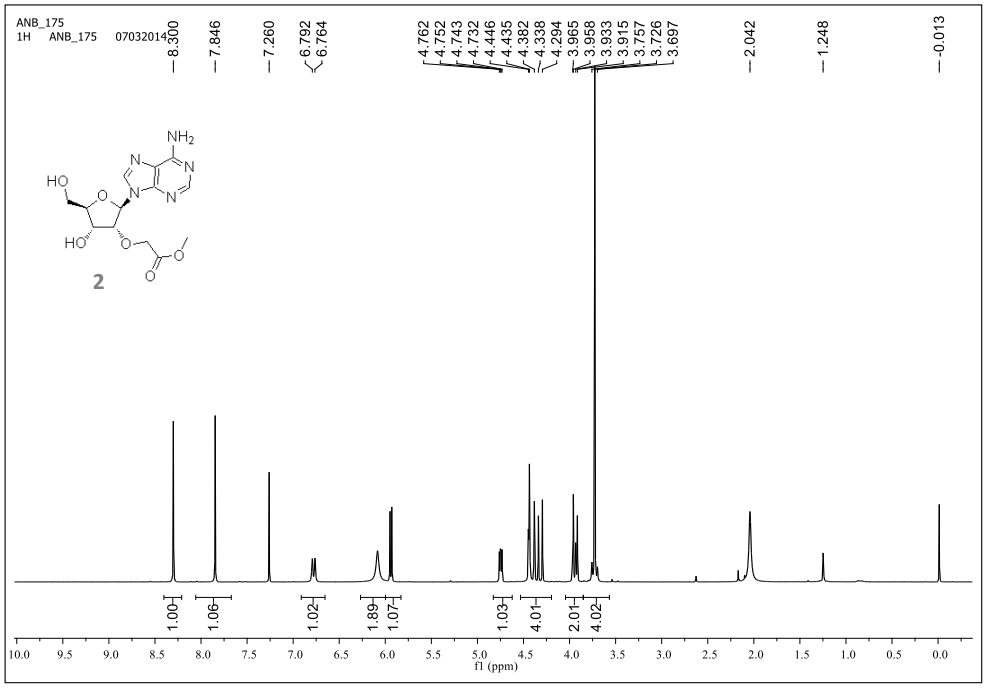
48. Griffey, R. H.; Monia, B. P.; Cummins, L. L.; Freier, S.; Greig, M. J.; Guinosso, C. J.; Lesnik, E.; Manalili, S. M.; Mohan, V.; Owens, S., 2 '-O-Aminopropyl ribonucleotides: a zwitterionic modification that enhances the exonuclease resistance and biological activity of antisense oligonucleotides. *J. Med. Chem.* **1996,** *39*, 5100-5109.

49. Arimitsu, K.; Kimura, H.; Kajimoto, T.; Ono, M.; Ohmomo, Y.; Yamashita, M.; Node, M.; Saji, H., Novel design and synthesis of a radioiodinated glycolipid analog as an acceptor

substrate for N-acetylglucosaminyltransferase V. *J. Labelled Comp. Radiopharm.* **2013,** *56*, 562-572.

2B.7 Appendix

1. ^{1}H-/ ^{13}C-/ESI-MS/HRMS spectra of compound **2**



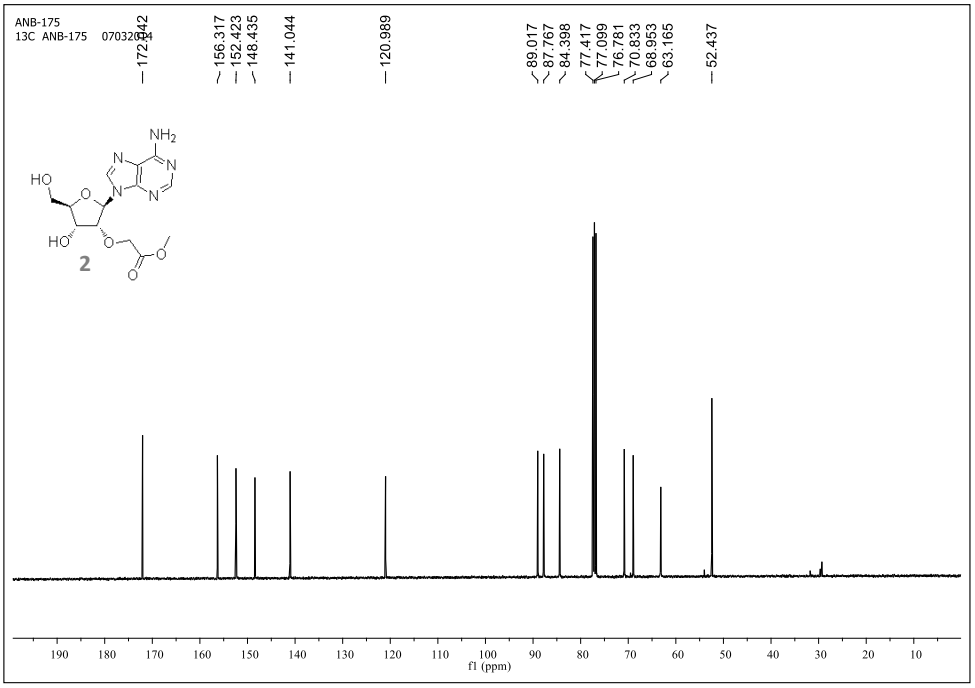


Figure S1. ^{1}H and ^{13}C NMR spectra of compound **2** in $CDCl_3$.

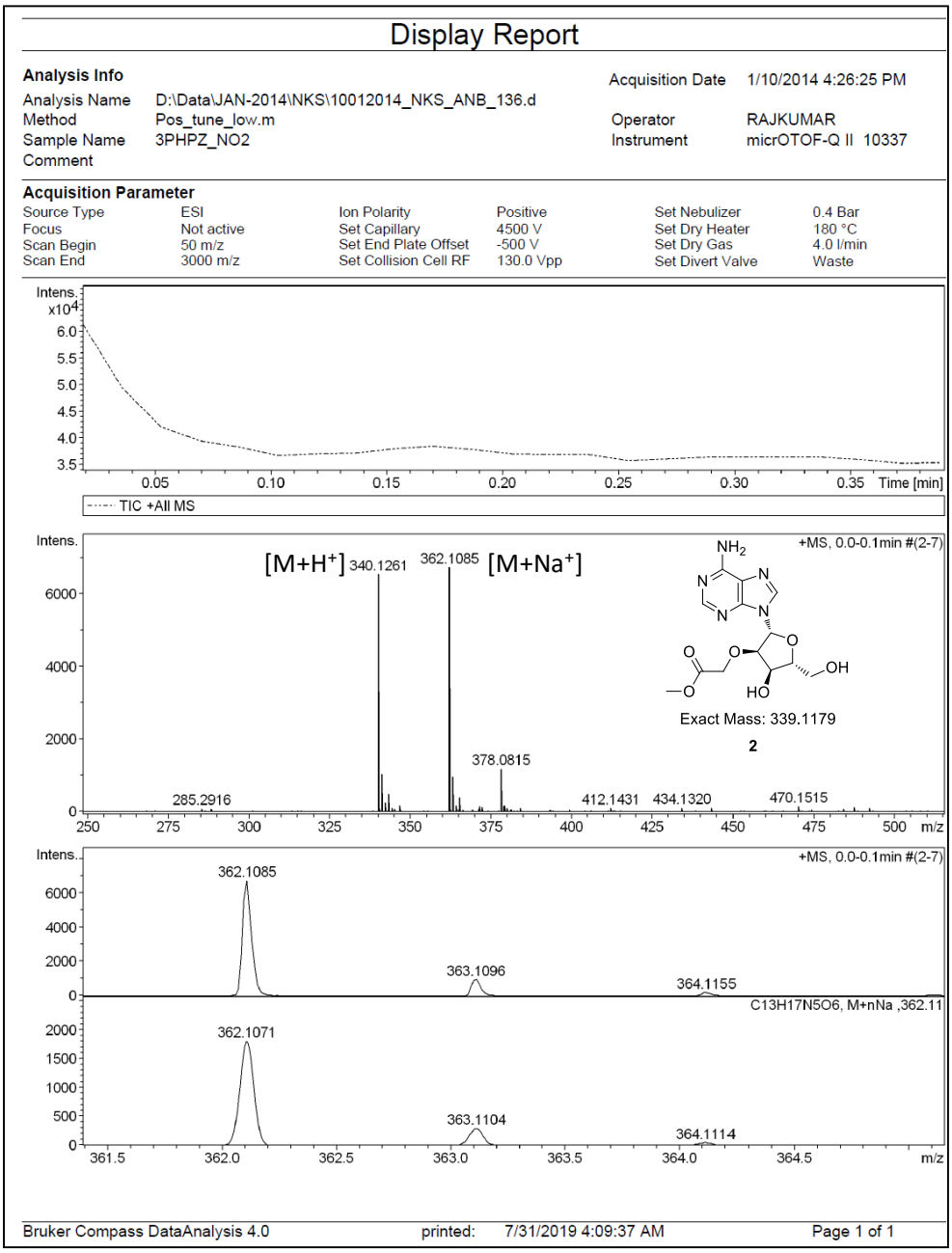


Figure S2. ESI-HRMS of compound **2**.

2.[1]H-/ESI-MS/HRMS spectra of compound **3**

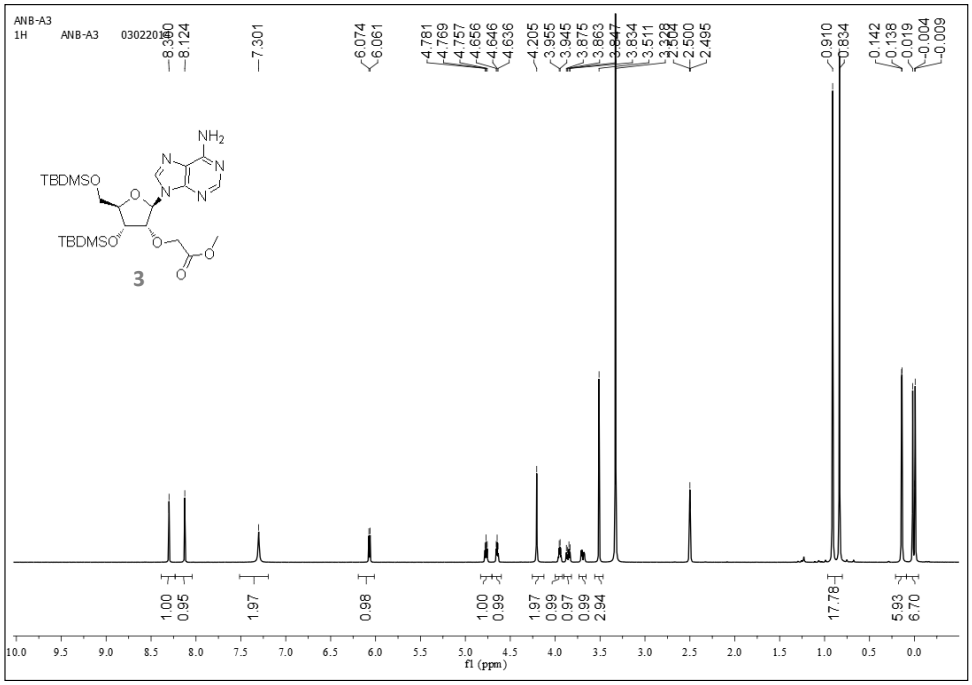


Figure S3. [1]H NMR spectra of compound **3** in $CDCl_3$.

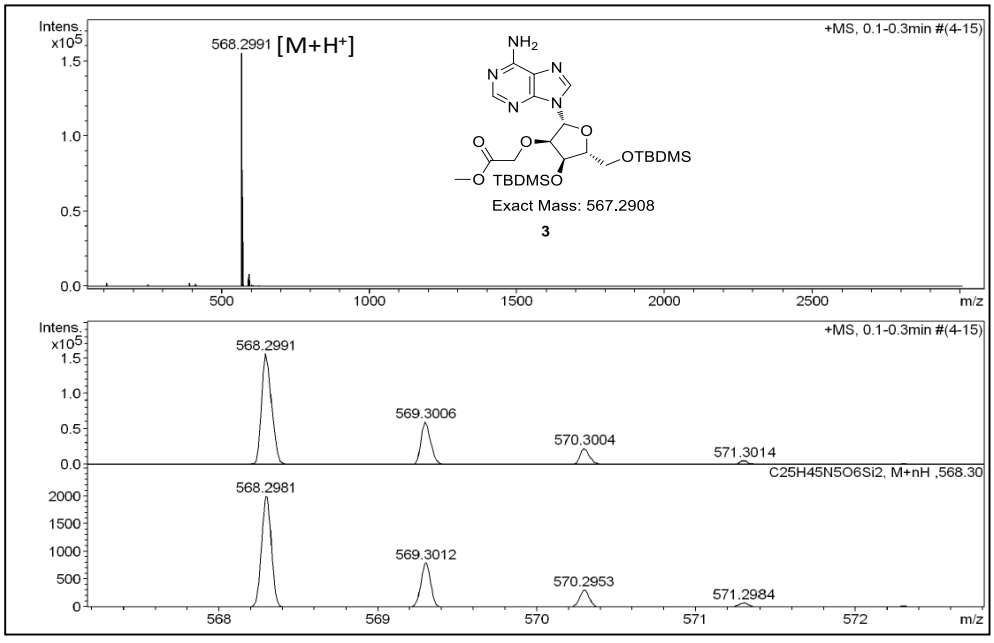
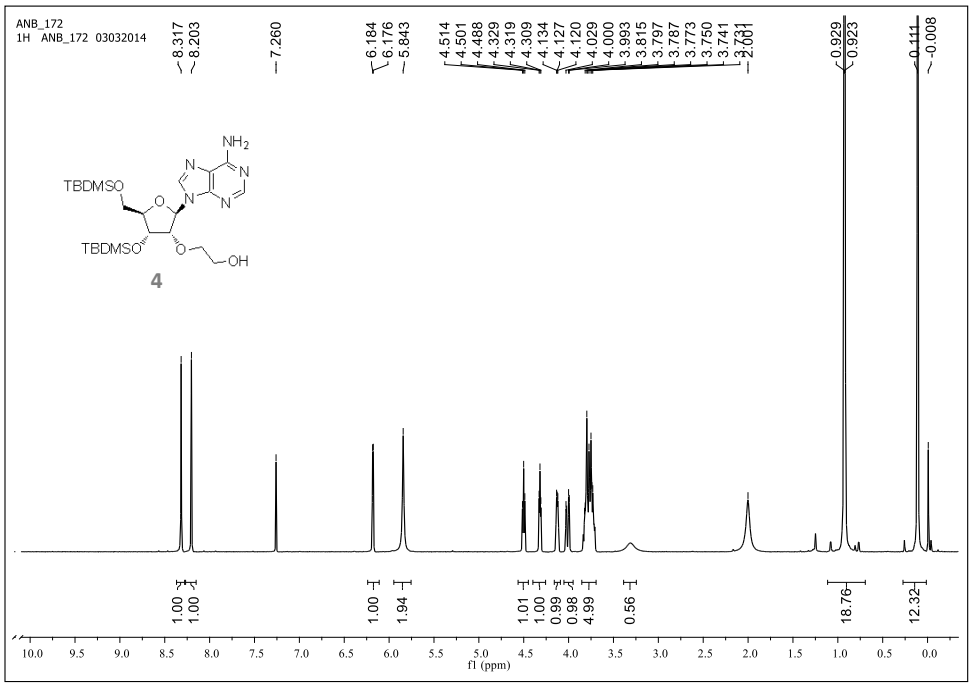


Figure S4. ESI-HRMS of compound **3**.

3.^{1}H-/ ^{13}C-/ESI-MS/HRMS spectra of compound **4**



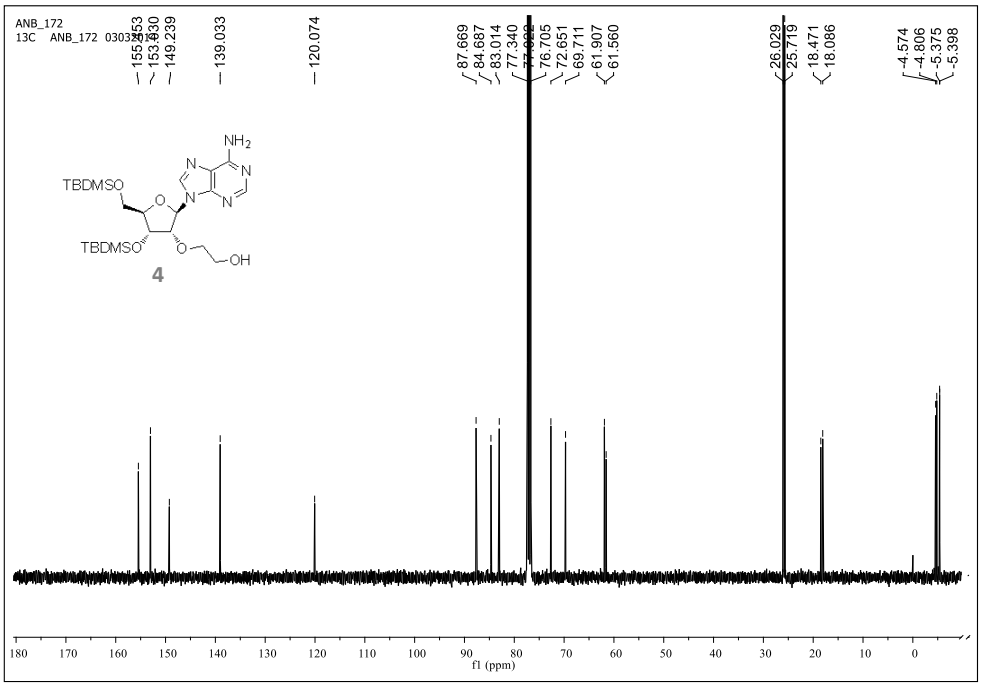


Figure S5. ^{1}H and ^{13}C NMR spectra of compound **4** in $CDCl_3$.

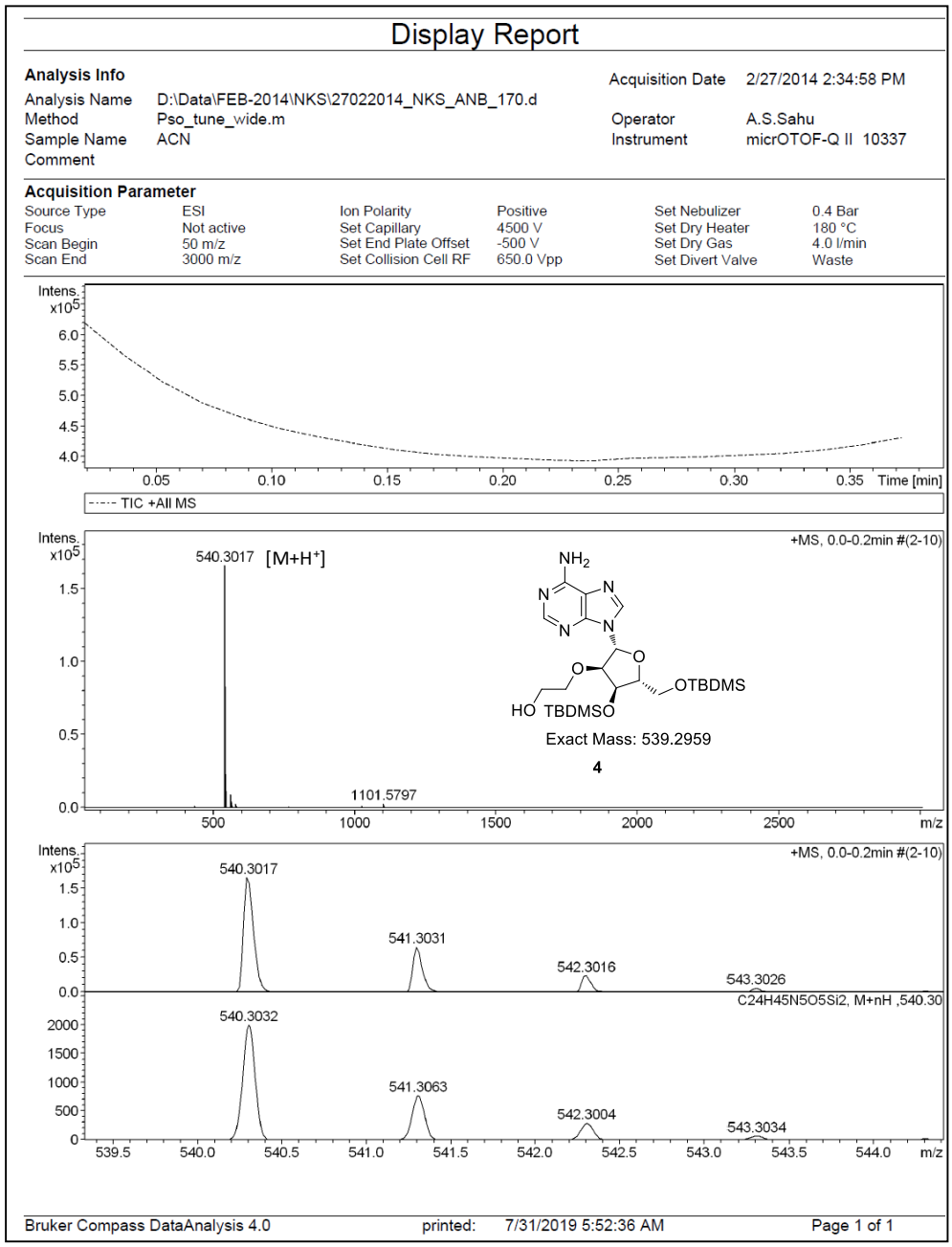


Figure S6. ESI-HRMS of compound **4**.

4. ^{1}H-/ ^{13}C-/ESI-MS/HRMS spectra of compound **5**

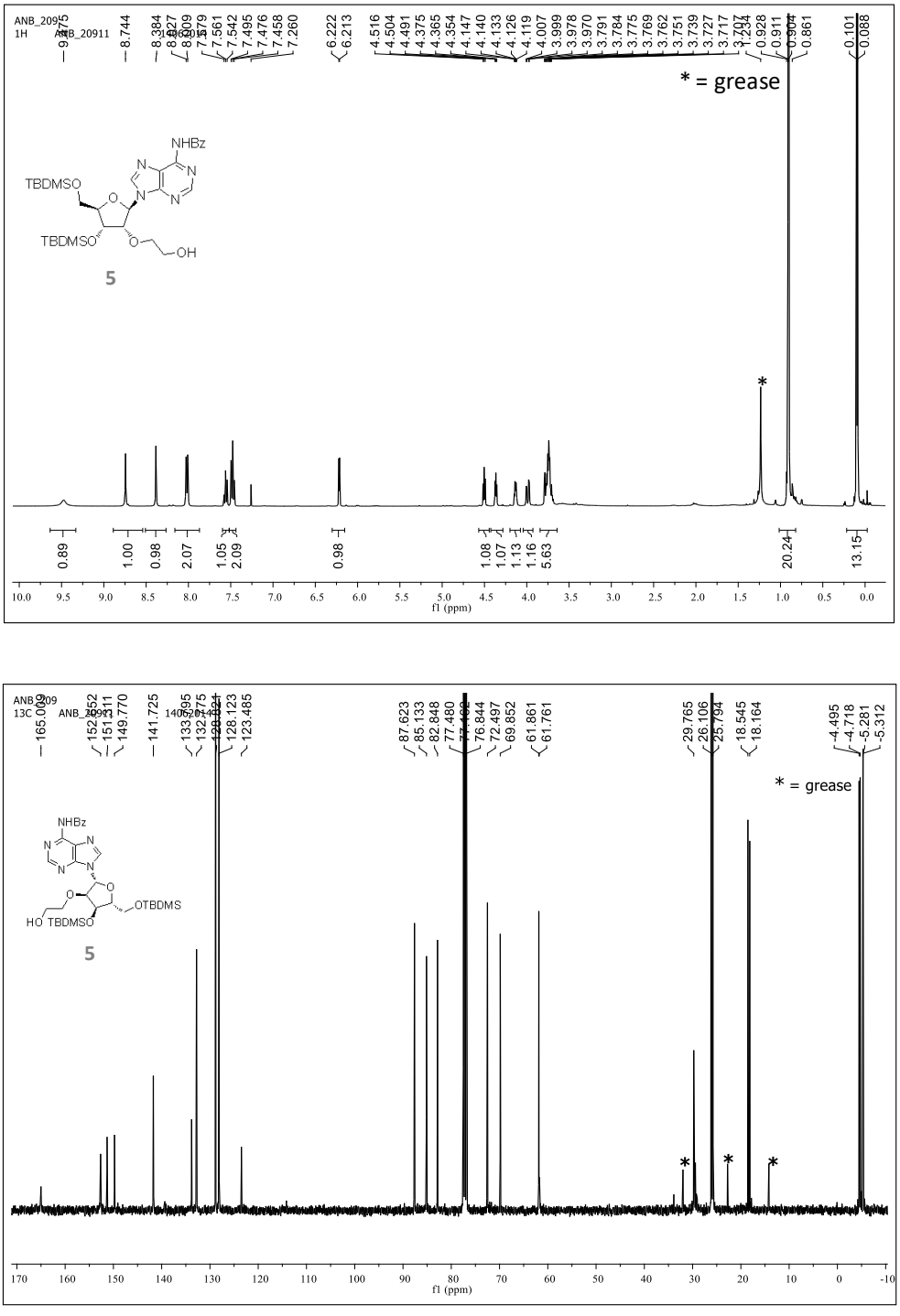


Figure S7. ^{1}H and ^{13}C NMR spectra of compound **5** in $CDCl_3$.

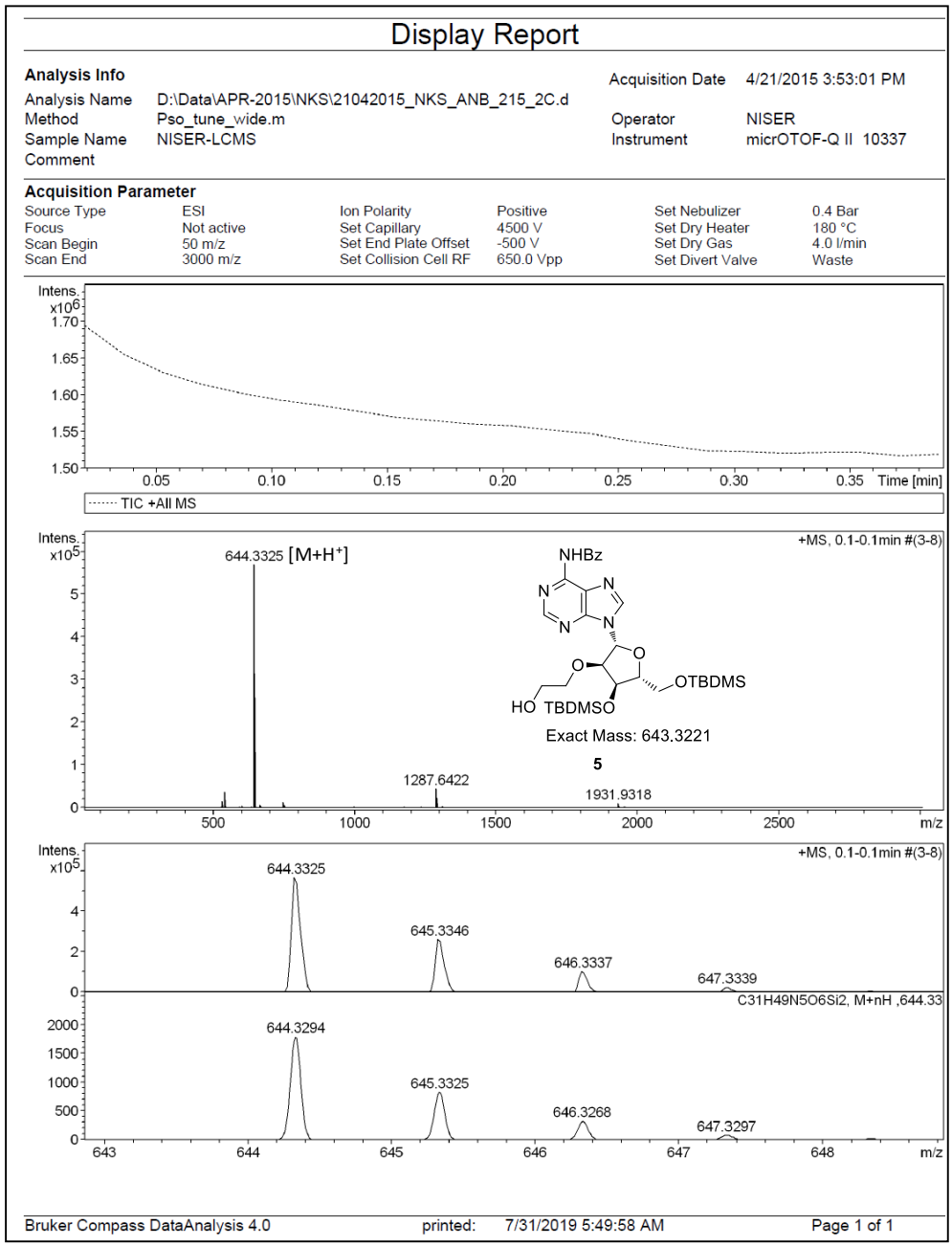


Figure S8. ESI-HRMS of compound **5**.

5.^{1}H-/ ^{13}C-/ESI-MS/HRMS spectra of NBn^{6}A derivative, **6a**

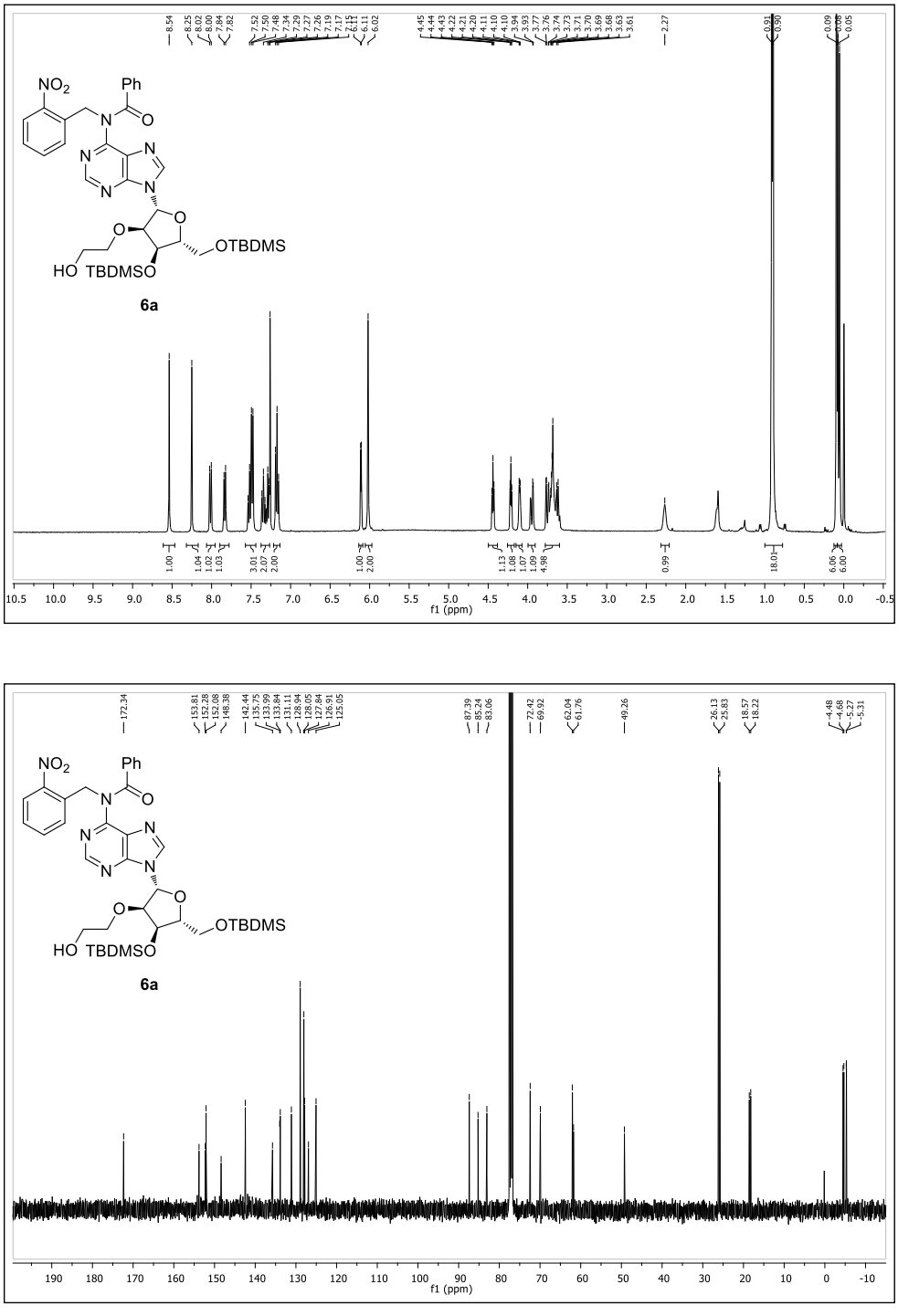


Figure S9. ^{1}H and ^{13}C NMR spectra of NBn^{6}A derivative, **6a** in $CDCl_3$.

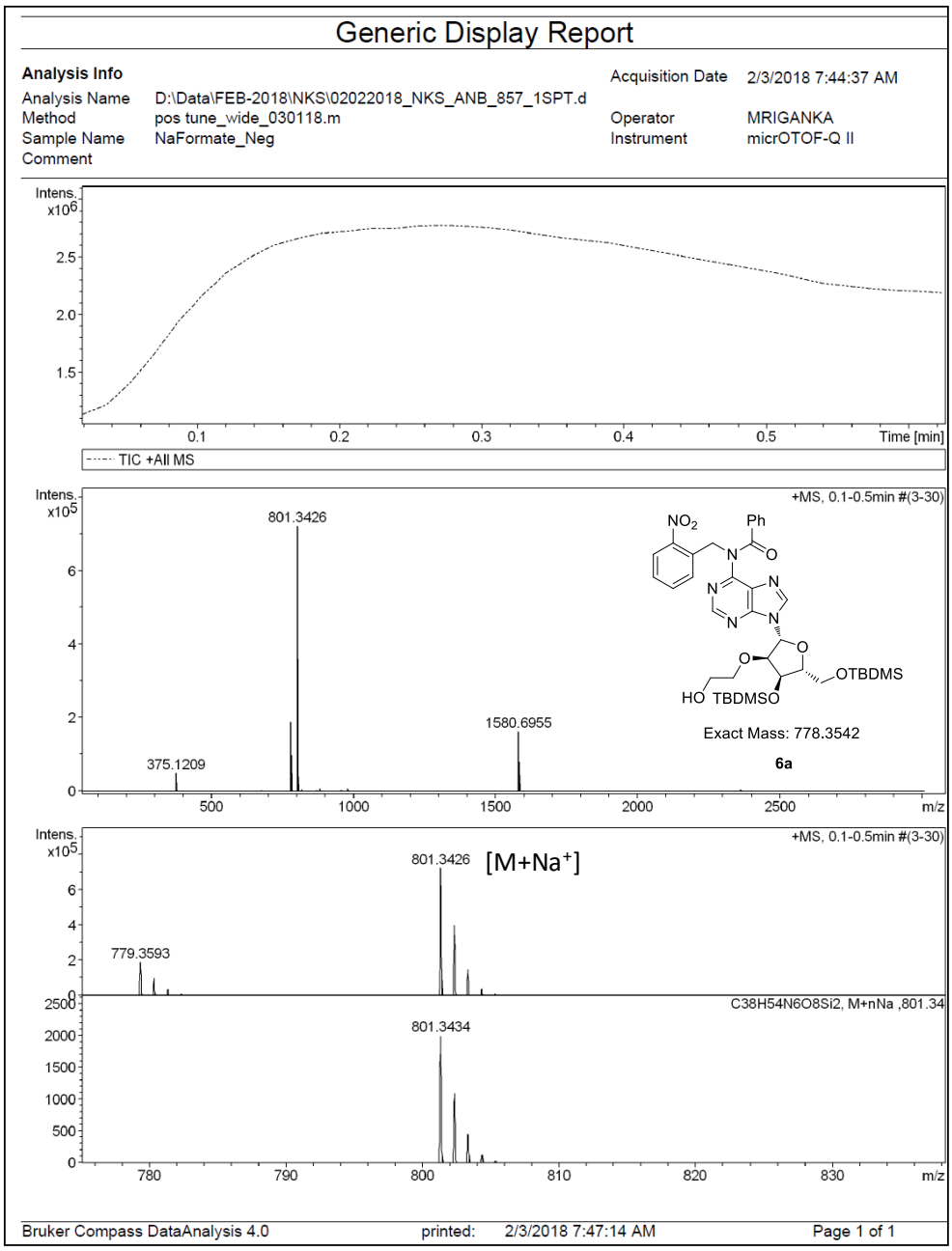


Figure S10. ESI-HRMS of NBn^{6}A derivative, **6a**.

6.^{1}H-/ ^{13}C-/ESI-MS/HRMS spectra of NBn^{6}A derivative, **7a**

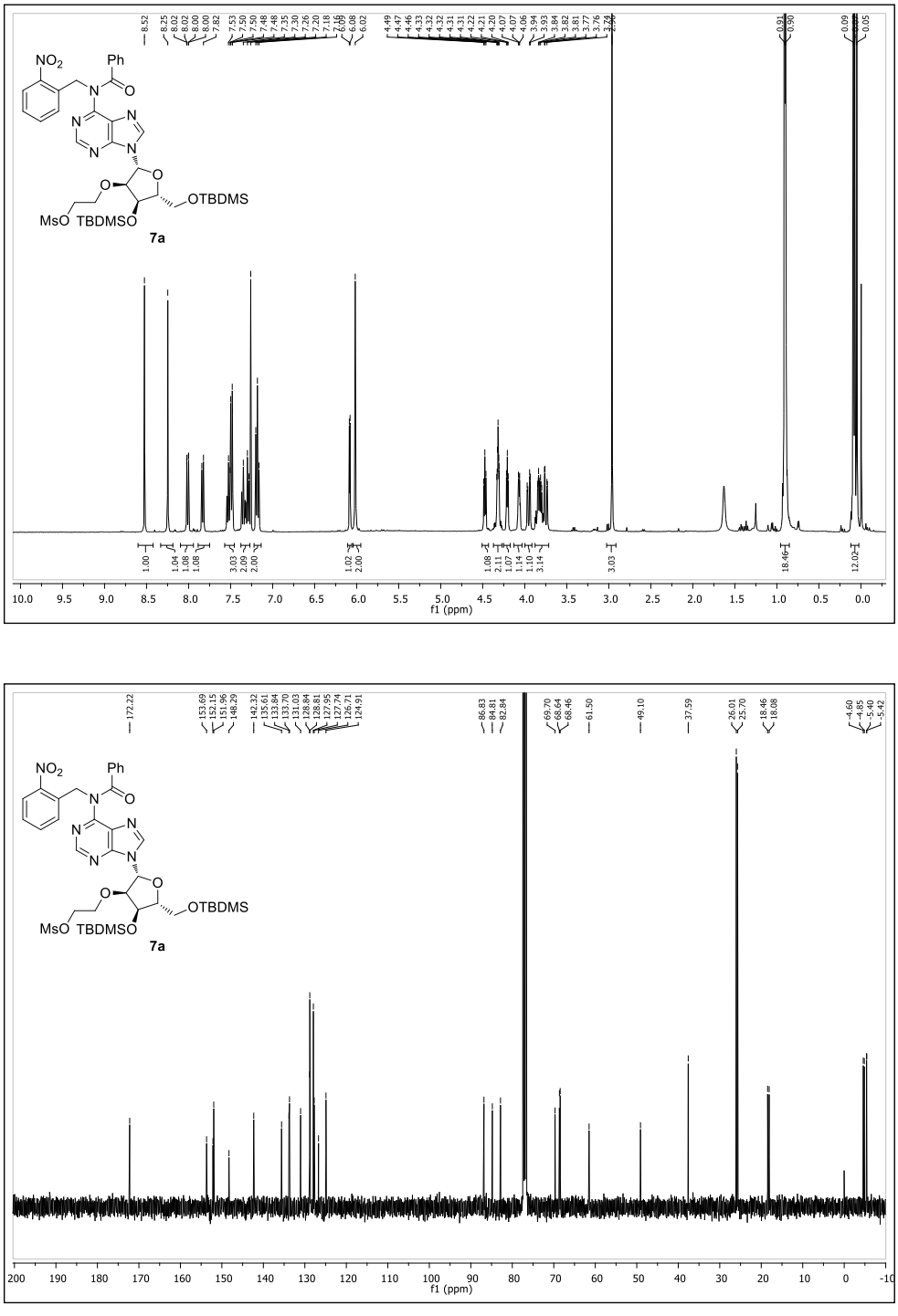


Figure S11. ^{1}H and ^{13}C NMR spectra of NBn^{6}A derivative, **7a** in $CDCl_3$.

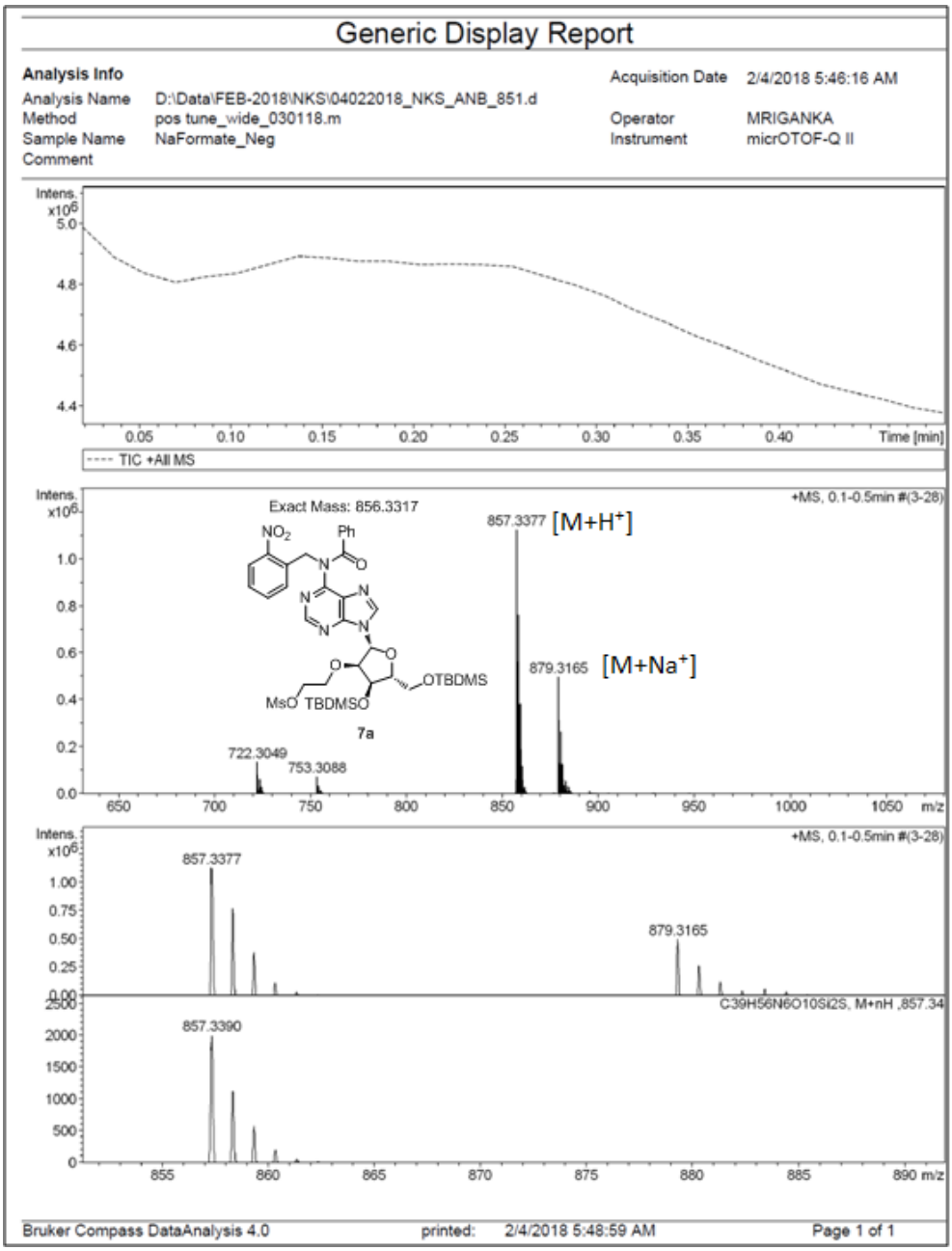
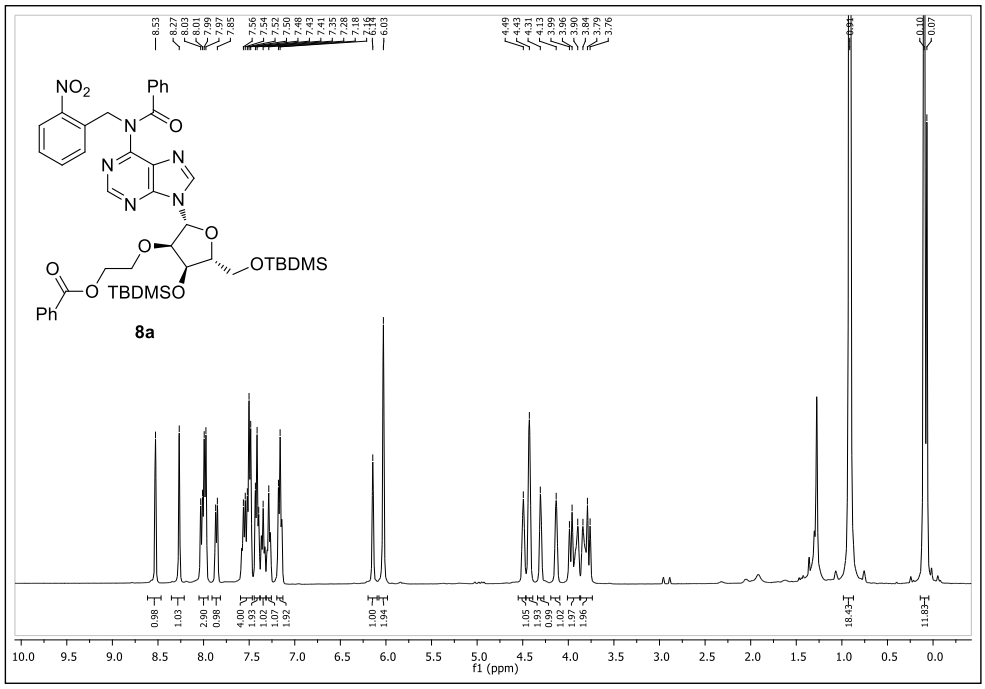


Figure S12. ESI-HRMS of NBn^{6}A derivative, **7a**.

7.^{1}H-/ ^{13}C-/ESI-MS/HRMS spectra of NBn^{6}A derivative, **8a**



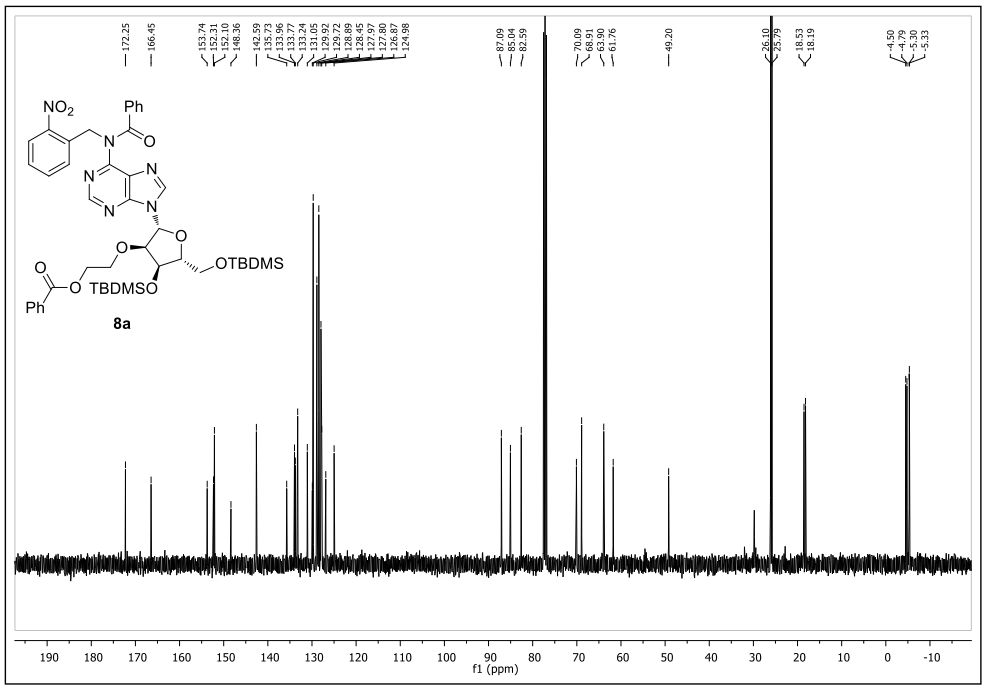


Figure S13. ^{1}H and ^{13}C NMR spectra of NBn^{6}A derivative, **8a** in $CDCl_3$.

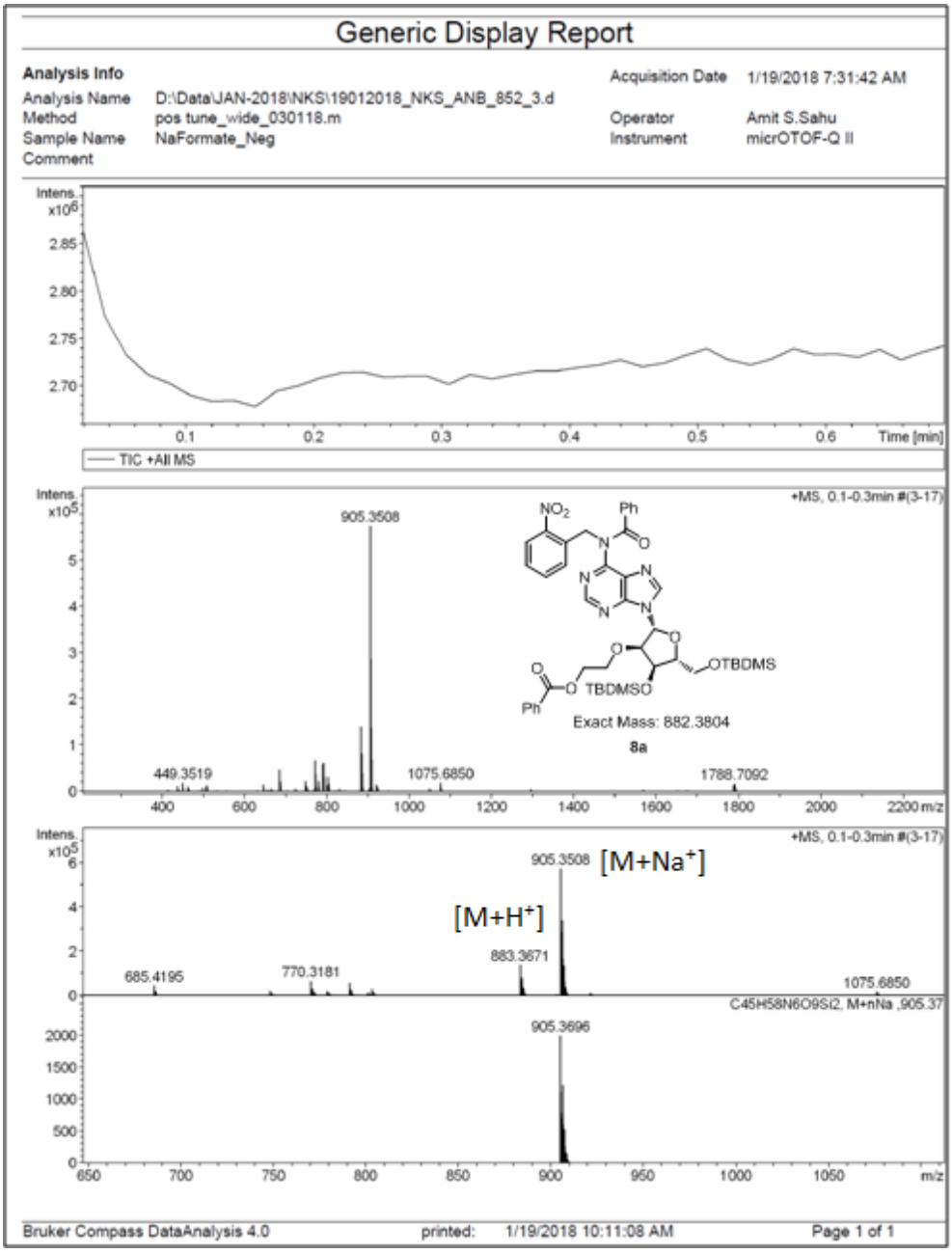


Figure S14. ESI-HRMS of NBn^{6}A derivative, **8a**.

8.^{1}H-/ ^{13}C-/ESI-MS/HRMS spectra of NBn^{6}A derivative, **9a**

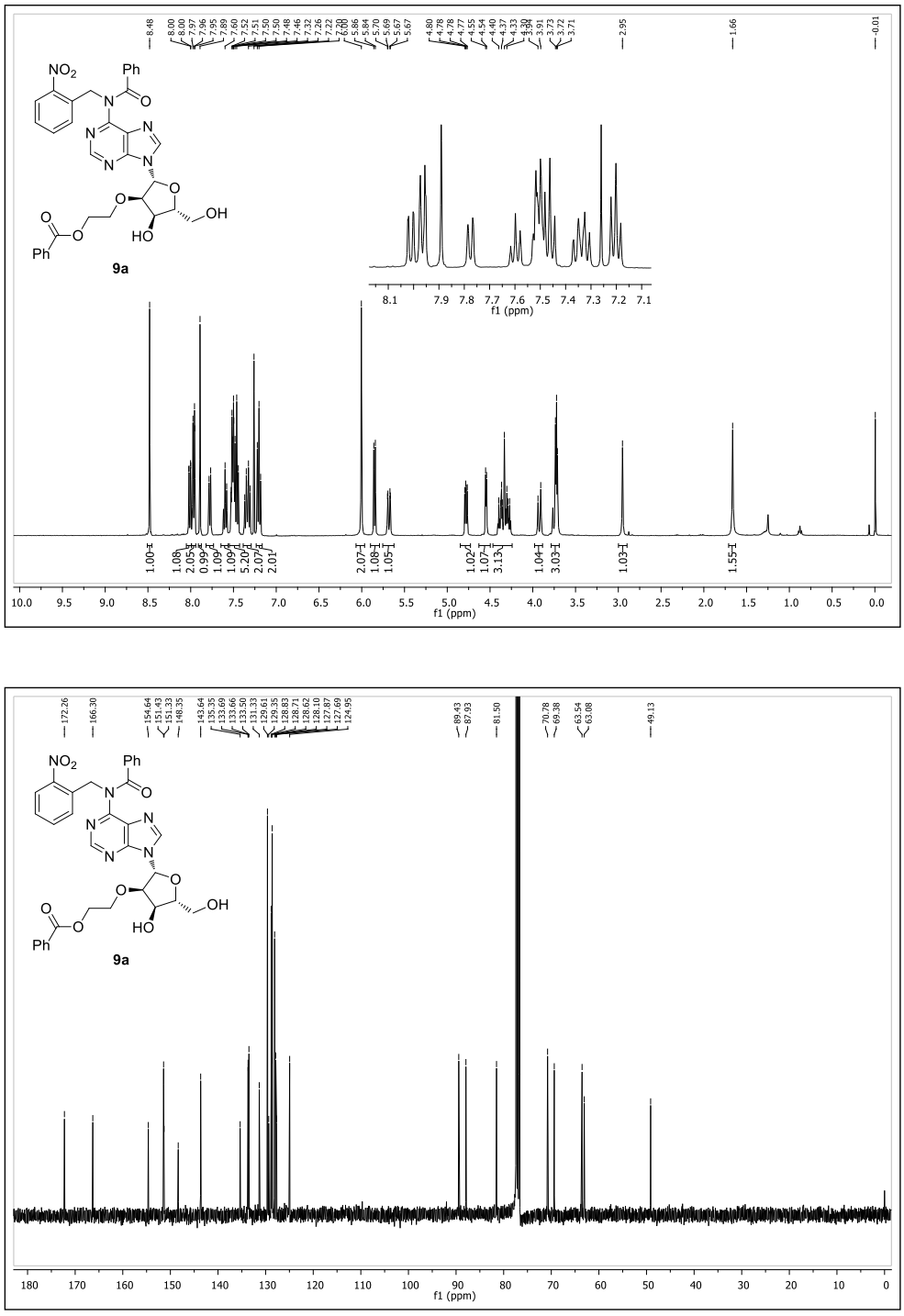


Figure S15. ^{1}H and ^{13}C NMR spectra of NBn^{6}A derivative, **9a** in $CDCl_3$.

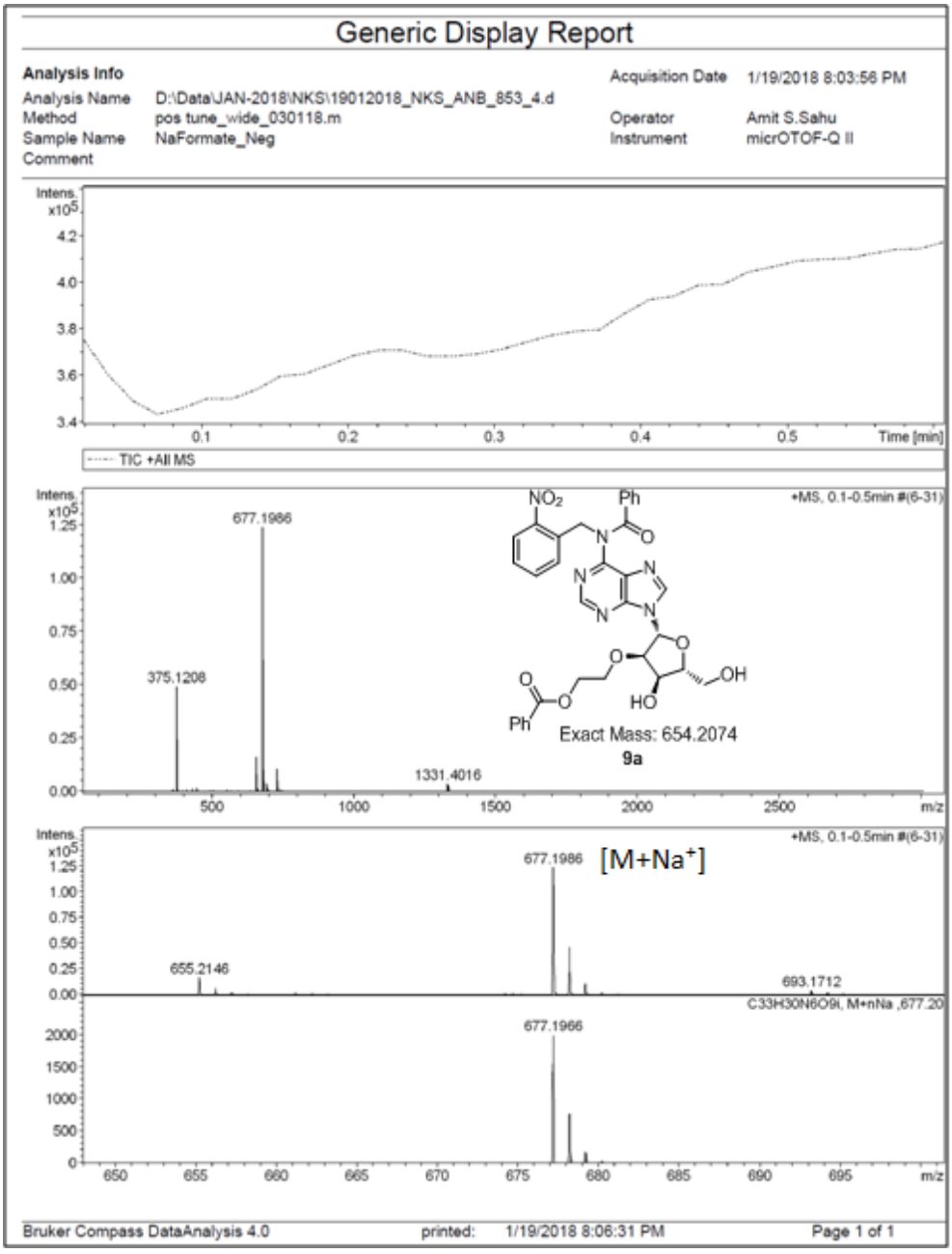


Figure S16. ESI-HRMS of NBn^{6}A derivative, **9a**.

9.^{1}H-/ ^{13}C-/ESI-MS/HRMS spectra of NBn^{6}A derivative, **10a**

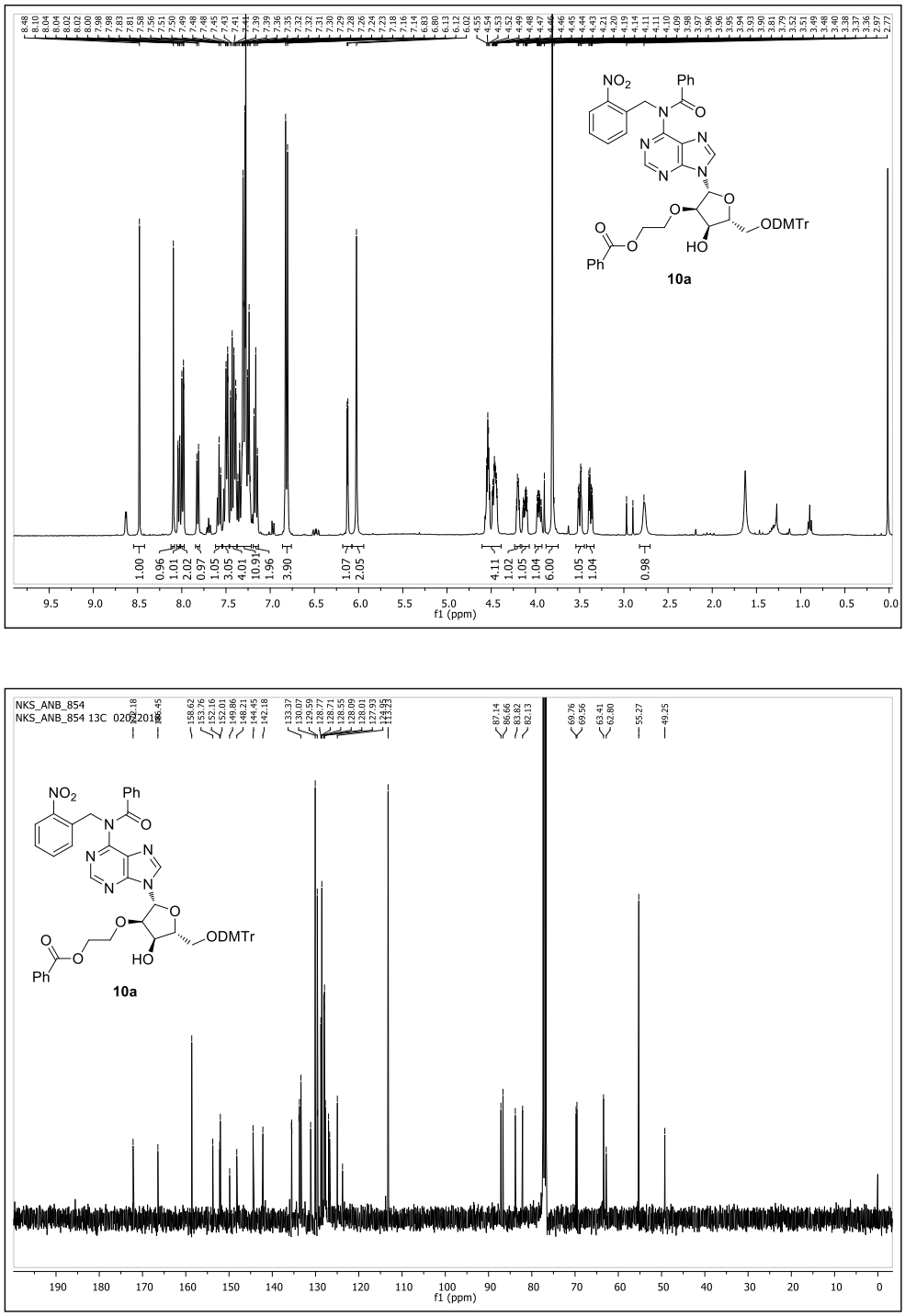


Figure S17. ^{1}H and ^{13}C NMR spectra of NBn^{6}A derivative, **10a** in $CDCl_3$.

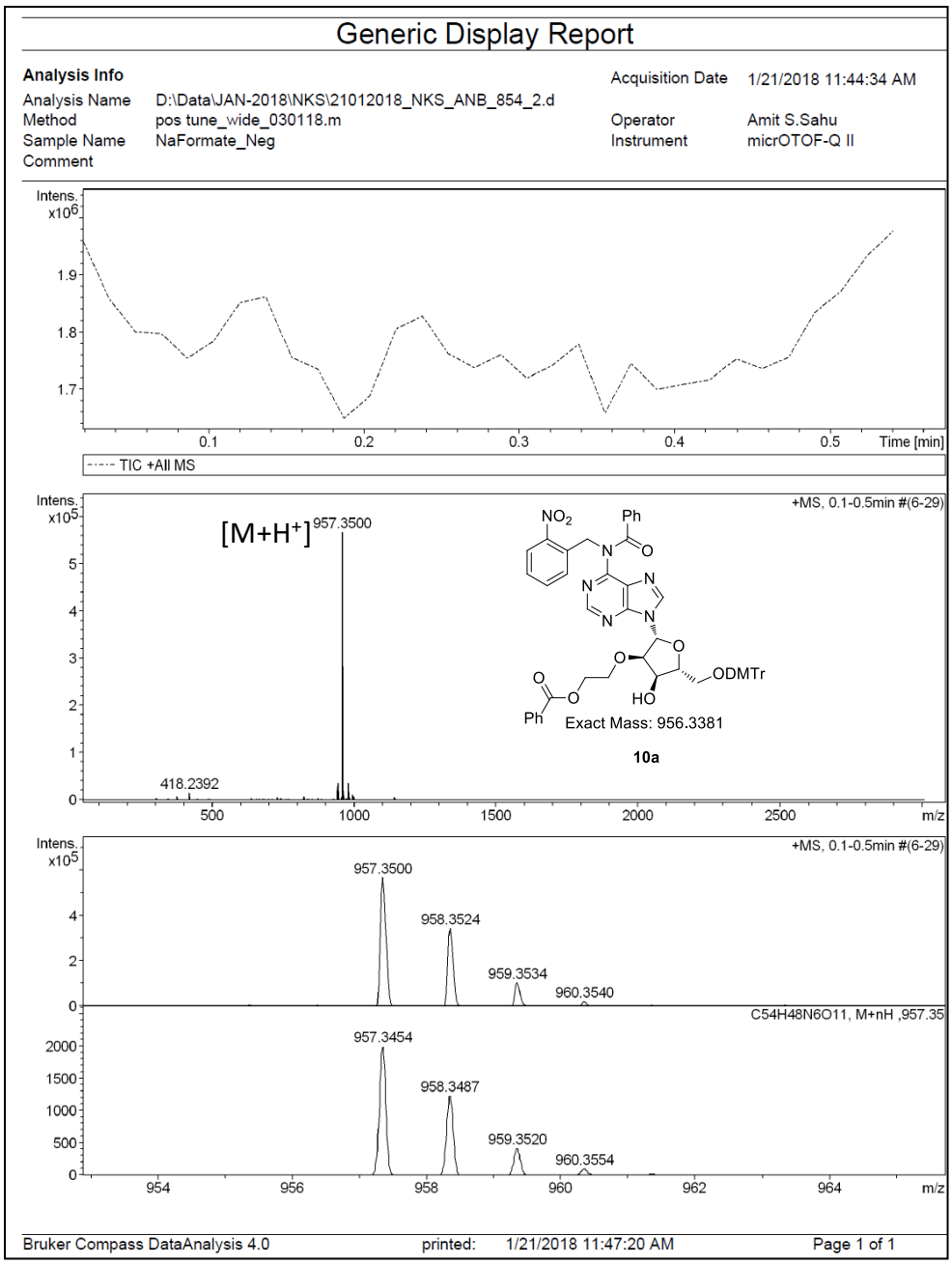


Figure S18. ESI-HRMS of 5'-ODMTr derivative of NBn^{6}A, **10a**.

10.^{1}H-/ ^{31}P-/ESI-MS/HRMS spectra of NBn^{6}A phophoramidite, **11a**

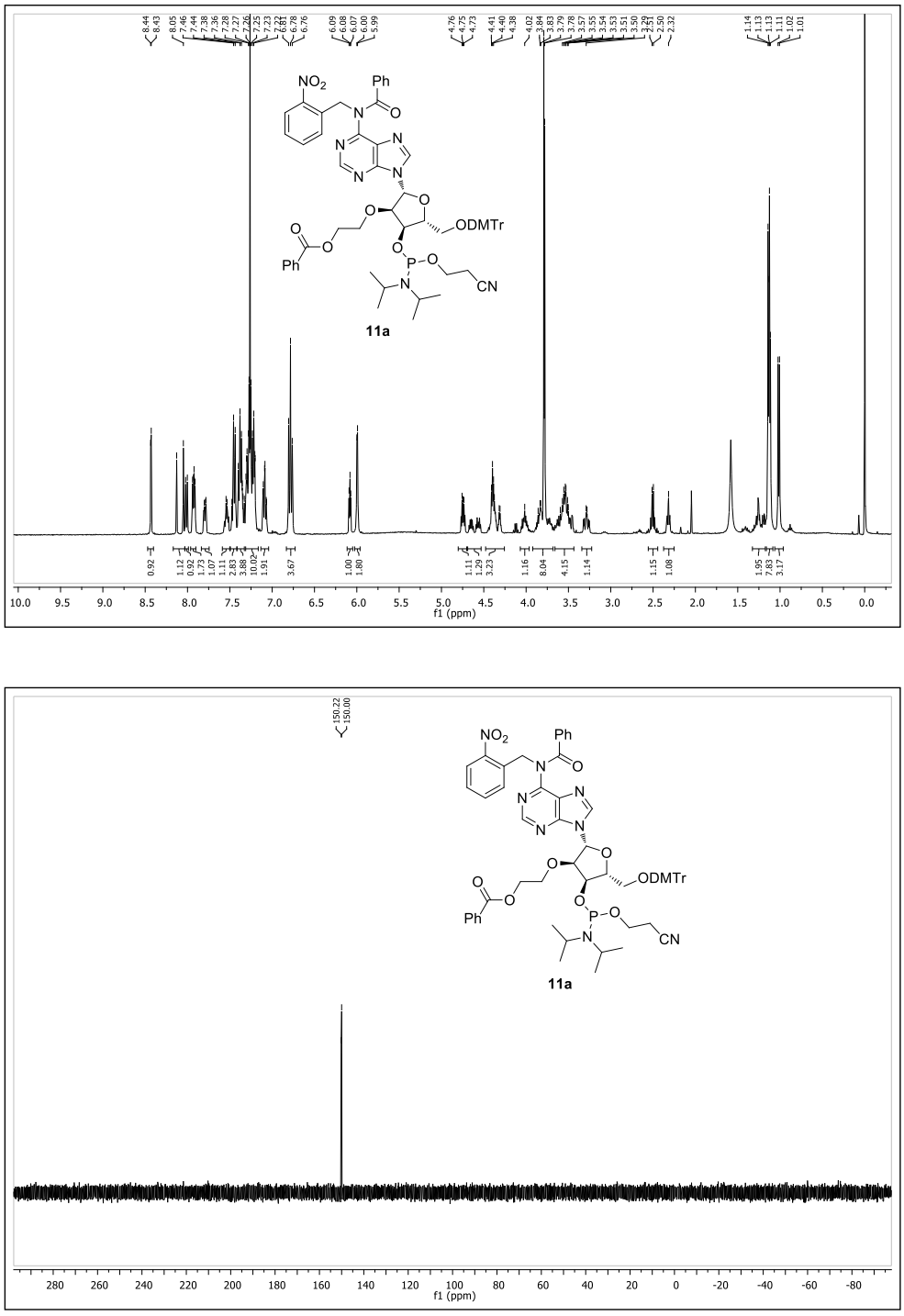


Figure S19. ^{1}H and ^{31}P NMR spectra of NBn^{6}A phosphoramidite derivative **11a** in $CDCl_3$.

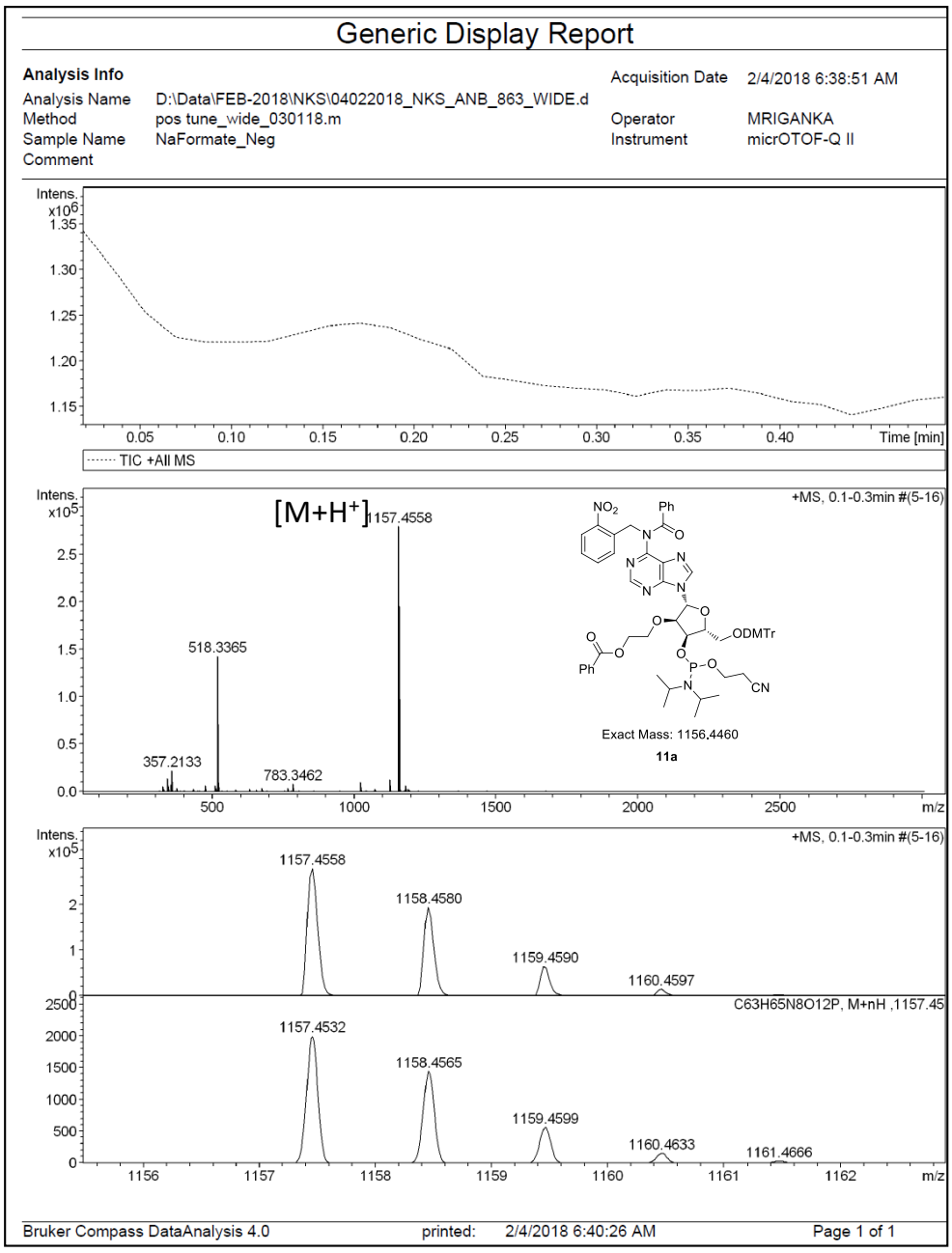


Figure S20. ESI-HRMS of NBn^{6}A phosphoramidite derivative **11a**.

11.^{1}H-/ ^{13}C-/ESI-MS/HRMS spectra of NBn^{1}A derivative, **6b**

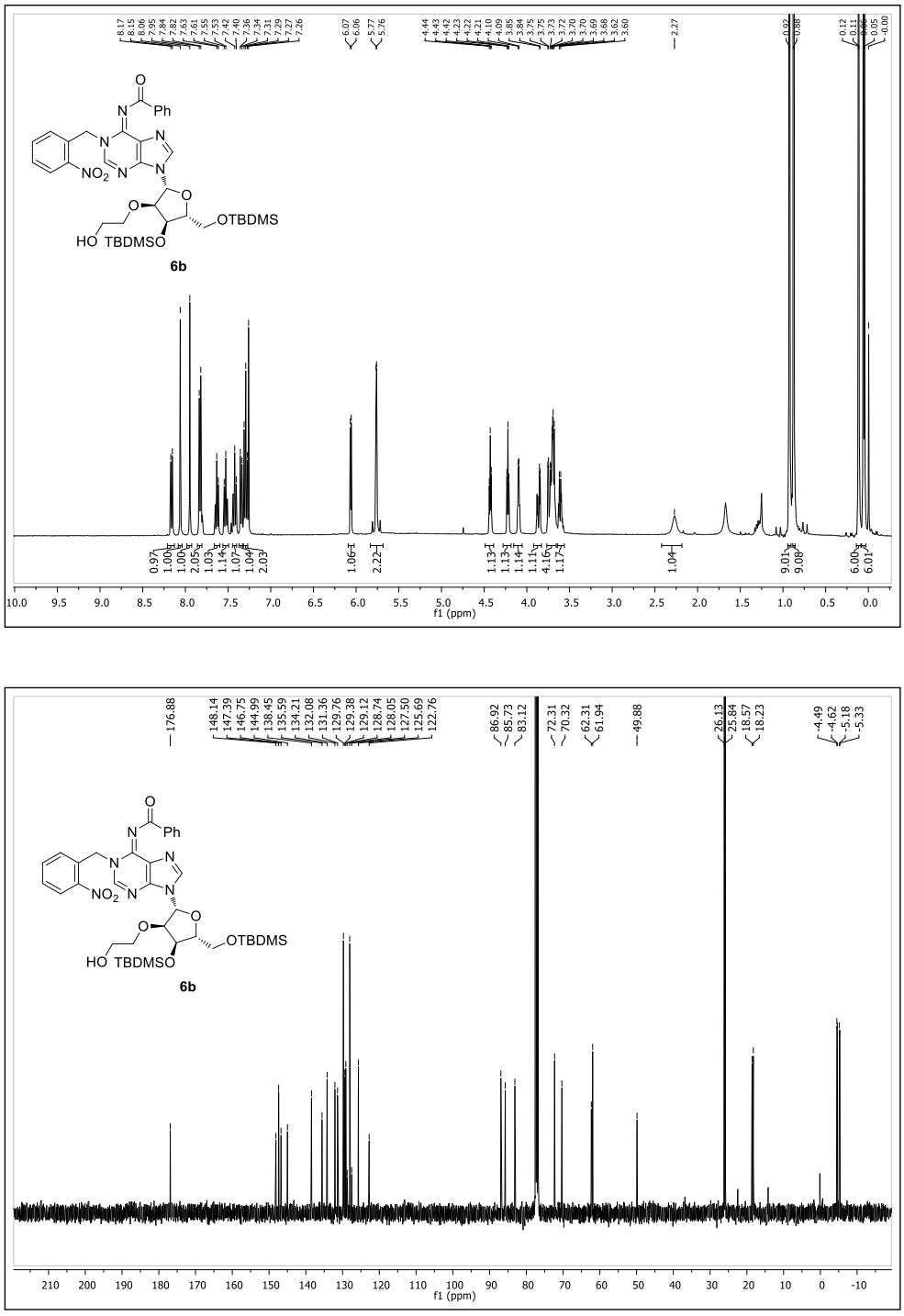


Figure S21. ^{1}H and ^{13}C NMR spectra of NBn^{1}A derivative, **6b** in $CDCl_3$.

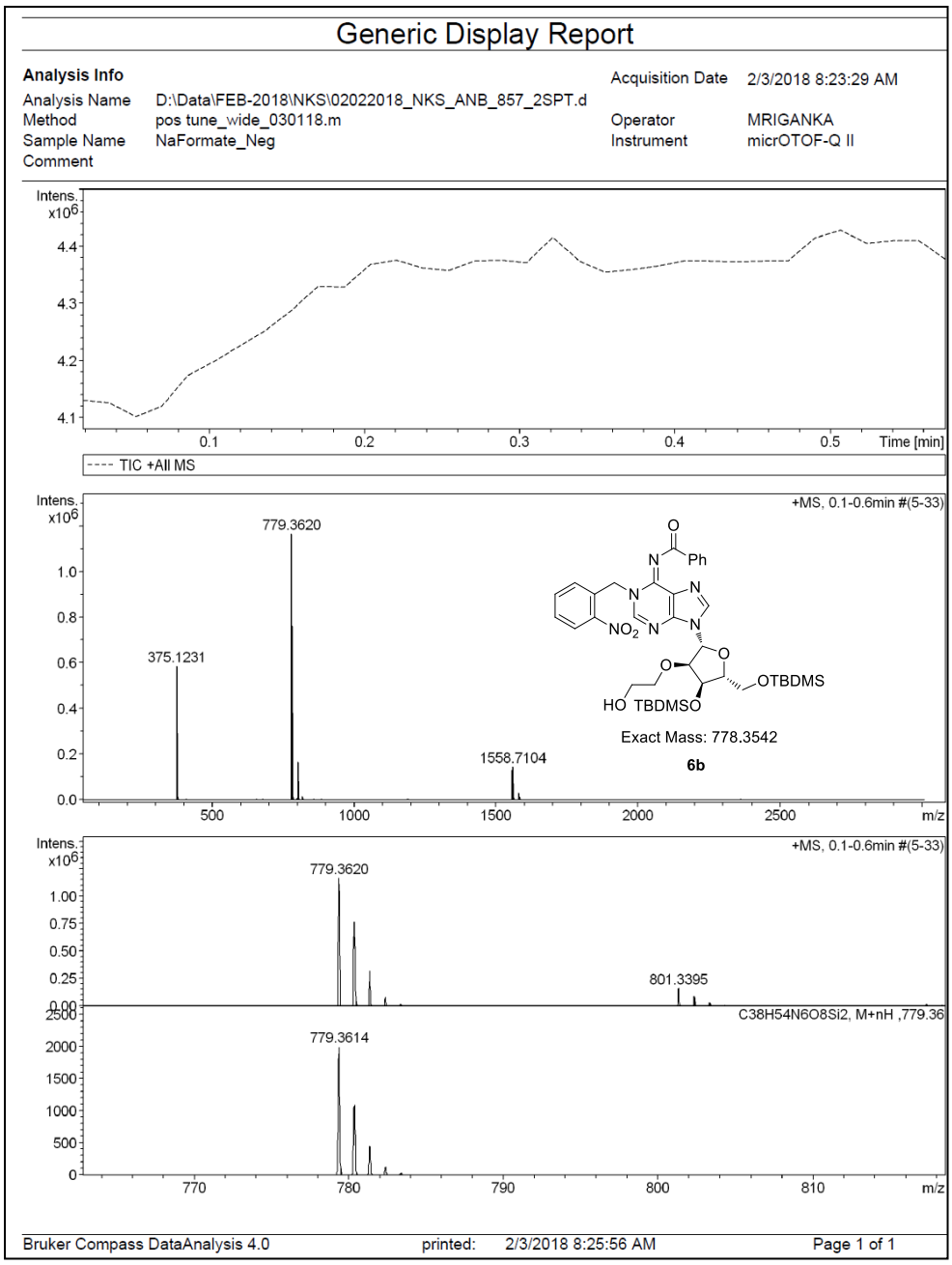


Figure S22. ESI-HRMS of NBn^{1}A derivative, **6b**.

12.^{1}H-/ ^{13}C-/ESI-MS/HRMS spectra of NBn^{1}A derivative, **7b**

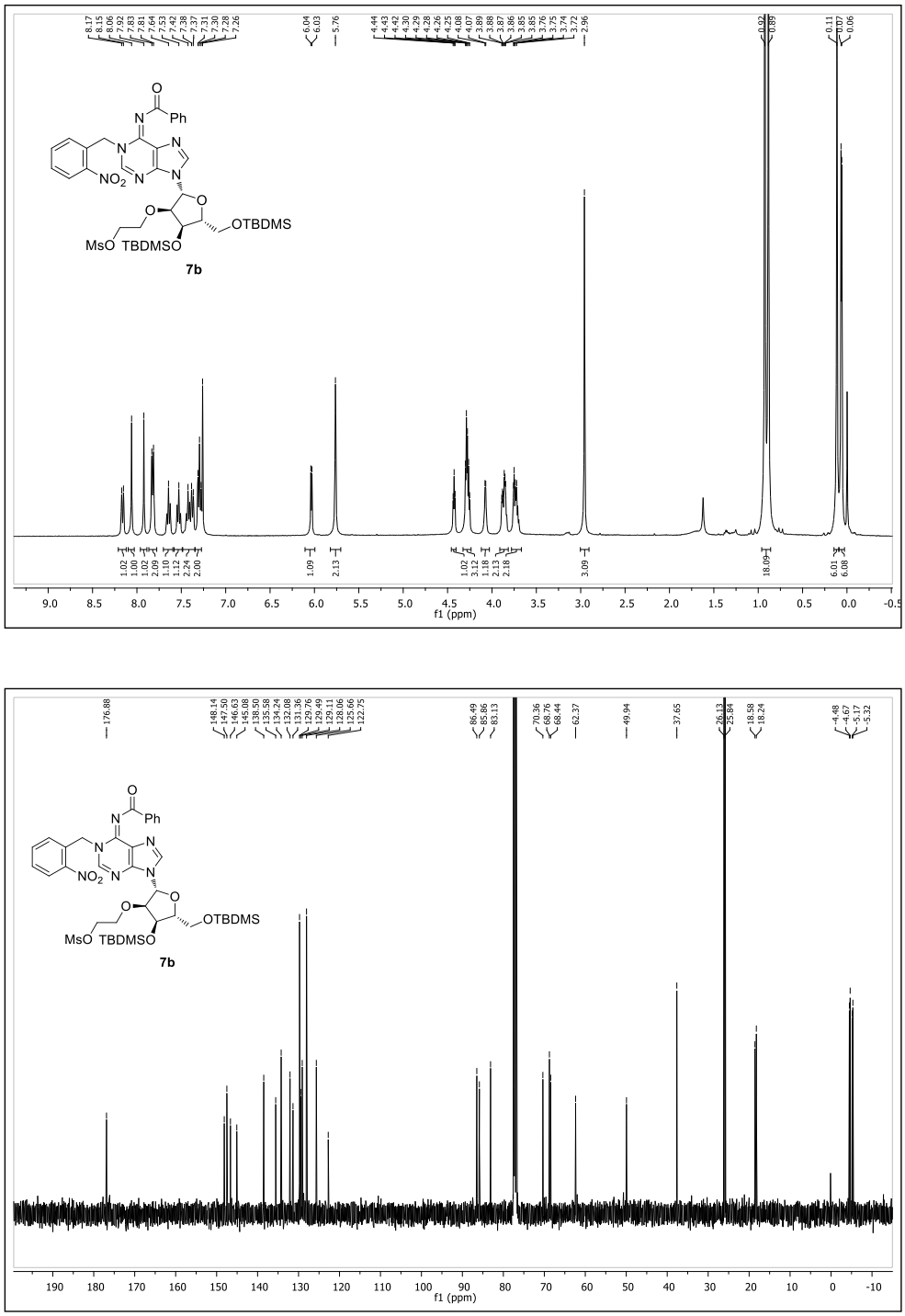


Figure S23. ^{1}H and ^{13}C NMR spectra of NBn^{1}A derivative, **7b** in $CDCl_3$.

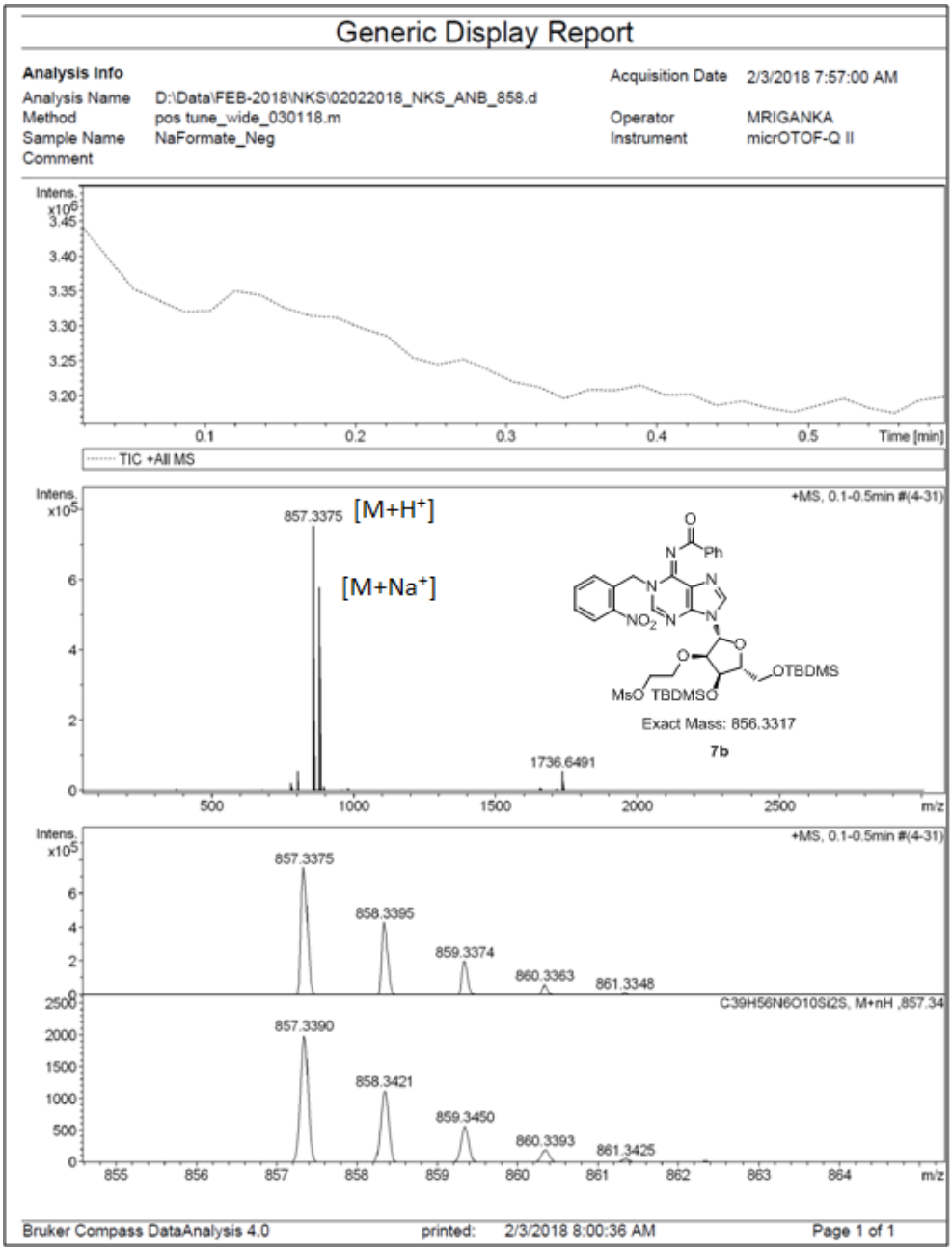


Figure S24. ESI-HRMS of NBn^{1}A derivative, **7b**.

13.^{1}H-/ ^{13}C-/ESI-MS/HRMS spectra of NBn^{1}A derivative, **8b**

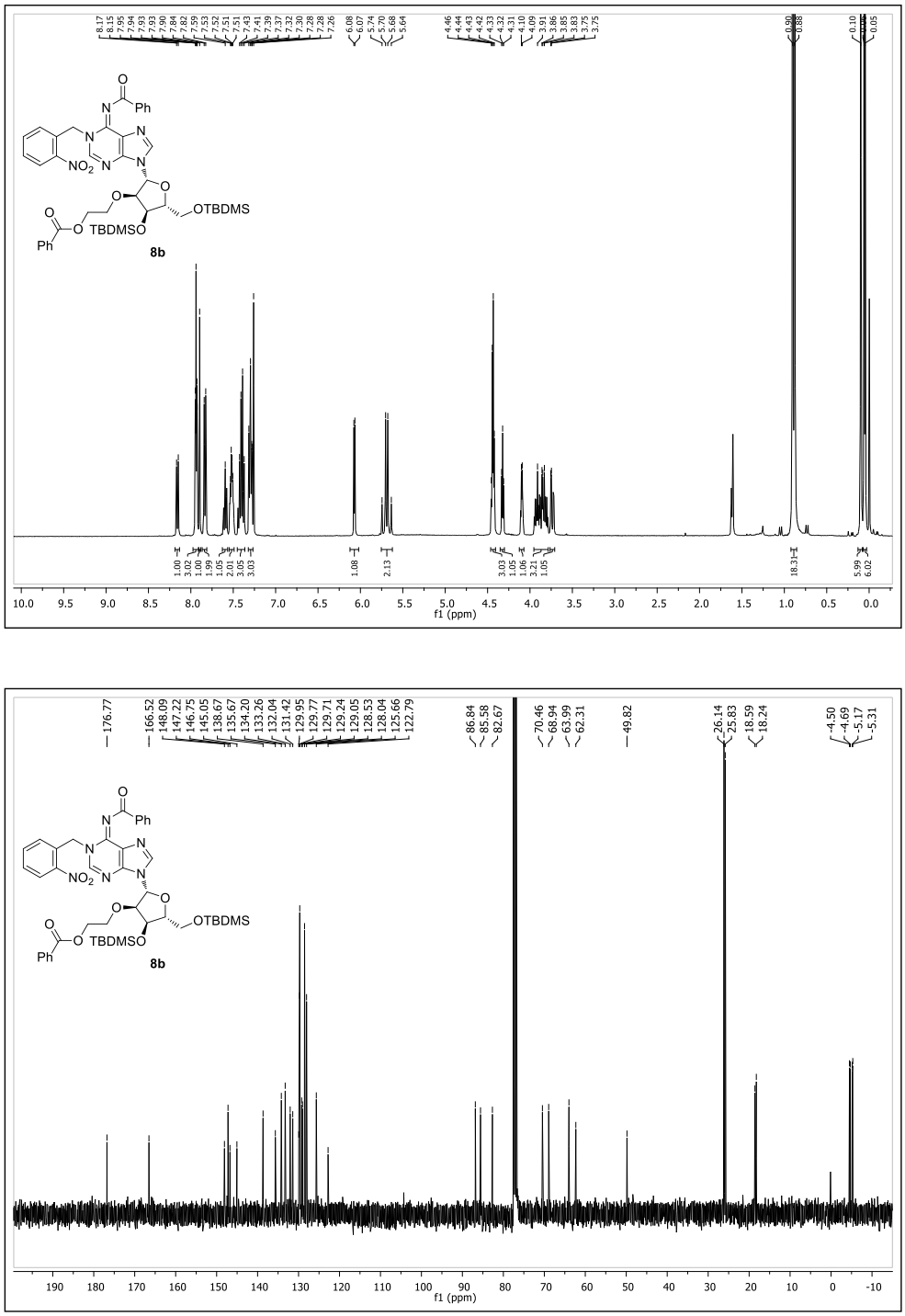


Figure S25. ^{1}H and ^{13}C NMR spectra of NBn^{1}A derivative, **8b** in $CDCl_3$.

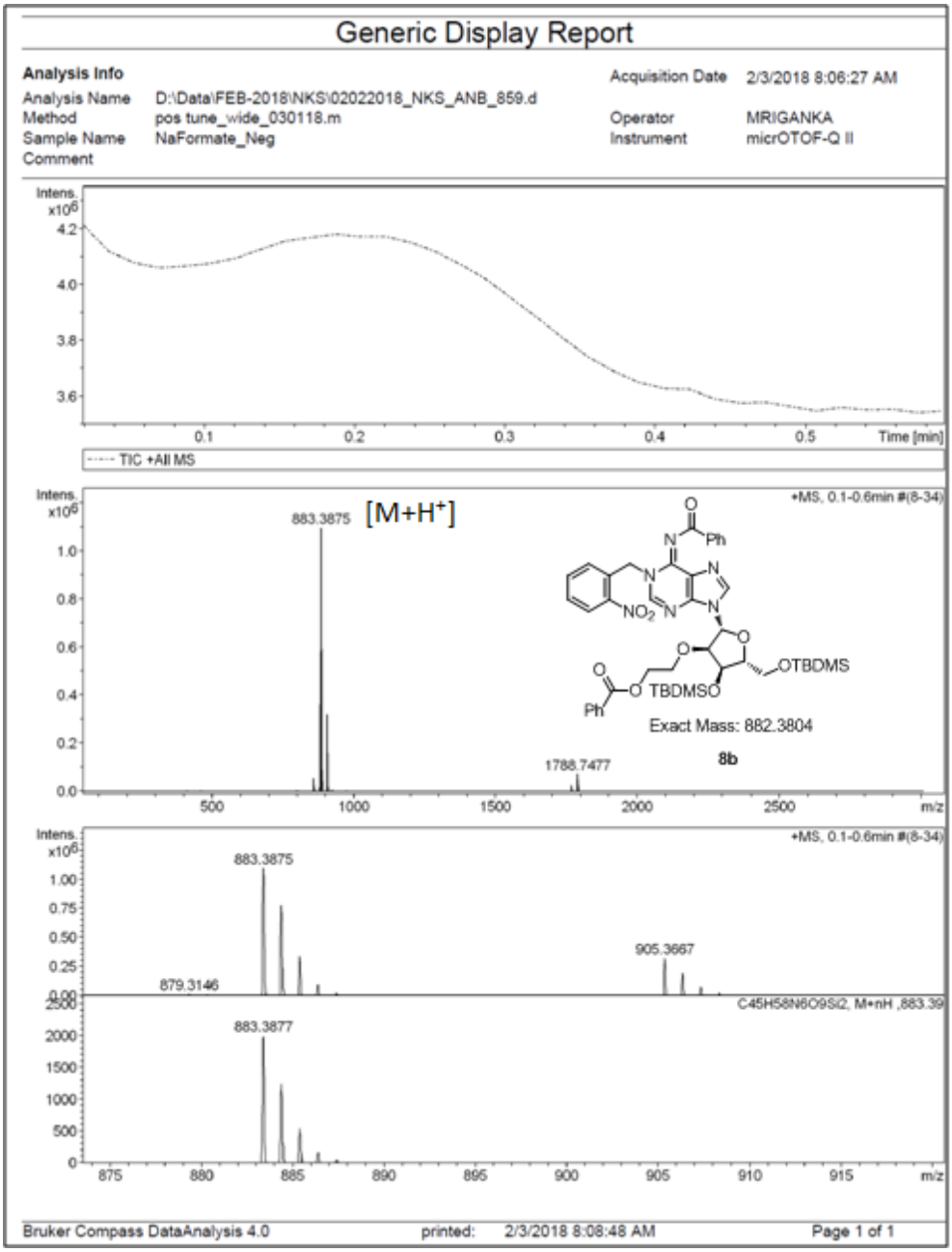


Figure S26. ESI-HRMS of NBn[1]A derivative, **8b**.

14.^{1}H-/ ^{13}C-/ESI-MS/HRMS spectra of NBn^{1}A derivative, **9b**

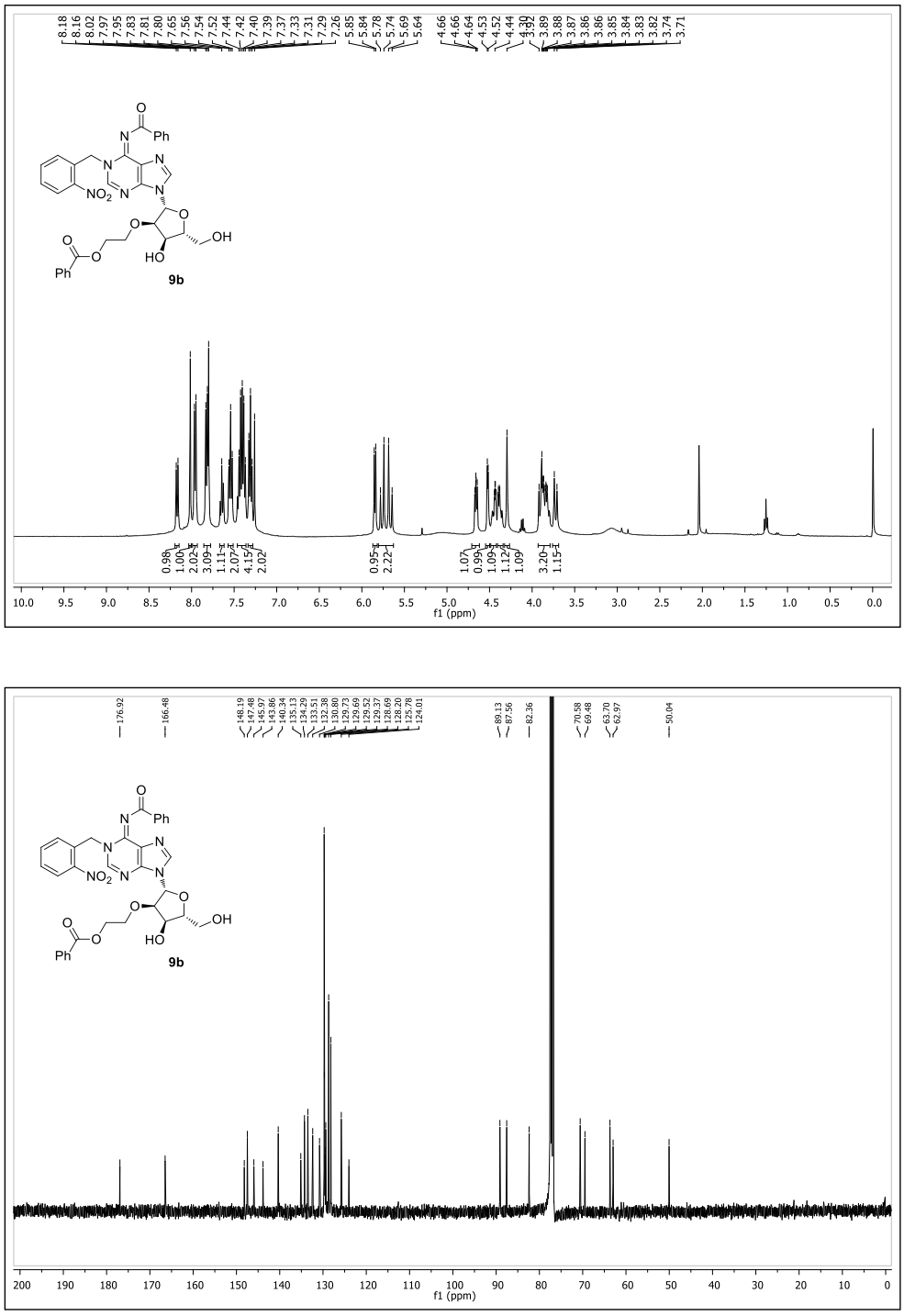


Figure S27. ^{1}H and ^{13}C NMR spectra of NBn^{1}A derivative, **9b** in $CDCl_3$.

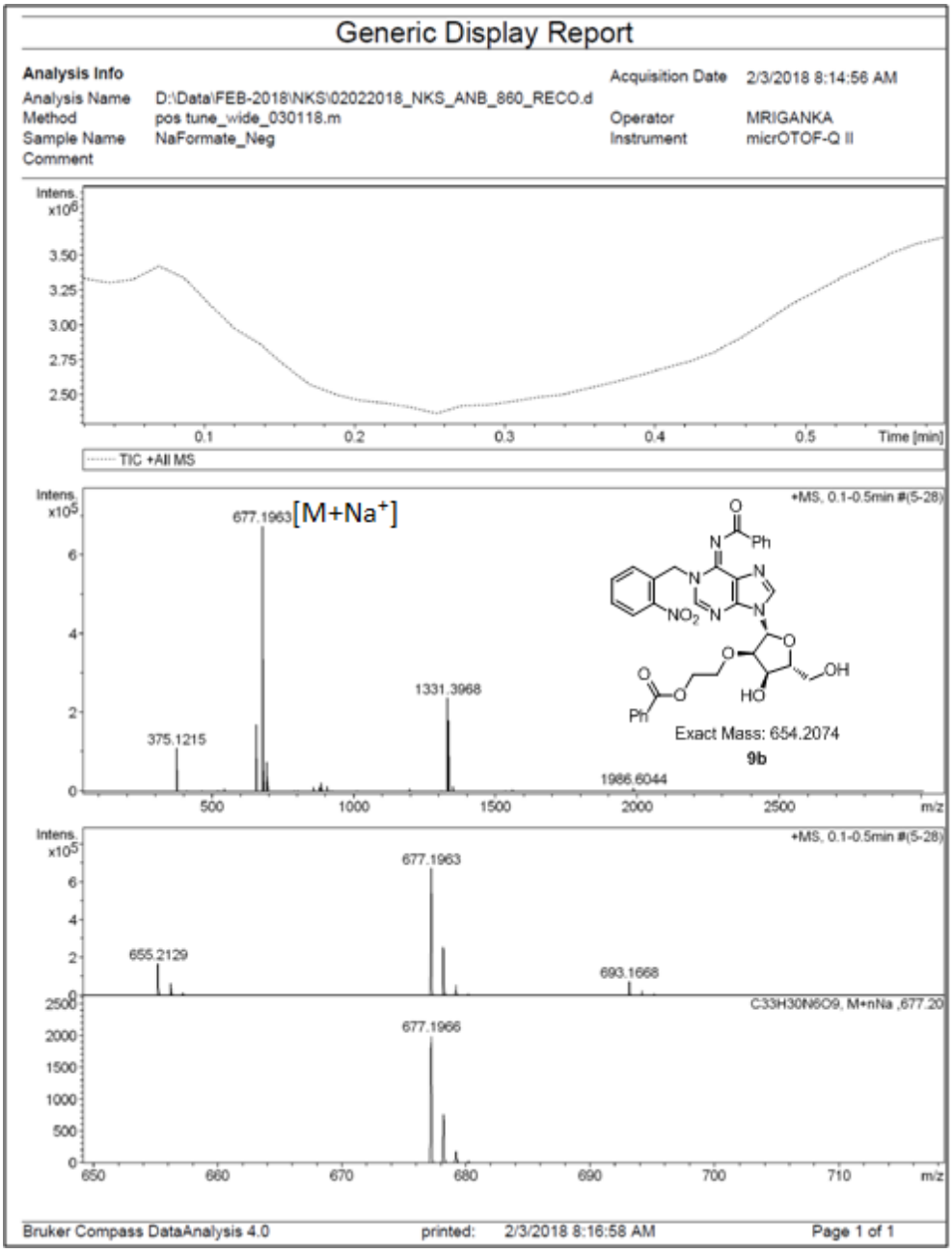


Figure S28. ESI-HRMS of NBn^{1}A derivative, **9b**.

15.^{1}H-/ ^{13}C-/ESI-MS/HRMS spectra of NBn^{1}A derivative, **10b**

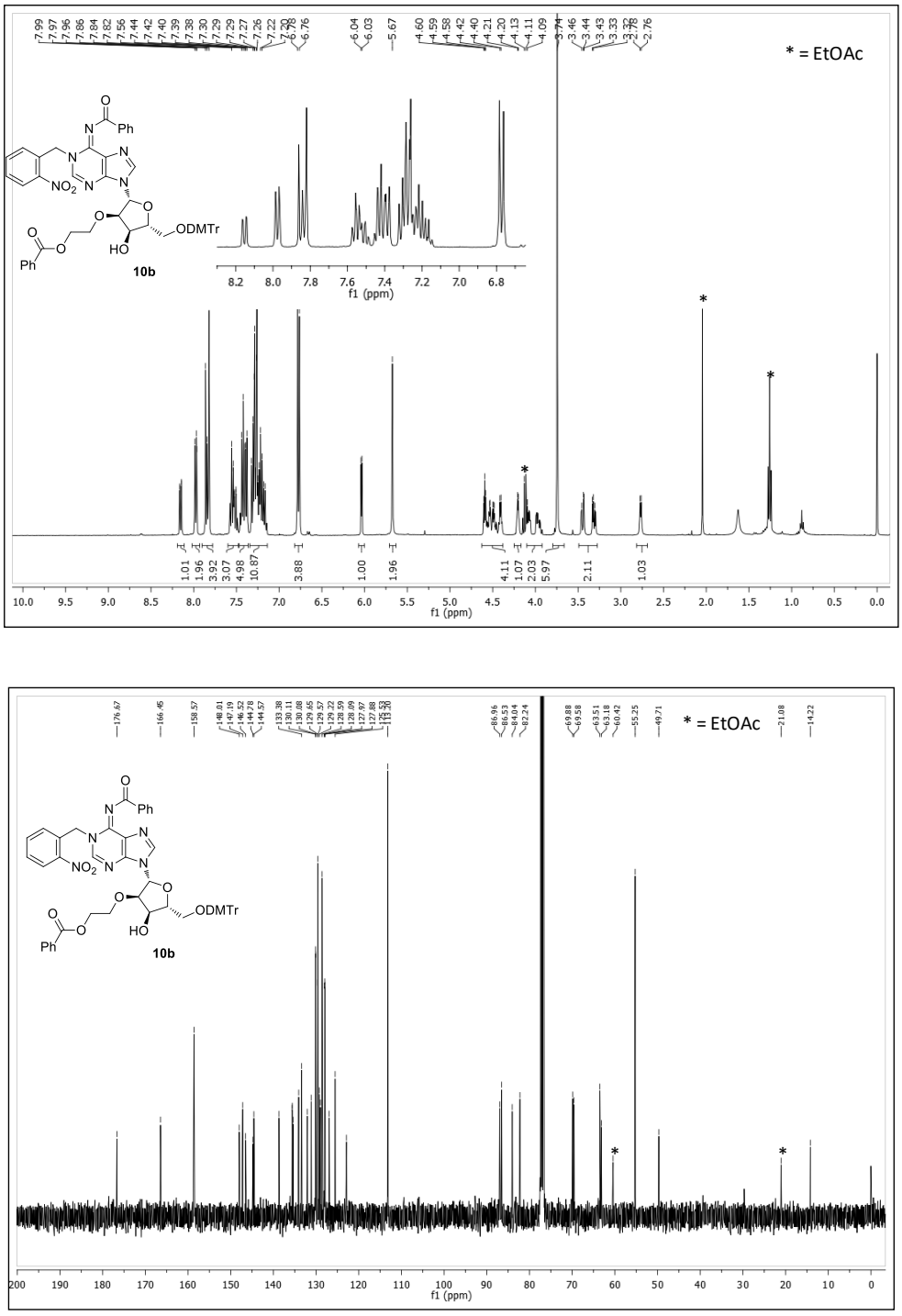


Figure S29. ^{1}H and ^{13}C NMR spectra of NBn^{1}A derivative, **10b** in $CDCl_3$.

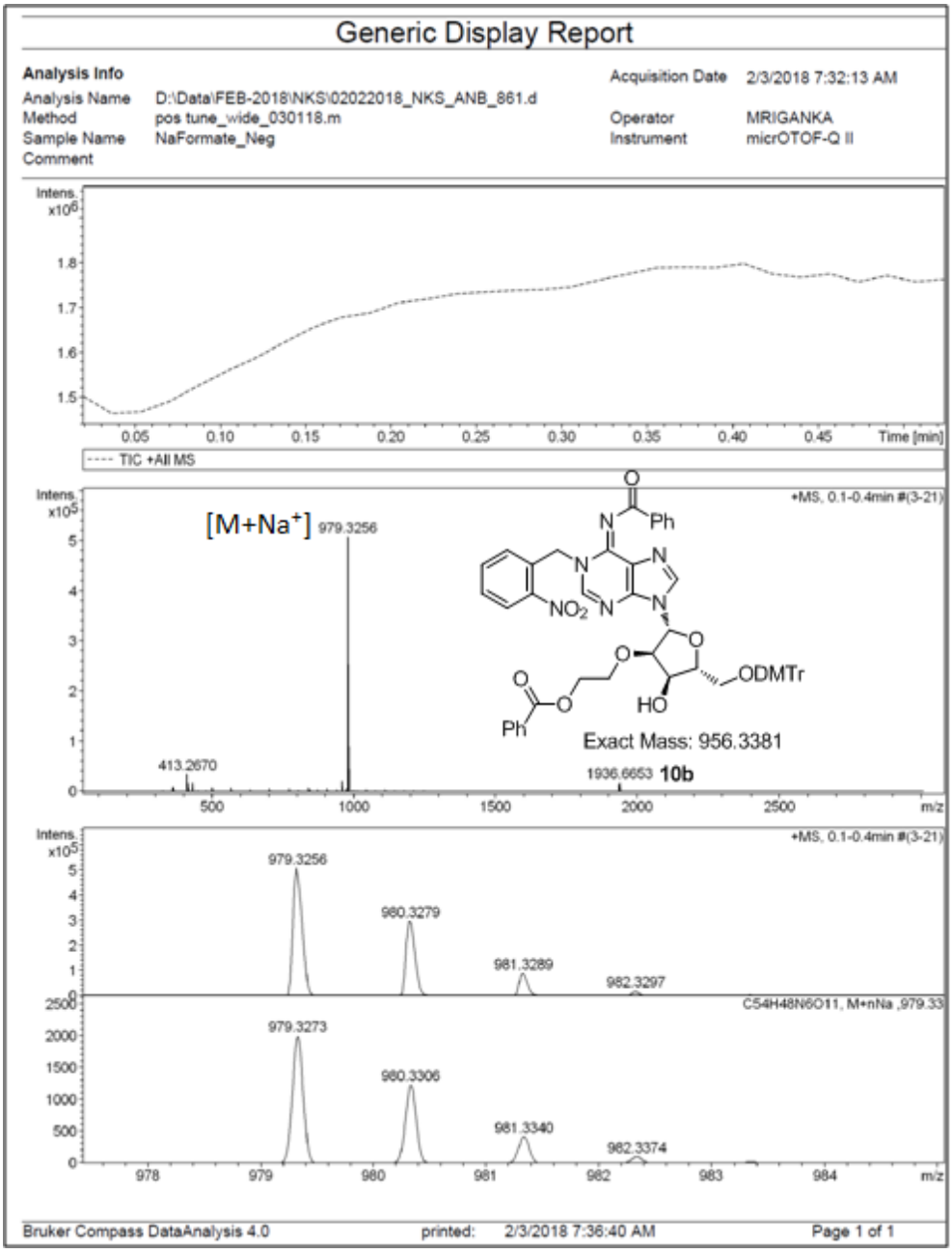


Figure S30. ESI-HRMS of 5'-ODMTr derivative of NBn^{1}A, **10b**.

16.^{1}H-/ ^{31}P-/ESI-MS/HRMS spectra of NBn^{1}A phosphoramidite, **11b**

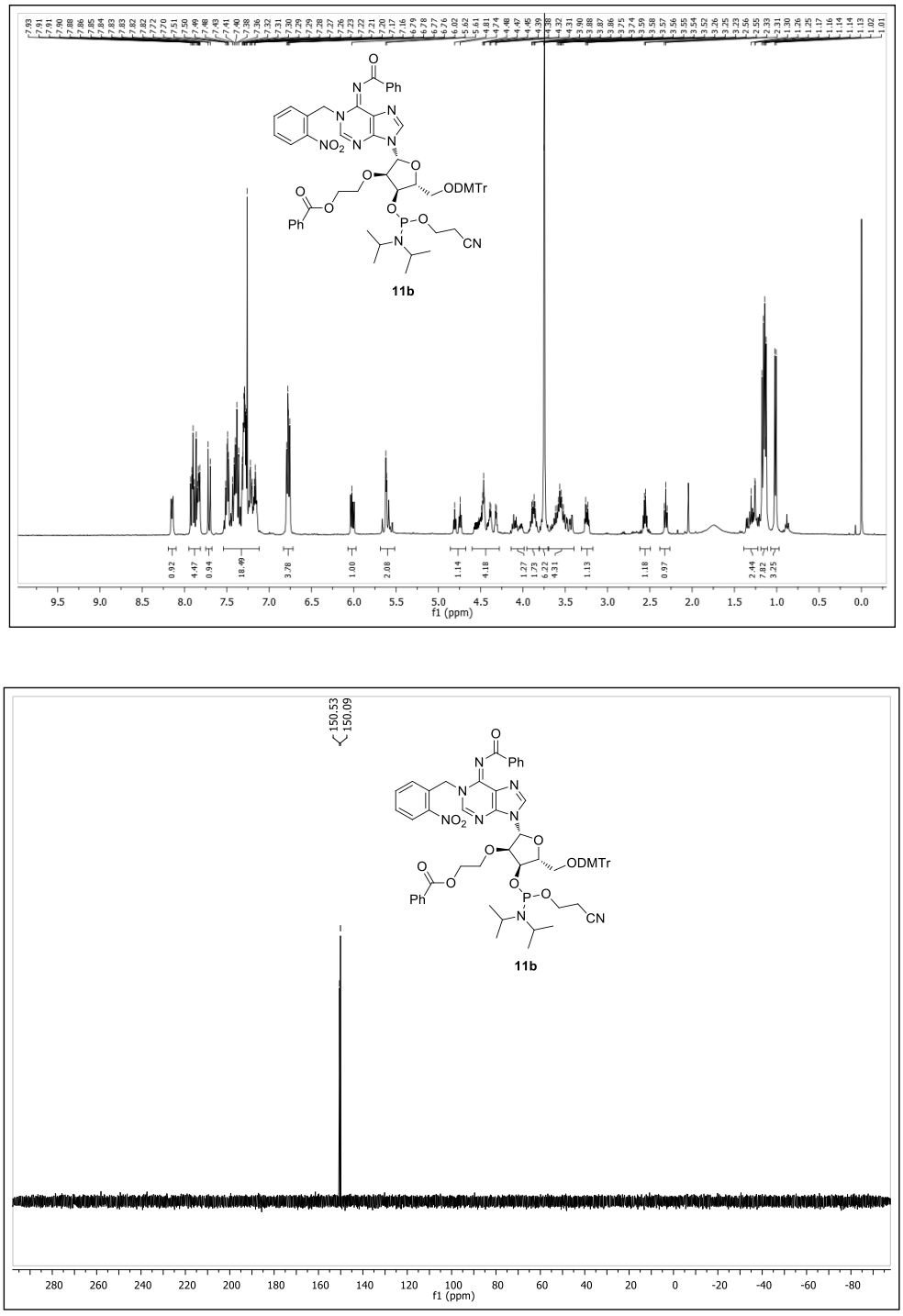


Figure S31. ^{1}H and ^{31}P NMR spectra of NBn^{1}A phosphoramidite derivative, **11b** in $CDCl_3$.

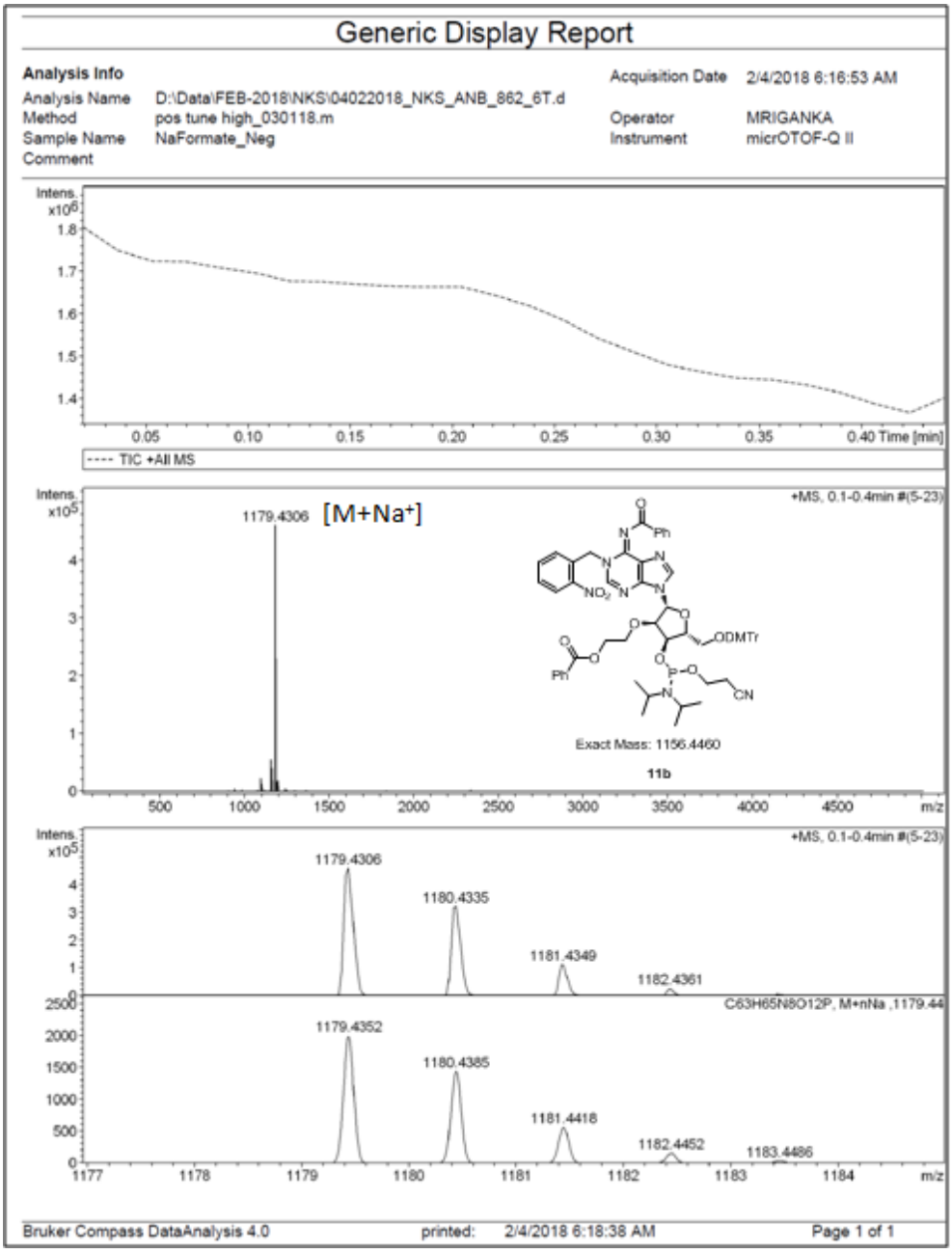


Figure S32. ESI-HRMS of NBn^{1}A phosphoramidite derivative **11b**.

17.^{1}H-/ ^{13}C- spectra of Compound **13**

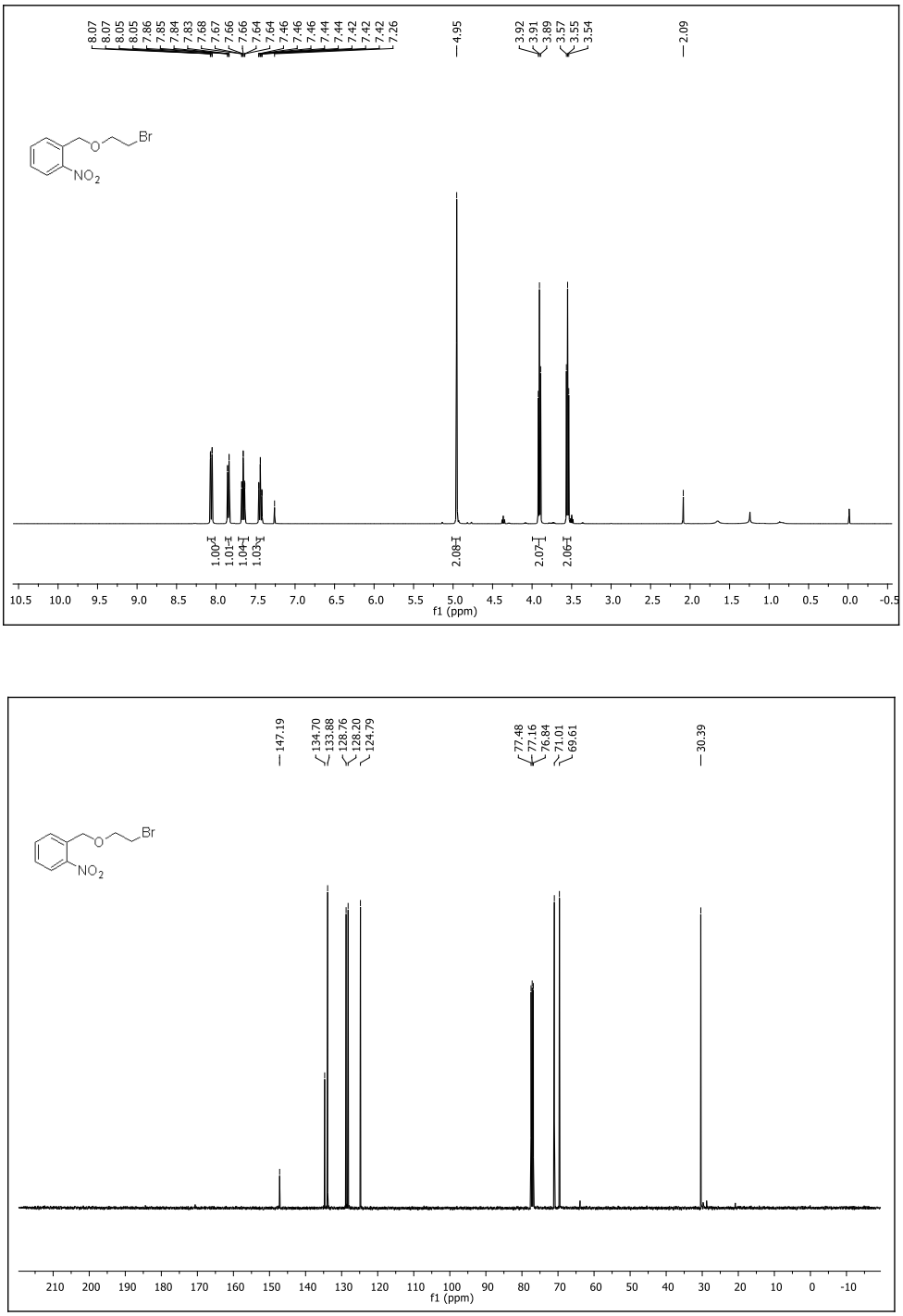


Figure S33. ^{1}H and ^{13}C NMR spectra of compound, **13** in $CDCl_3$.

18.^{1}H-/ ^{13}C-/ESI-MS/HRMS spectra of A-NBnOE nucleoside **14**

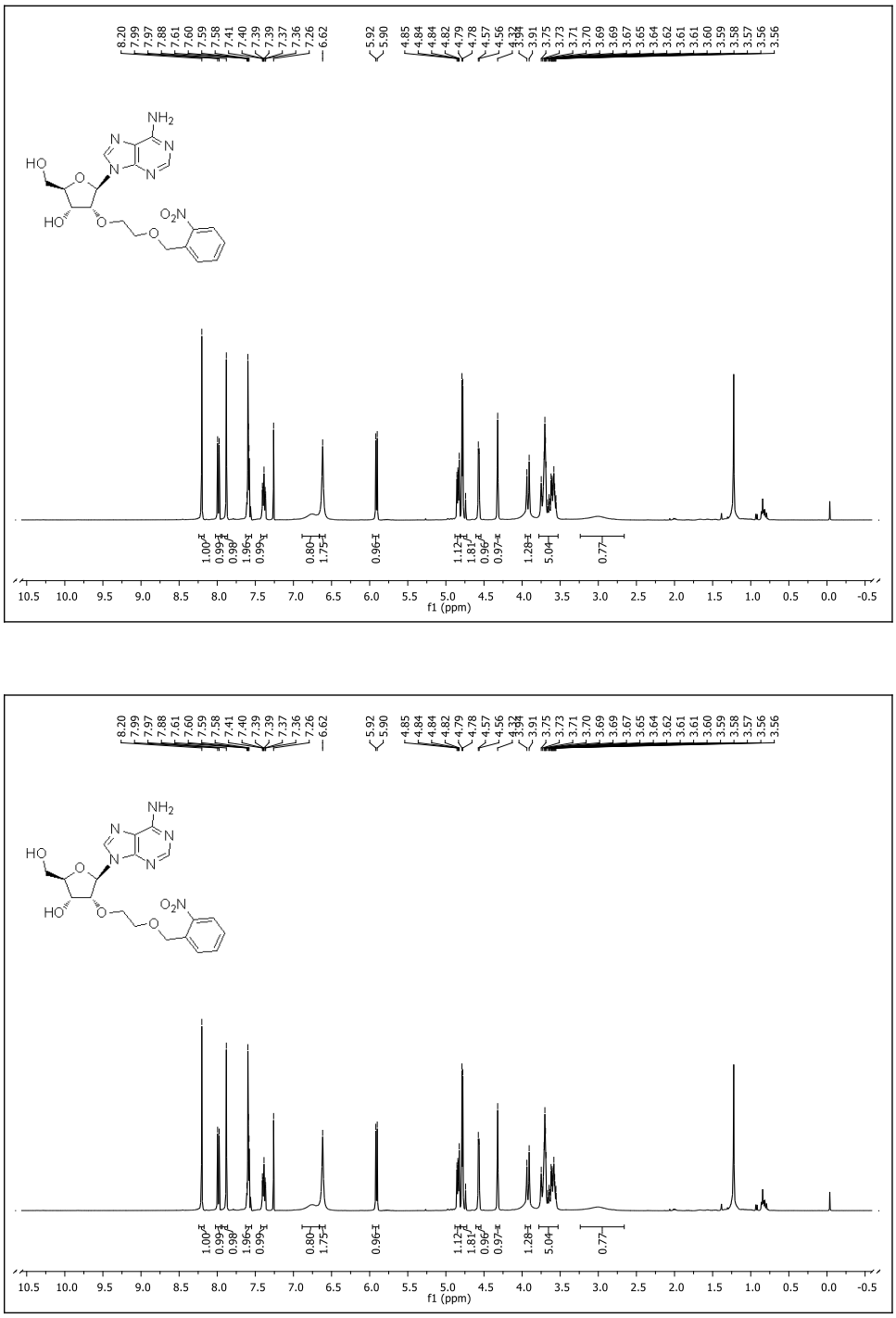


Figure S34. ^{1}H and ^{13}C NMR spectra of A-NBnOE nucleoside **14** in $CDCl_3$.

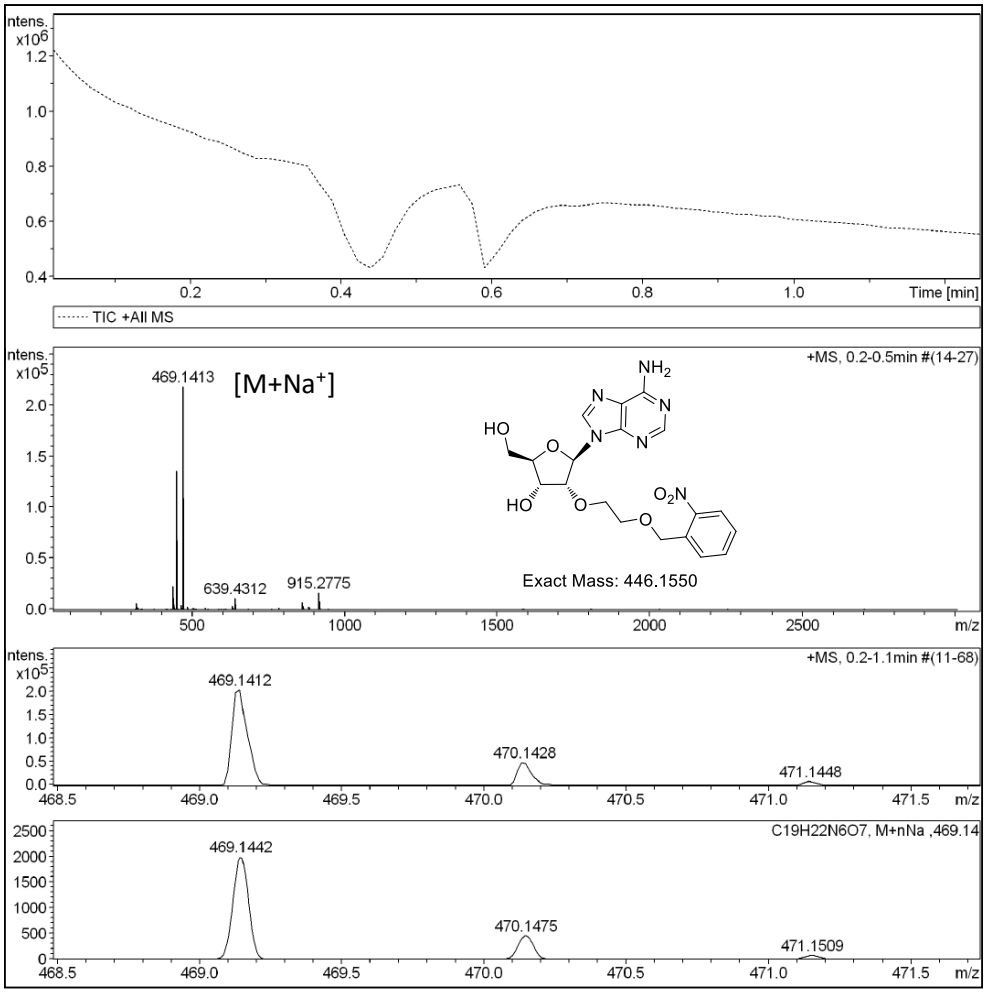


Figure S35. ESI-HRMS of A-NBnOE nucleoside **14**.

19. ^{1}H-/ ^{13}C-/ESI-MS/HRMS spectra of A-NBnOE derivative, **15**

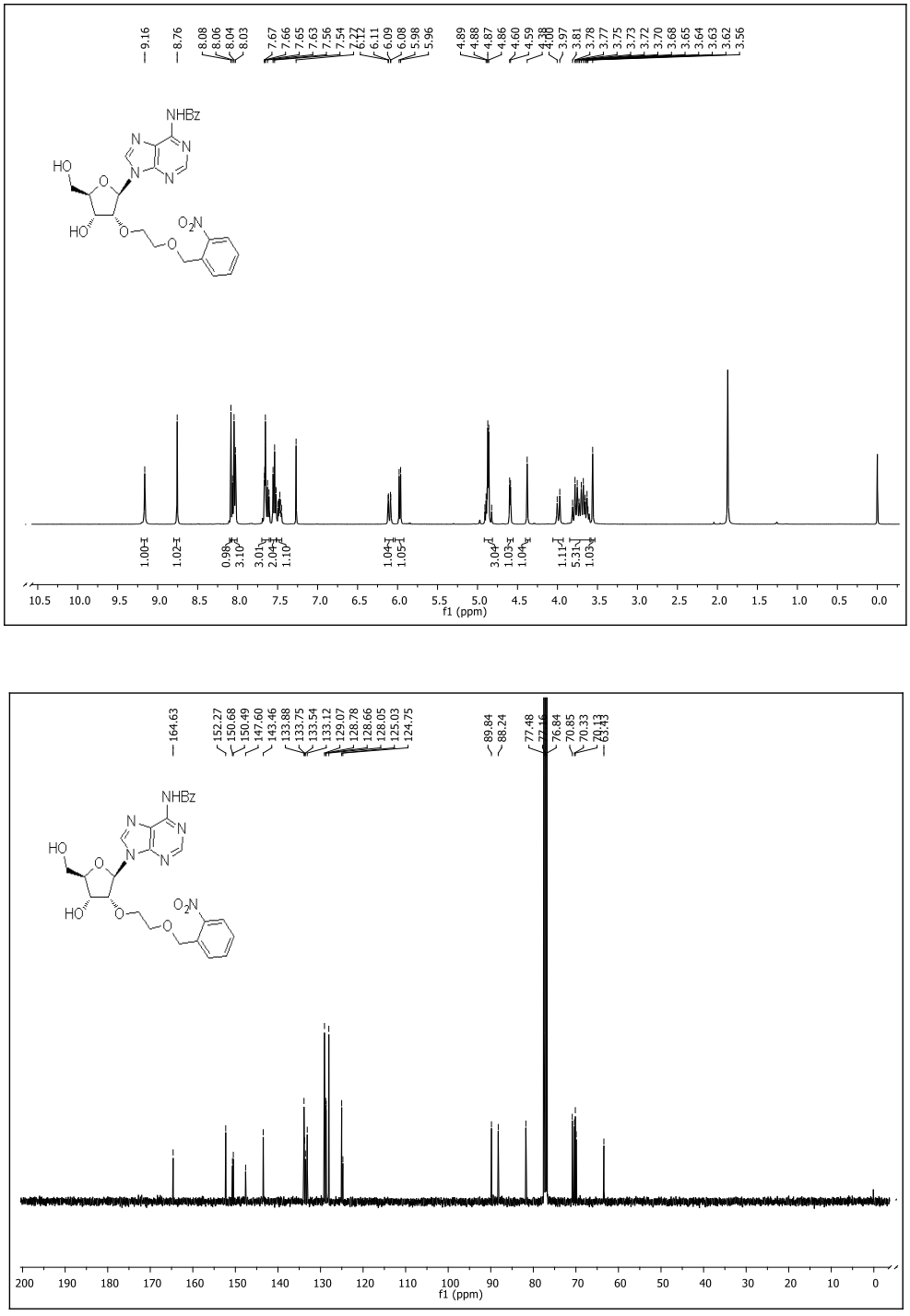


Figure S36. ^{1}H and ^{13}C NMR spectra of A-NBnOE nucleoside, **15** in $CDCl_3$

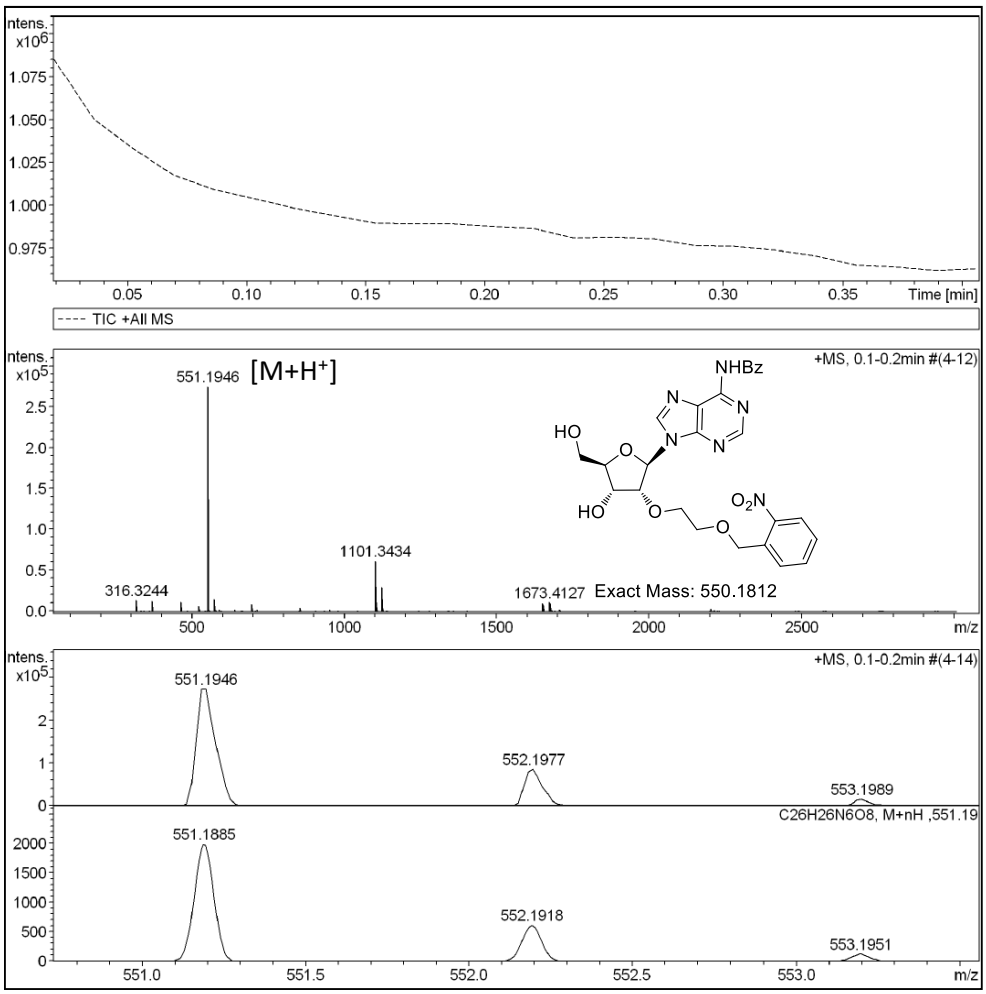


Figure S37. ESI-HRMS of A-NBnOE nucleoside, **15**.

20.^{1}H-/ ^{13}C-/ESI-MS/HRMS spectra of A-NBnOE derivative, **16**

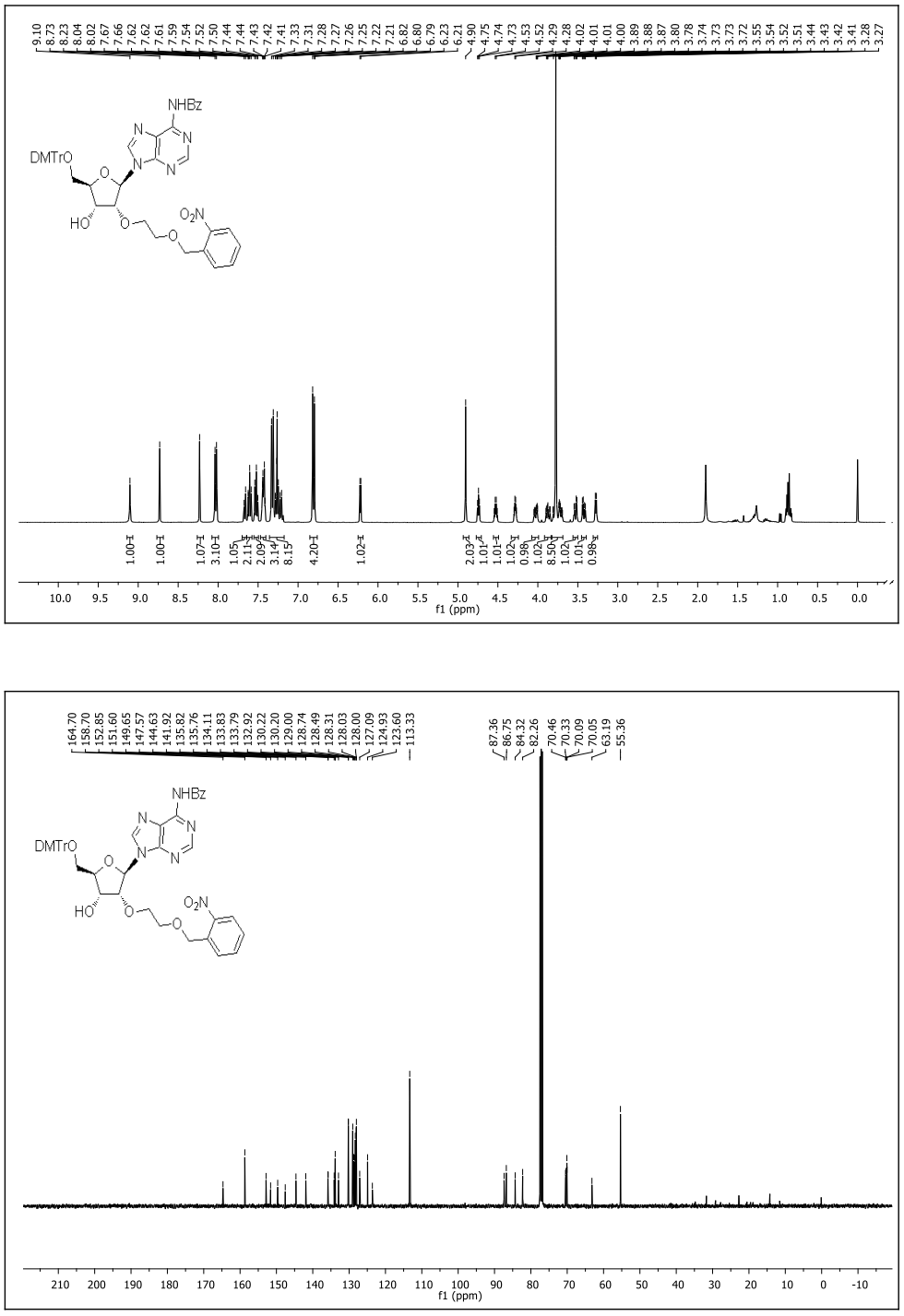


Figure S38. ^{1}H and ^{13}C NMR spectra of A-NBnOE nucleoside derivative, **16,** in $CDCl_3$.

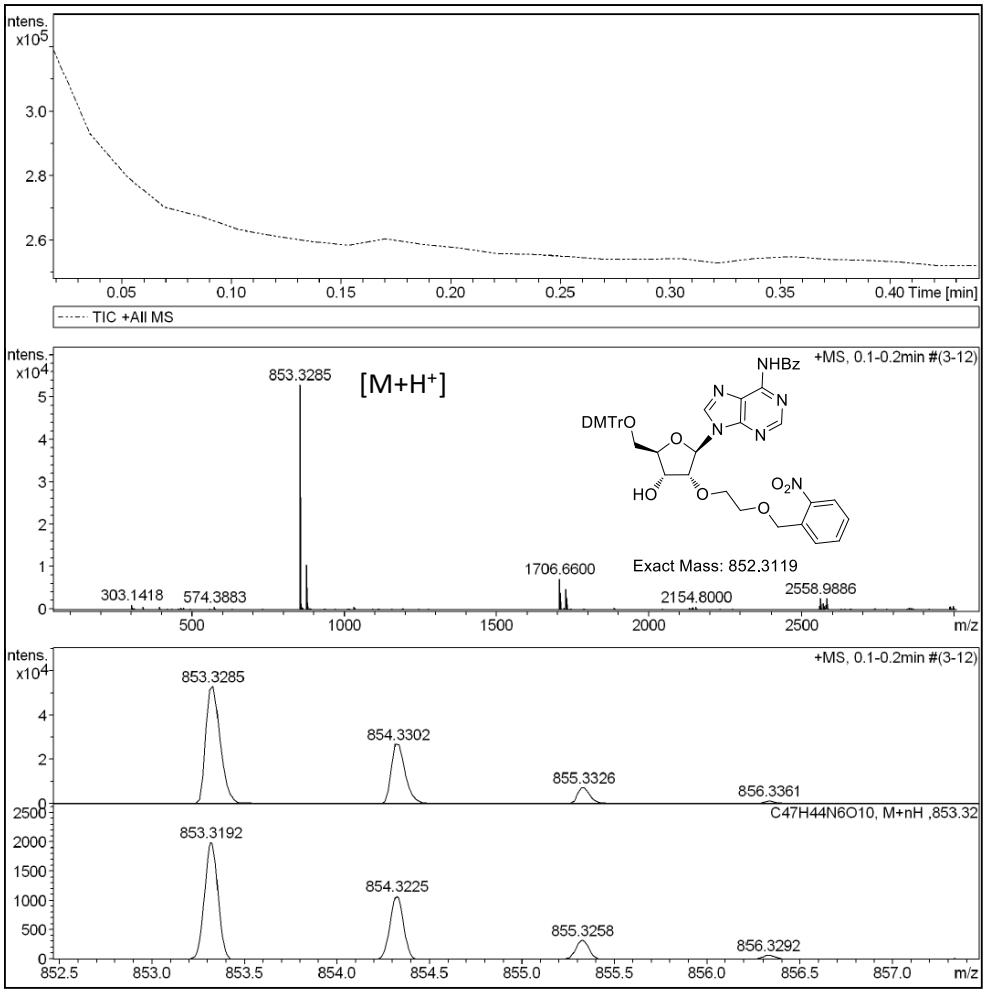


Figure S39. ESI-HRMS of A-NBnOE nucleoside derivative, **16**

21.^{1}H-/ ^{13}C-/^{31}P-/ESI-MS/HRMS spectra of A-NBnOE phosphoramidite, **16**

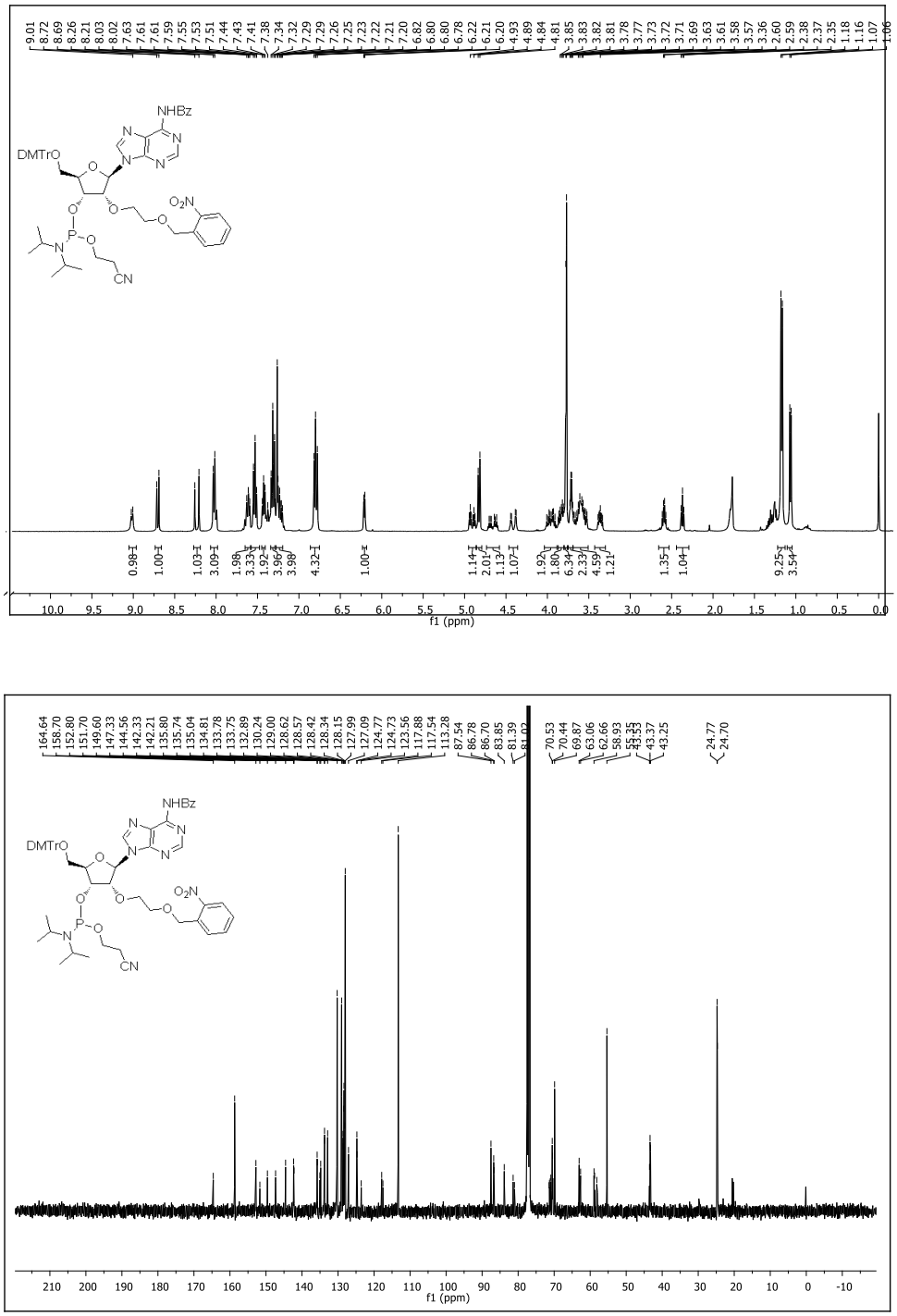


Figure S40. ^{1}H and ^{13}C NMR spectra of A-NBnOE nucleoside derivative, **17** in $CDCl_3$.

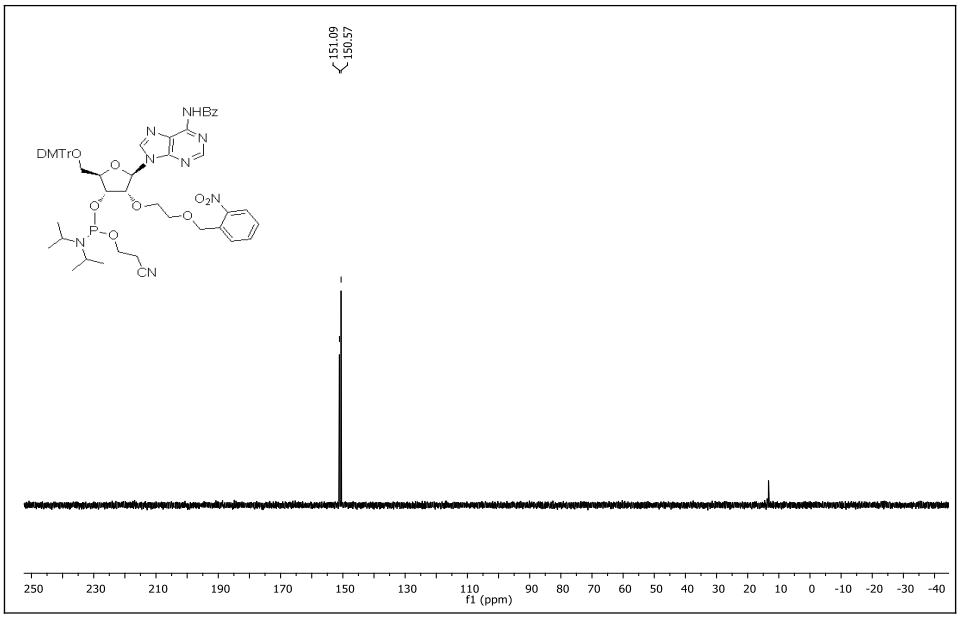


Figure S41. ^{31}P NMR spectra of A-NBnOE nucleoside phosphoramidite, **17** in $CDCl_3$.

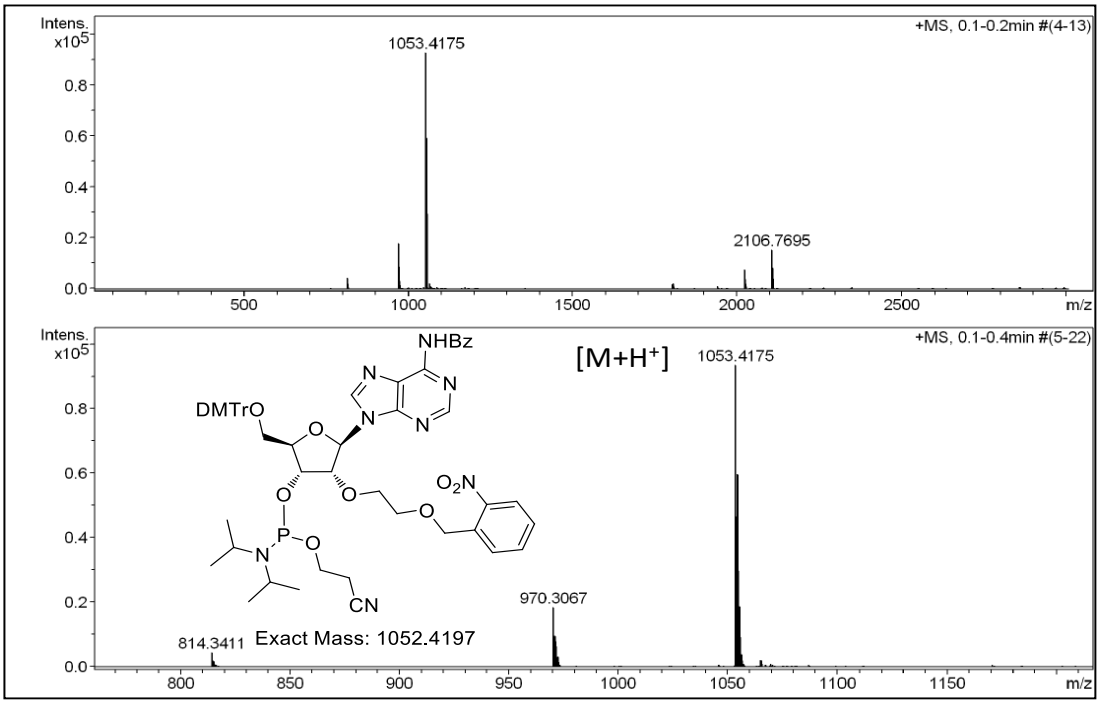


Figure S42. ESI-MS of A-NBnOE nucleoside phosphoramidite, **17**

22.Stacked ^{1}H-/^{13}C NMR spectra of **6a/6b** for comparative understandings in $CDCl_3$

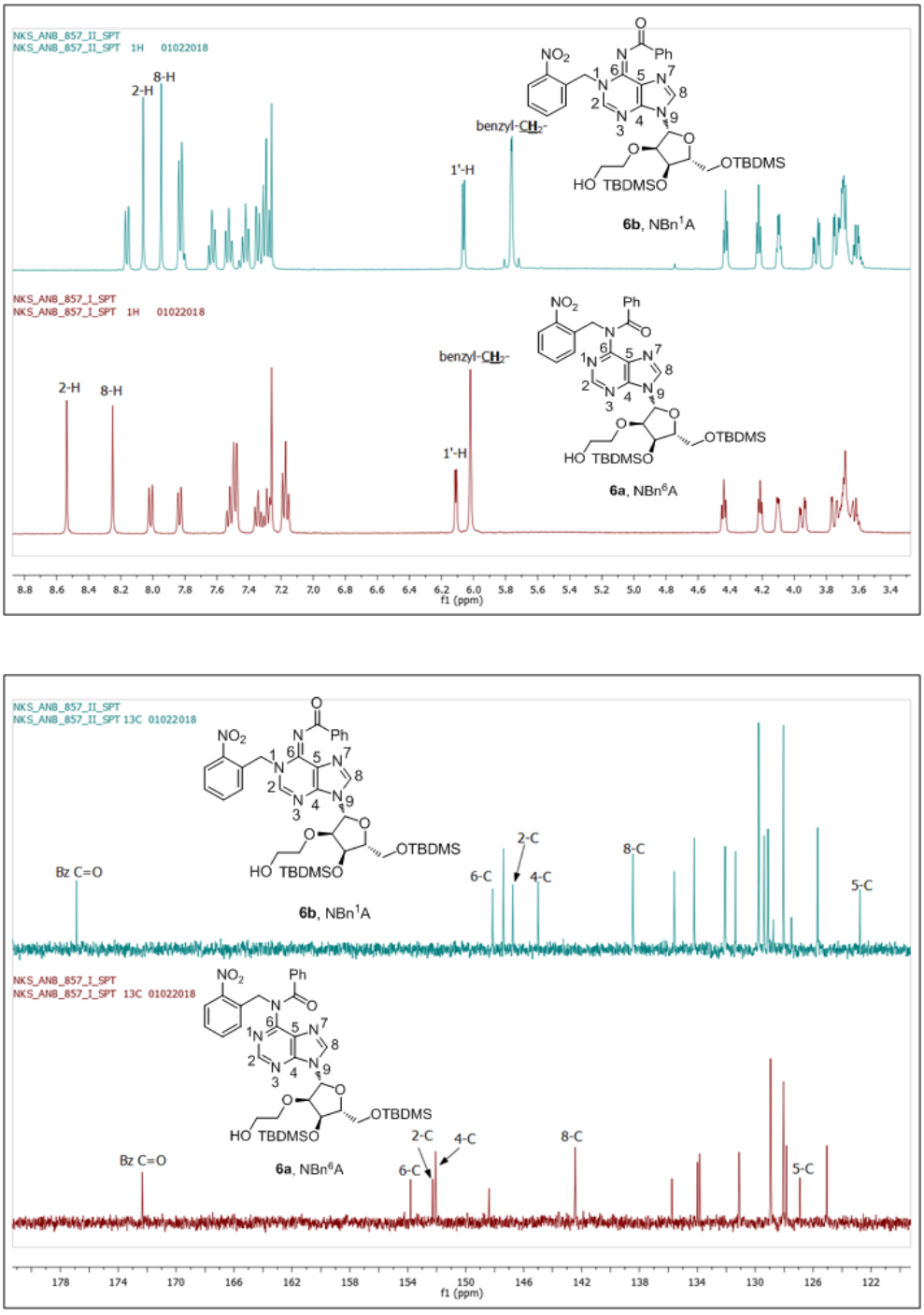


Figure S43. ^{1}H-/^{13}C NMR stacked spectra of **6a/6b** and their chemical shift assignment of important peaks for comparative understandings in $CDCl_3$.

23.D_2O exchange ^1H NMR and photo cleavage studies at 365 nm for nucleoside **9a**

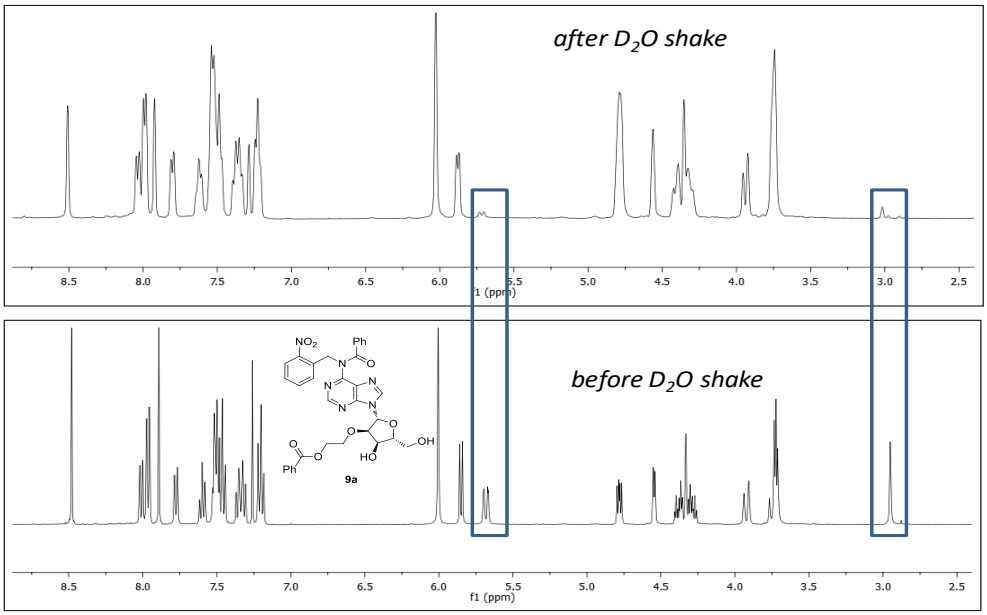


Figure S44. D_2O exchange ^1H NMR spectra of nucleoside **9a**, exchangeable 3'-OH at δ5.7 ppm, and 5'-OH at δ2.9 ppm before and after D_2O shake are framed.

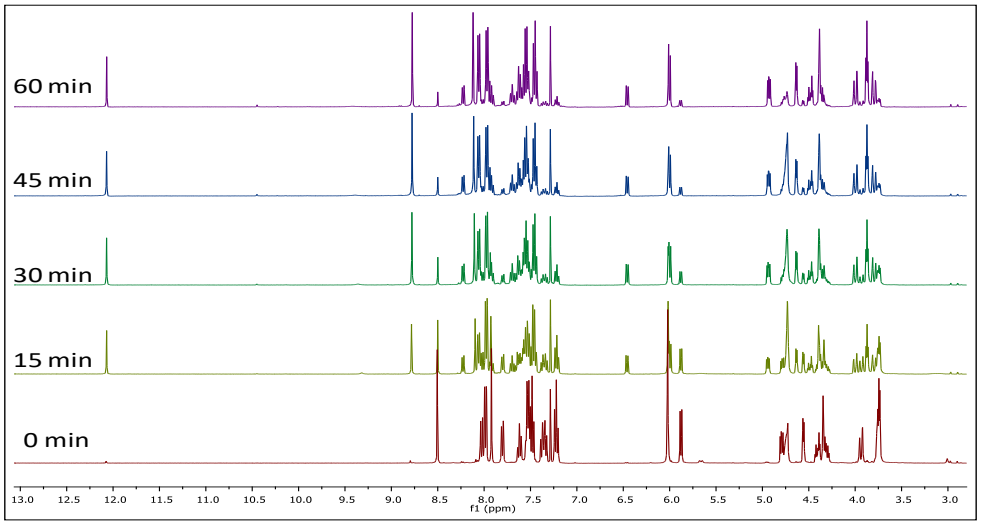


Figure S45. 1H-NMR spectral array of compound **9a** at different time intervals irradiated with 365 nm in $CDCl_3$, peak at δ12.1 ppm belongs to aldehydic proton of photoproduct, 2-nitrosobenzaldehyde.

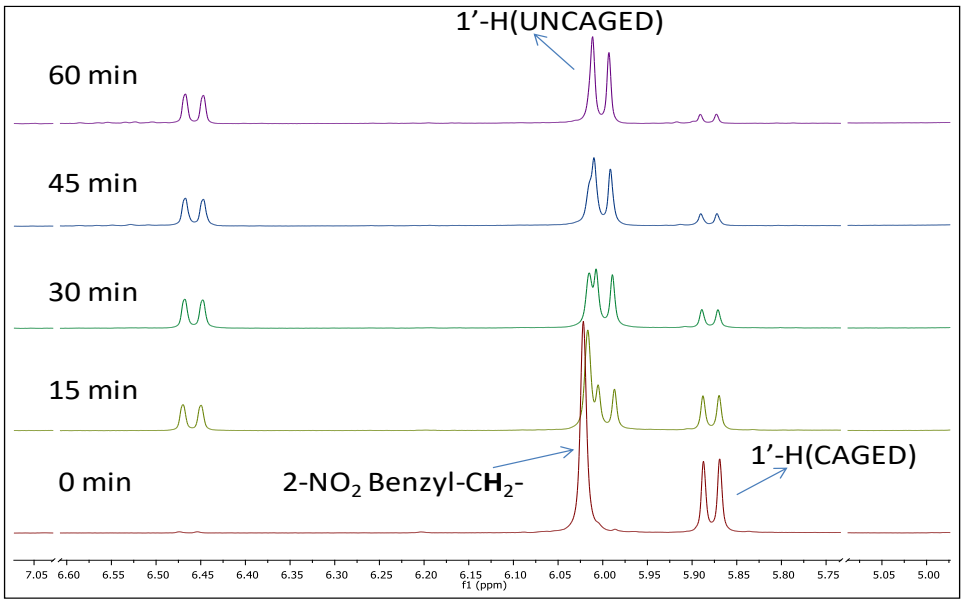


Figure S46. Expanded portion of ^{1}H-NMR spectra shown in Figure XX. These spectra indicate the photo cleavage of compound **9a** with increasing exposure time with 365 nm UV light in $CDCl_3$.

24.D_2O exchange ^1H NMR and photo cleavage studies at 365 nm for nucleoside **9b**

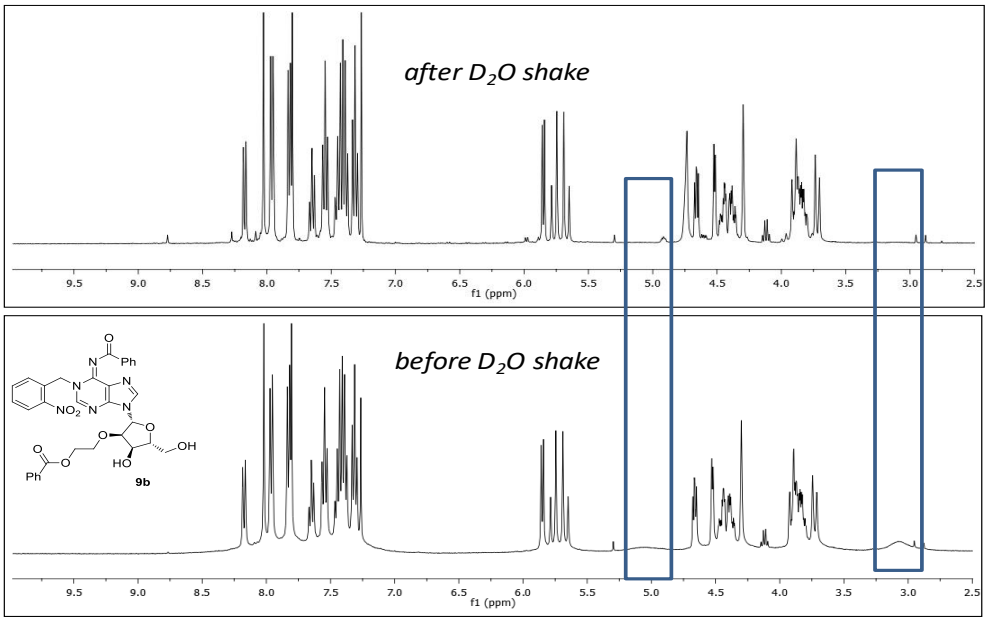


Figure S47. D_2O exchange ^1H NMR spectra of nucleoside **9b**, exchangeable 3'-OH at δ5.1 ppm, and 5'-OH at δ3.1 ppm before and after D_2O shake are framed.

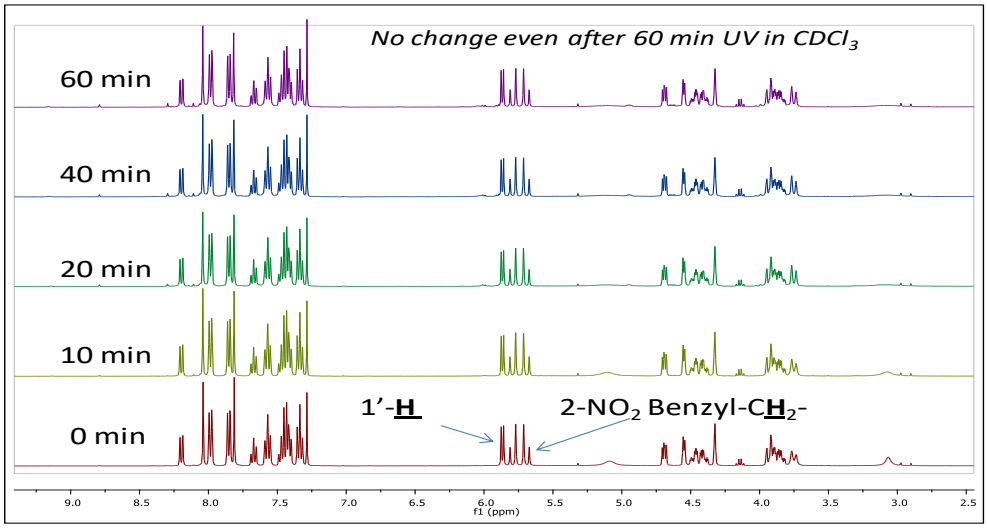


Figure S48. 1H-NMR spectral array of compound **9b** at different time intervals irradiated with 365 nm in $CDCl_3$, indicating no change in ^1H-NMR spectra even after irradiating for 60 min. (no photo cleavage)

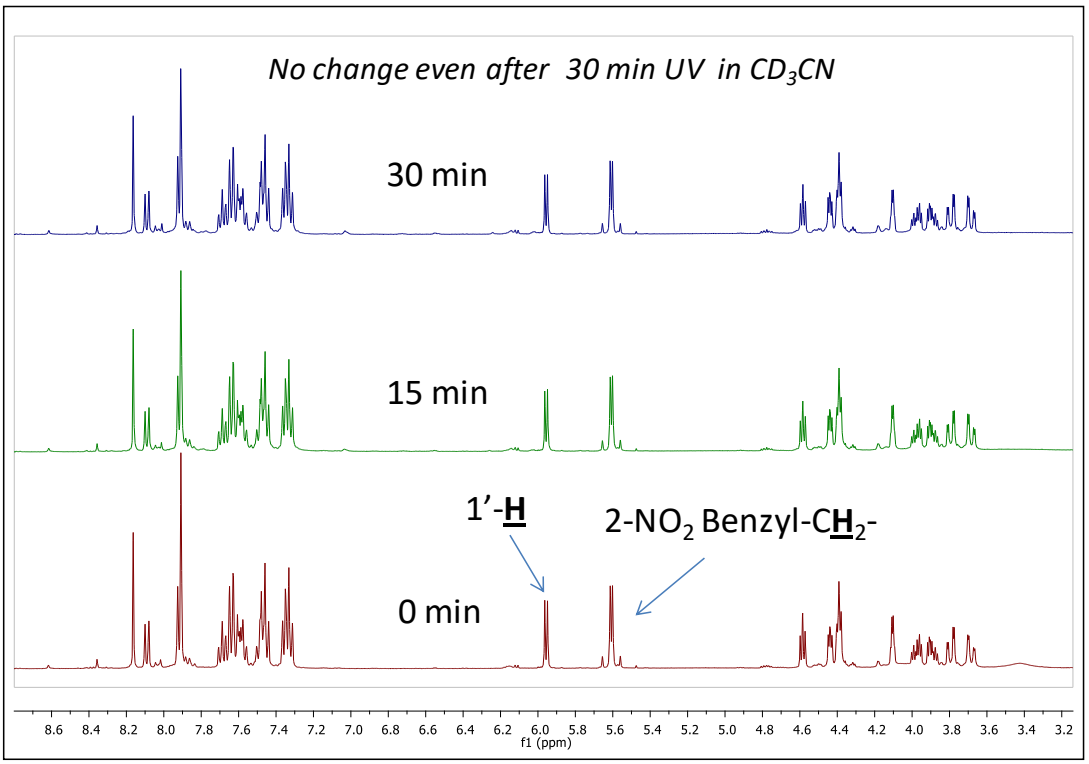


Figure S49. 1H-NMR spectral array of compound **9b** at different time intervals irradiated with 365 nm in CD_3CN, indicating no change in ^{1}H-NMR spectra after 30 min of irradiation. (no photo cleavage)

25.ESI-MS of Compound **9b** irradiated with 365 nm.

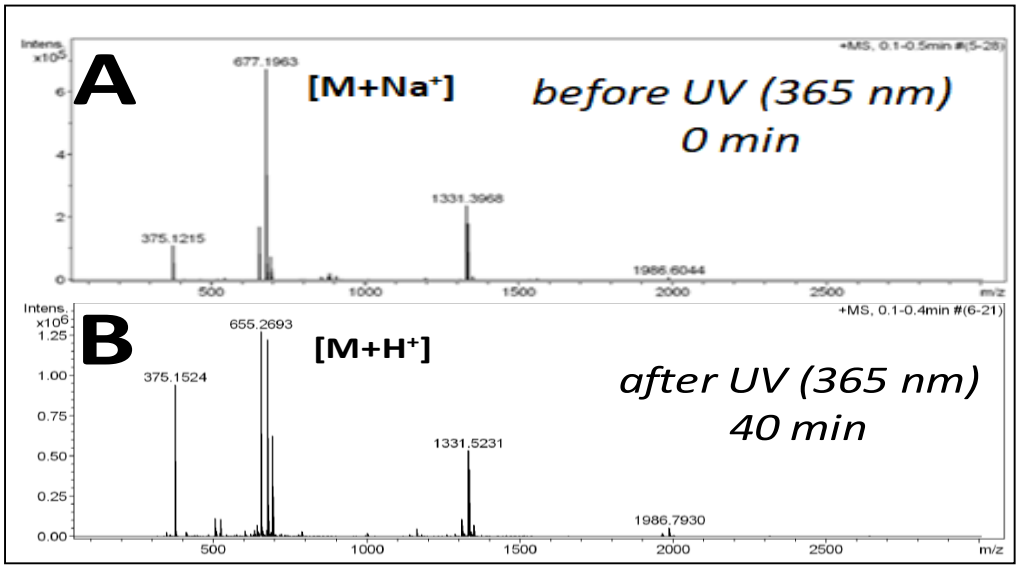


Figure S50. ESI-MS of compound **9b** irradiated with 365 nm in MeOH before (A) and after (B) UV studies, indicating no change in ESI-MS after 40 min of irradiation. (No photo cleavage).

26.UV spectra of Compound **6a**/**6b**/**9b** irradiated with 365 nm at different time intervals.

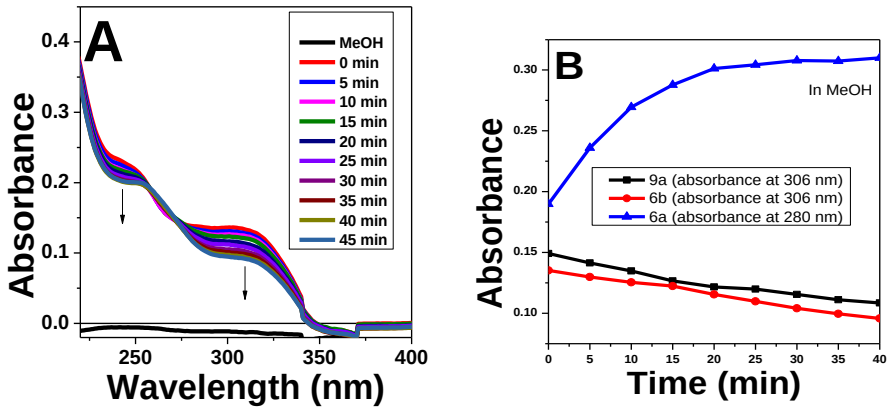


Figure S51. UV spectra of compound **6b** irradiated with 365 nm at different time intervals in MeOH (A) Absorbance of compound **6a**/**6b**/**9b** at different exposure times at 365 nm (B)

27.Stability of NBn^{6}A derivative **10a** in acidic and basic conditions

A. Acid treatment

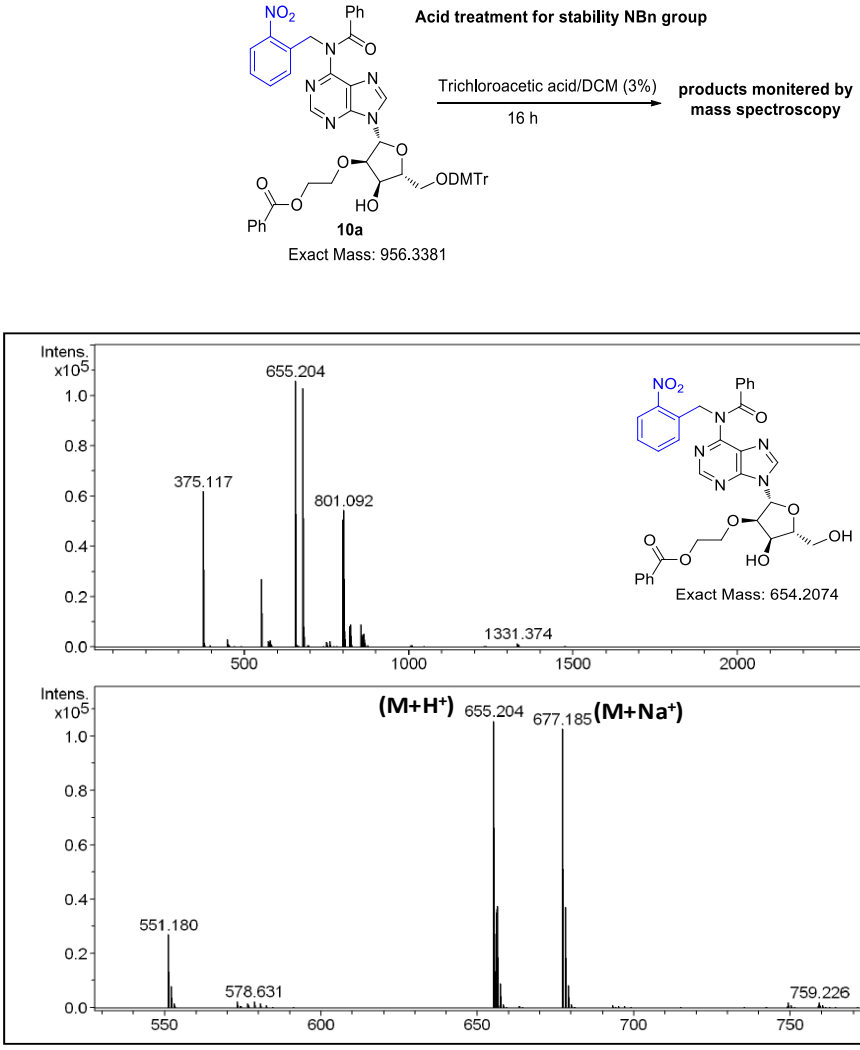
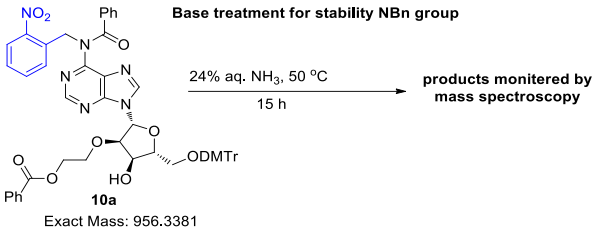


Figure S52. ESI Mass of compound **10a** (NBn^{6}A derivative) after treatment with 3% TCA/DCM, mass spectra recorded after 16 h. The mass spectra indicate that the *o*-nitrobenzyl group is intact with adenosine moiety in acidic conditions even after 16 h of acid treatment. Prolonged exposure to 3% TCA, the cleavage of benzoyl group (m/z 551.18) and addition of trichloroacyl group (m/z 801.09) is observed after DMTr deprotection. The mass spectral analyses indicate the presence of *o*-nitrobenzyl group in observed in mass peaks.

A. Base treatment:



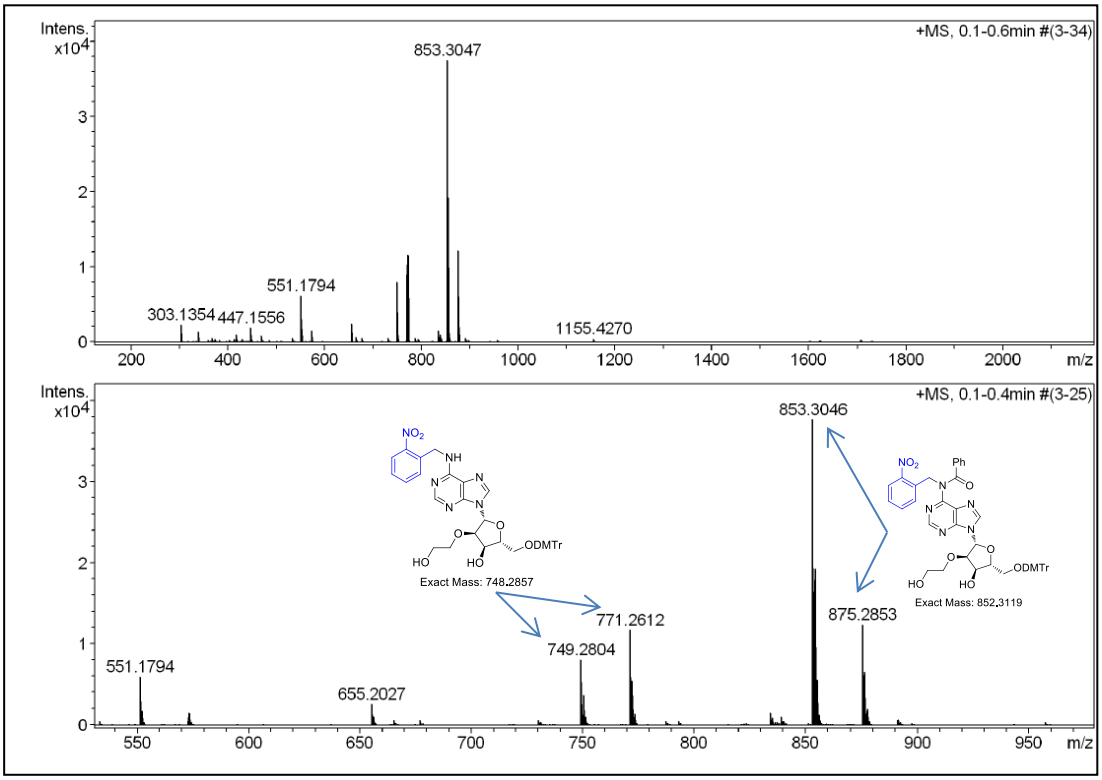


Figure S53. ESI Mass of compound **10a** (NBn^{6}A derivative) after treatment with aq.ammonia at 50 °C, mass spectra recorded after 15 h. The mass spectra indicate that the *o*-nitrobenzyl group is intact with adenosine moiety in basic conditions even after 15 h of base treatment. Under basic conditions the benzoyl group of ester is hydrolyzed m/z 749.28, ($M+H^+$) and 771.26 ($M+Na^+$); amide was also hydrolyzed m/z 853.304 ($M+H^+$) and 875.285 ($M+Na^+$). The mass spectral analyses indicate the presence of *o*-nitrobenzyl group in observed mass peaks.

28.Stability of NBn^{1}A derivative **10b** in acidic and basic conditions

A. Acid treatment:

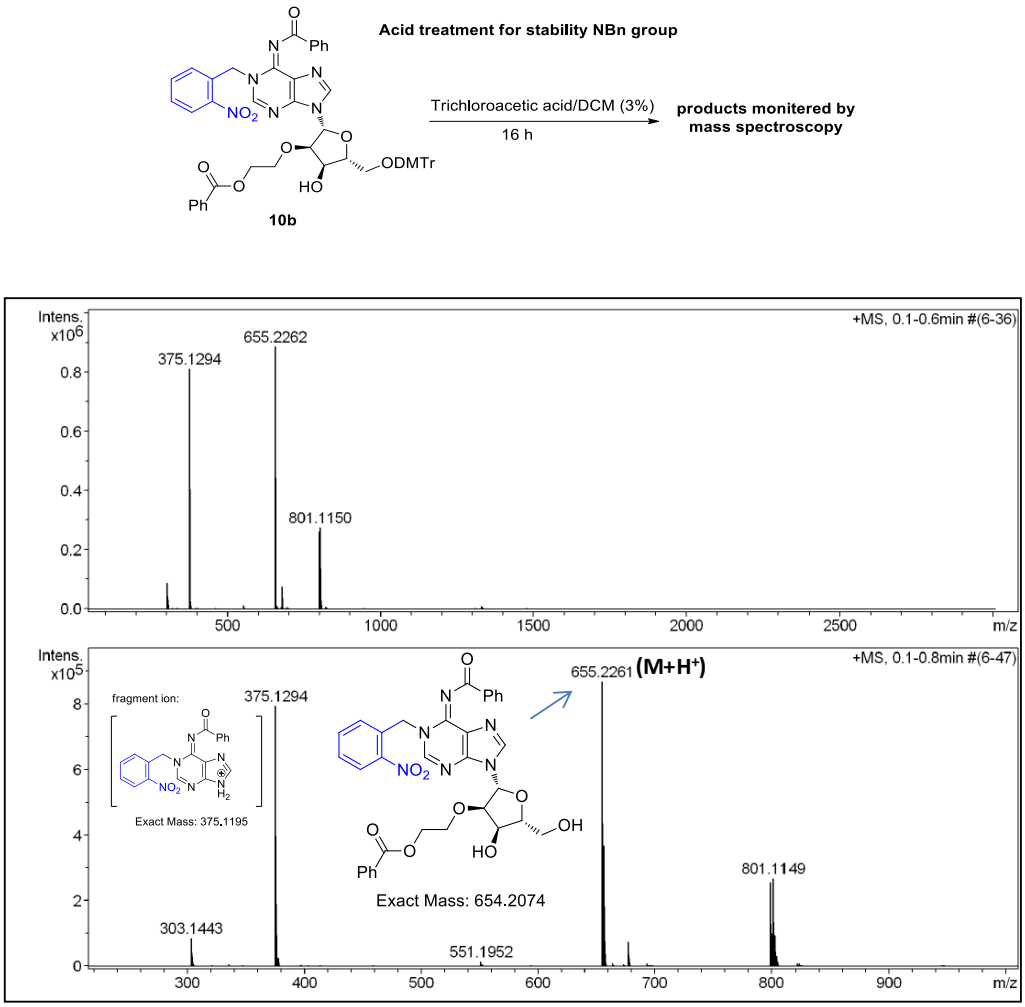


Figure S54. ESI Mass of compound **10b** (NBn^{1}A derivative) after treatment with 3% TCA/DCM, mass spectra recorded after 10 h. The mass spectra indicate that the *o*-nitrobenzyl group is intact with adenosine moiety in acidic conditions even after 10 h of acid treatment. Prolonged exposure to 3% TCA, addition of trichloroacyl group (m/z 801.09) is observed after DMTr deprotection. The mass spectral analyses indicate the presence of *o*-nitrobenzyl group in observed in mass peaks.

B. Base treatment:

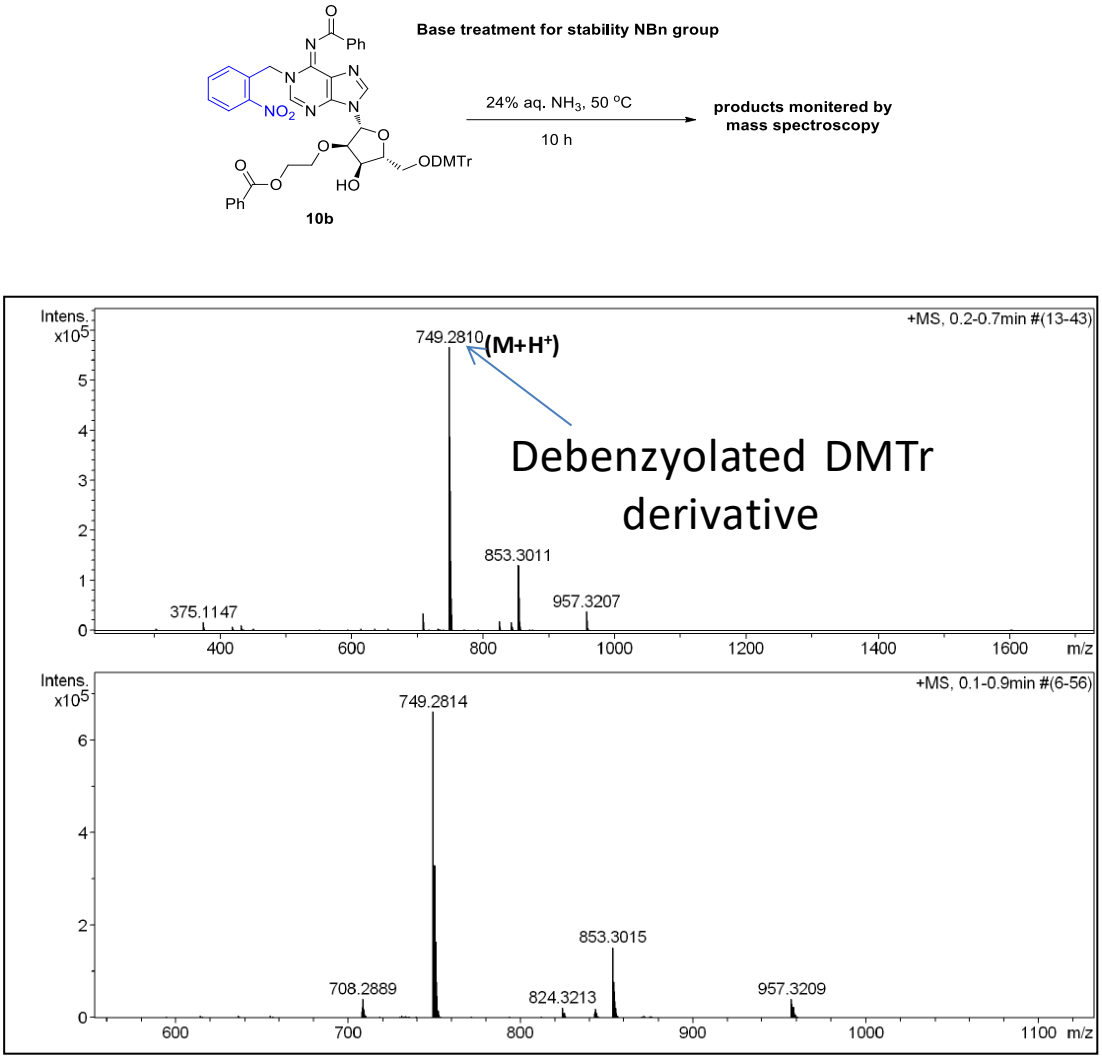


Figure S55. ESI Mass of compound **10b** (NBn^{6}A derivative) after treatment with aq.ammonia at 50 ^{o}C, mass spectra recorded after 10 h. The mass spectra indicate that the *o*-nitrobenzyl group is intact with adenosine moiety in basic conditions even after 10 h of base treatment. The mass spectral analyses indicate the presence of *o*-nitrobenzyl group in observed mass peaks.

29.ESI-MS of RNA oligonucleotides and their photo cleavage studies with 365 nm UV light.

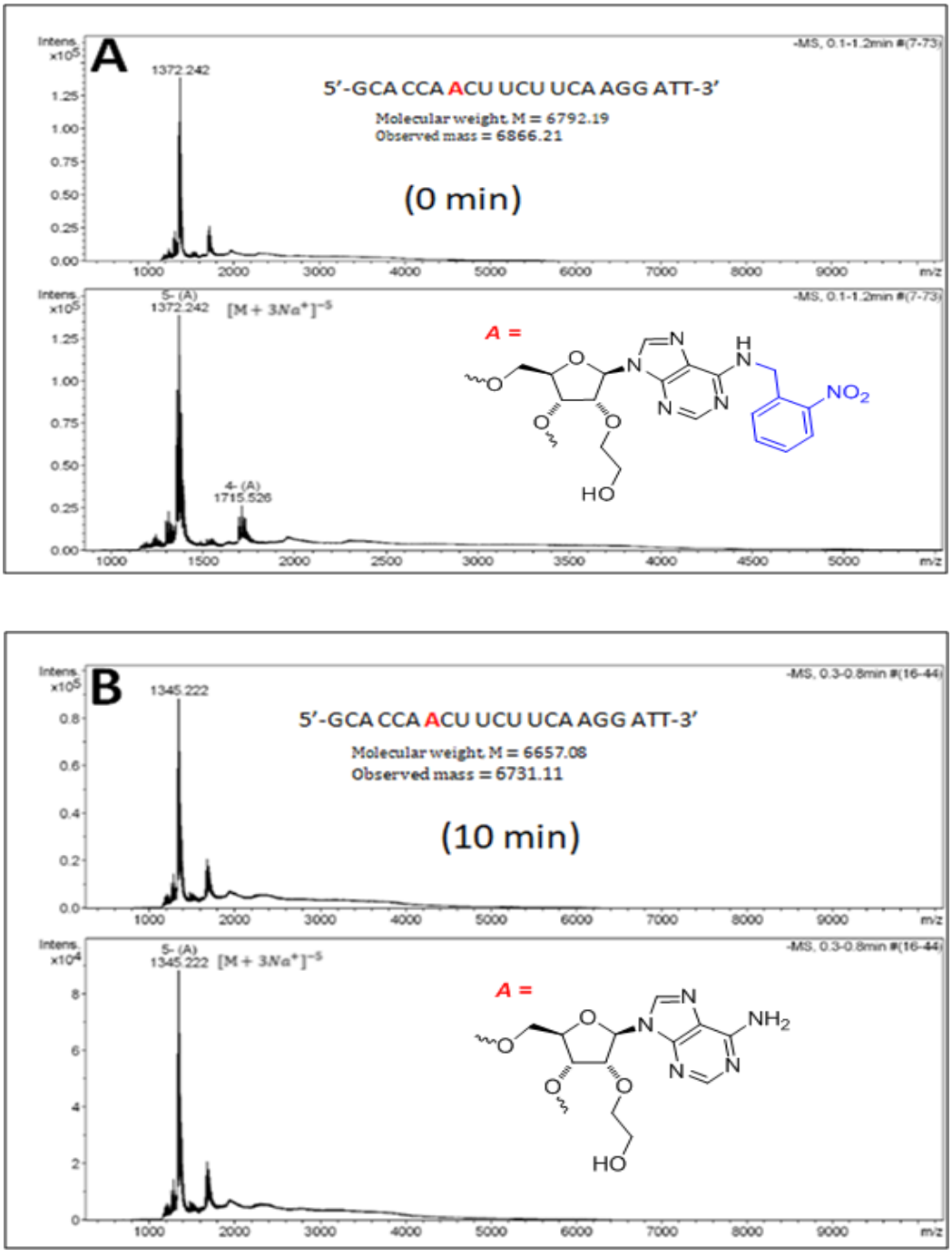


Figure S56. ESI-MS of NBn^{6}A-RNA-1, before (A) and after (B) irradiating at 365 nm for 10 min. (6866.21-6731.11 = 135.1 mass difference per one nitrobenzyl group) indicating the photo cleavage of 2-nitrobenzyl group.

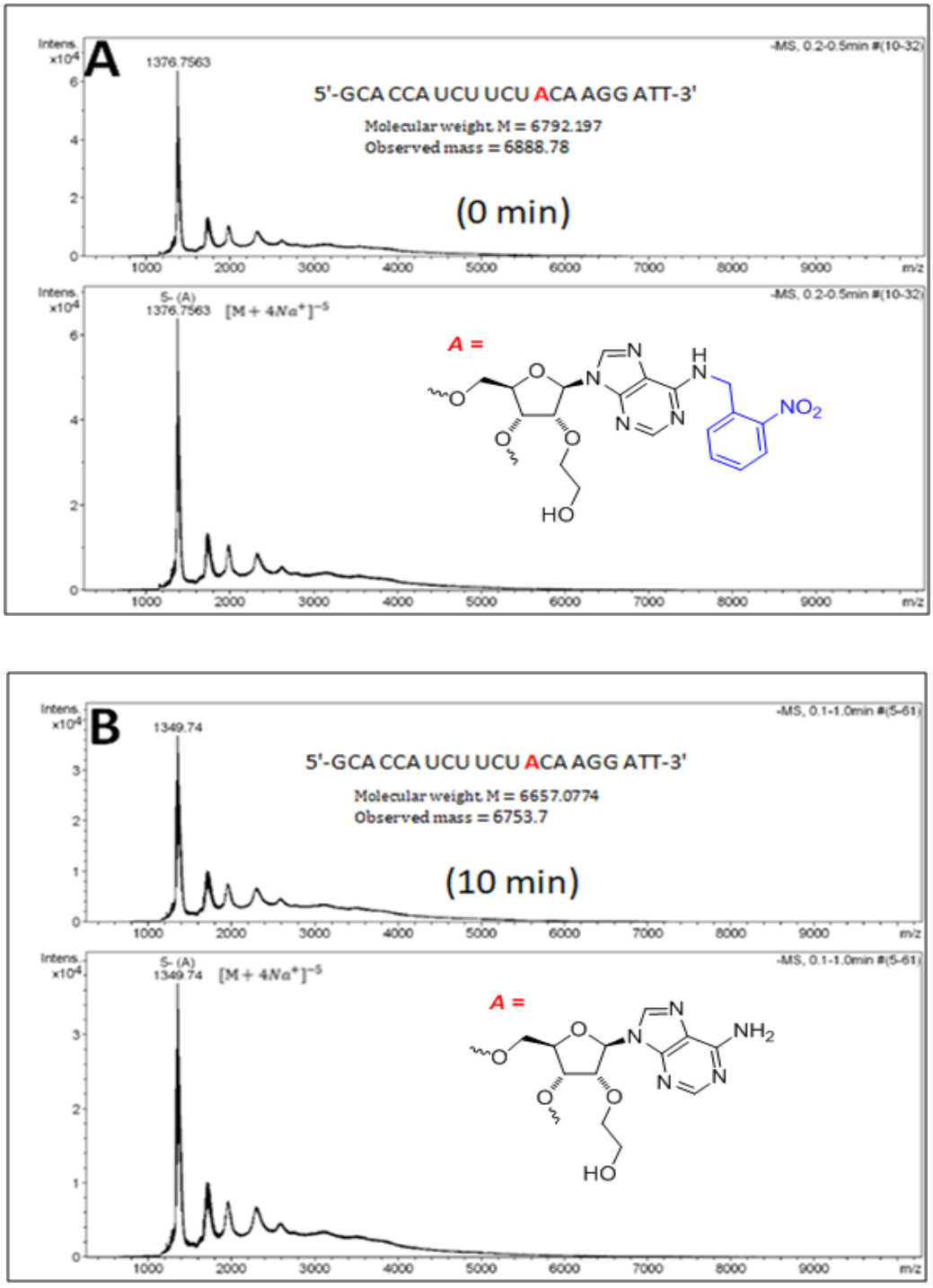


Figure S57. ESI-MS of NBn^{6}A-RNA-2, before (A) and after (B) irradiating at 365 nm for 10 min. (6888.78 -6753.7 = 135.1 mass difference per one nitrobenzyl group) indicating the complete photo cleavage of 2-nitrobenzyl group.

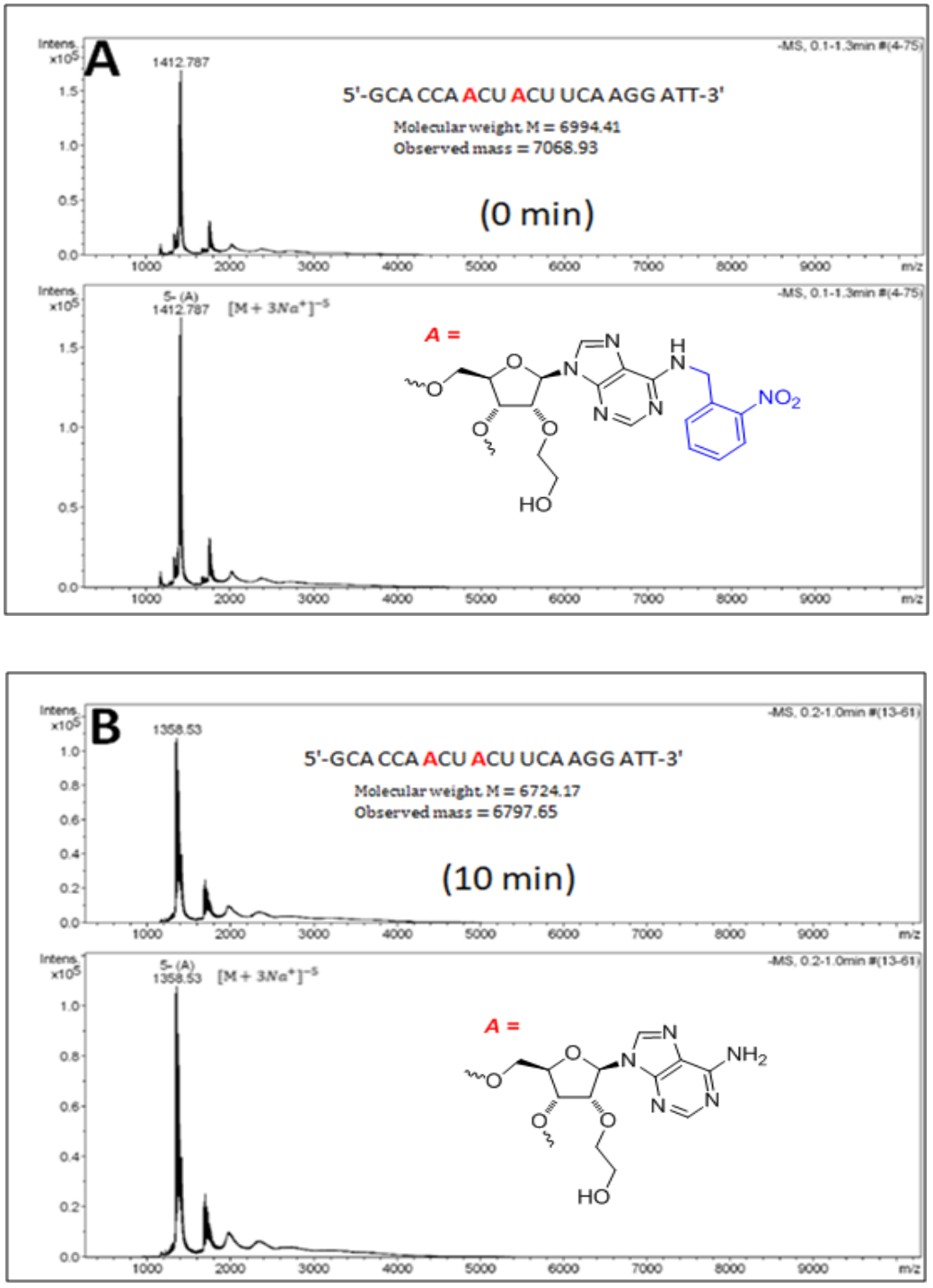


Figure S58. ESI-MS of NBn^{6}A-RNA-3, before (A) and after (B) irradiating at 365 nm for 10 min. ((7068.93 -6797.65 = 271.28)/2 modifications = 135.64 mass difference per one 2-nitrobenzyl group) indicating the complete photo cleavage of two 2-nitrobenzyl groups.

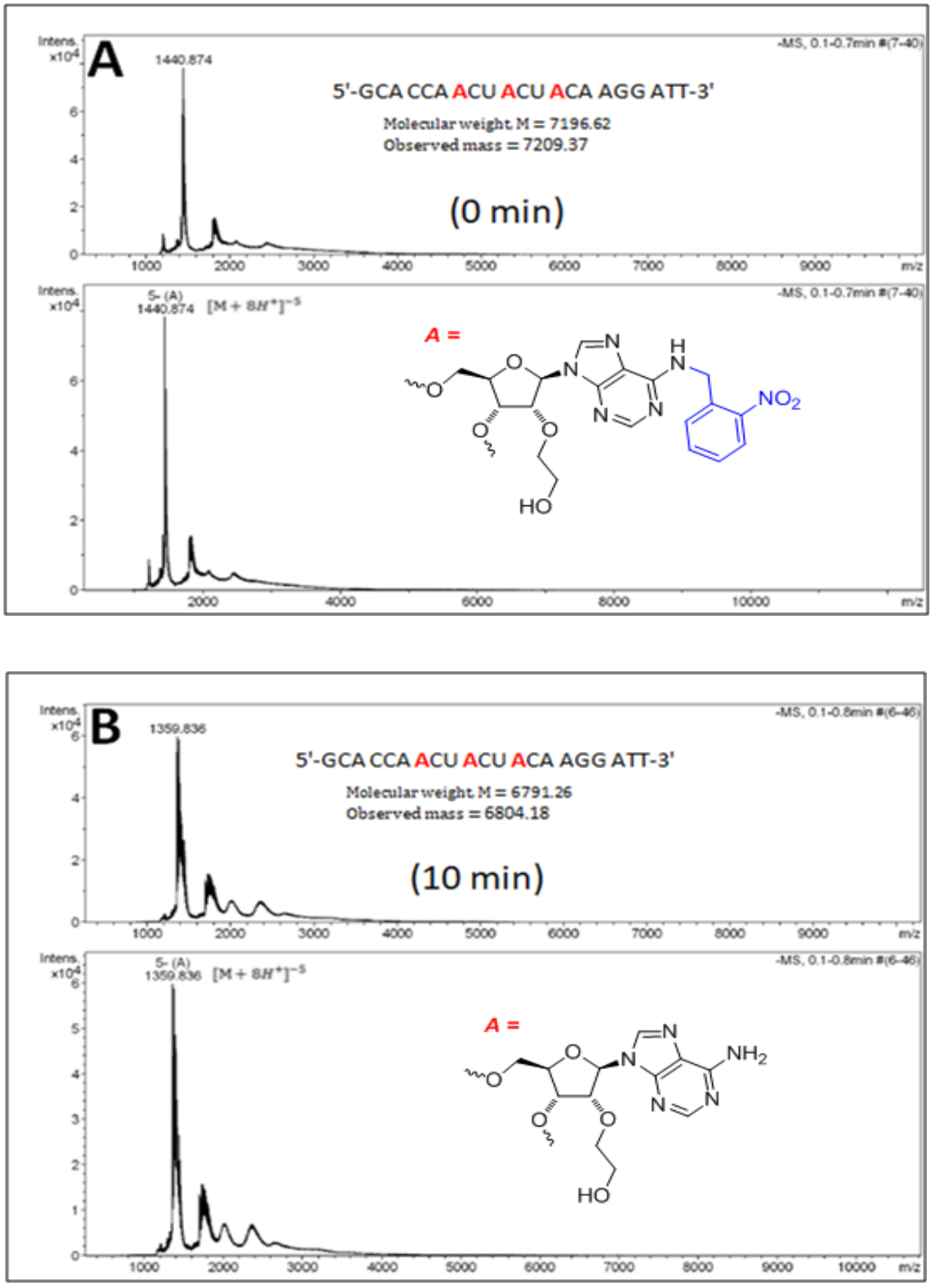


Figure S59. ESI-MS of NBn^{6}A-RNA-4, before (A) and after (B) irradiating at 365 nm for 10 min. ((7209.37 -6804.18 = 405.2)/3 modifications = 135.06 mass difference per one 2-nitrobenzyl group) indicating the complete photo cleavage of three 2-nitrobenzyl groups.

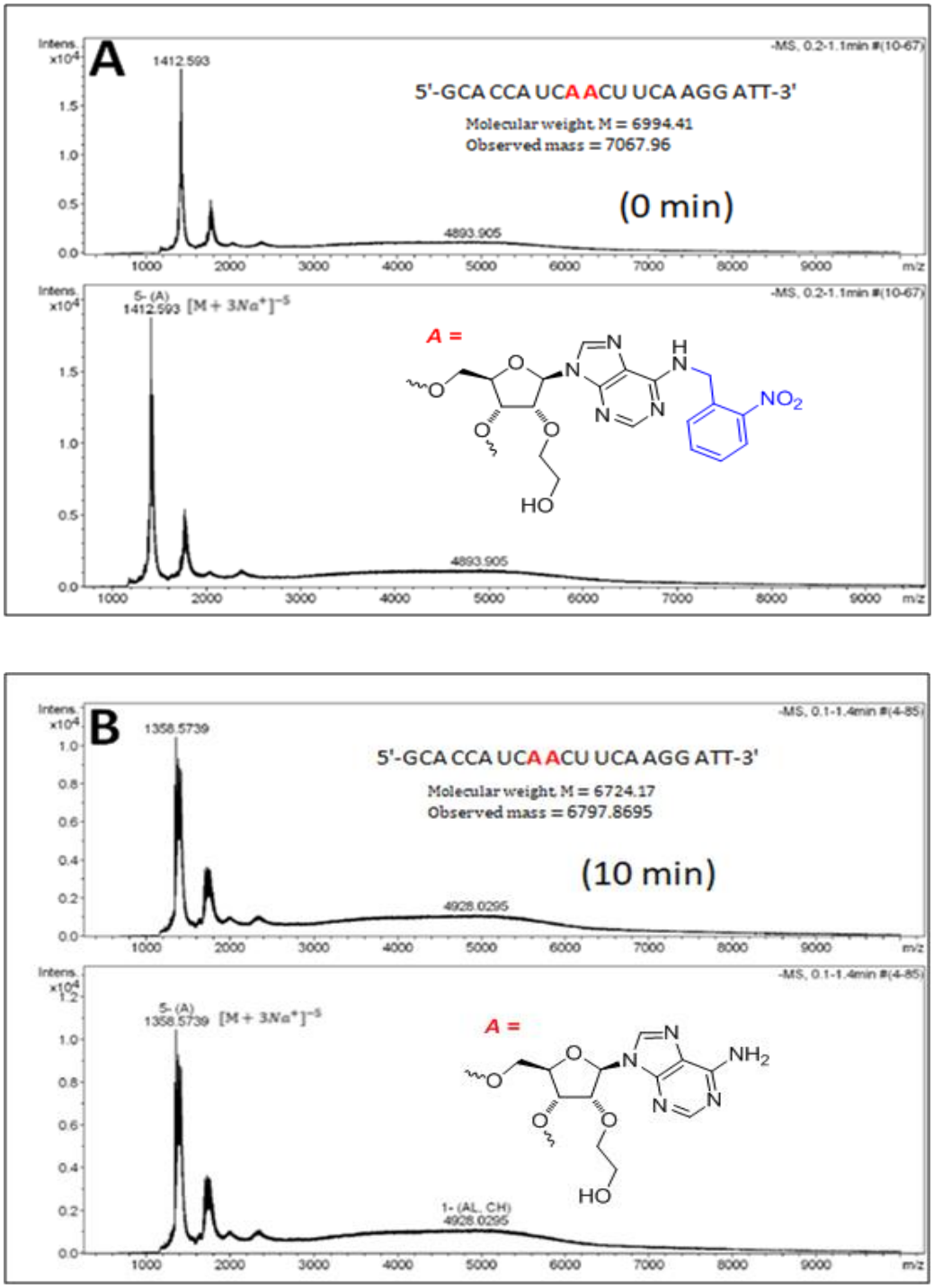


Figure S60. ESI-MS of NBn^{6}A-RNA-5, before (A) and after (B) irradiating at 365 nm for 10 min. ((7067.96 -6797.87 = 270.1)/2 modifications = 135.06 mass difference per one 2-nitrobenzyl group) indicating the complete photo cleavage of two 2-nitrobenzyl groups.

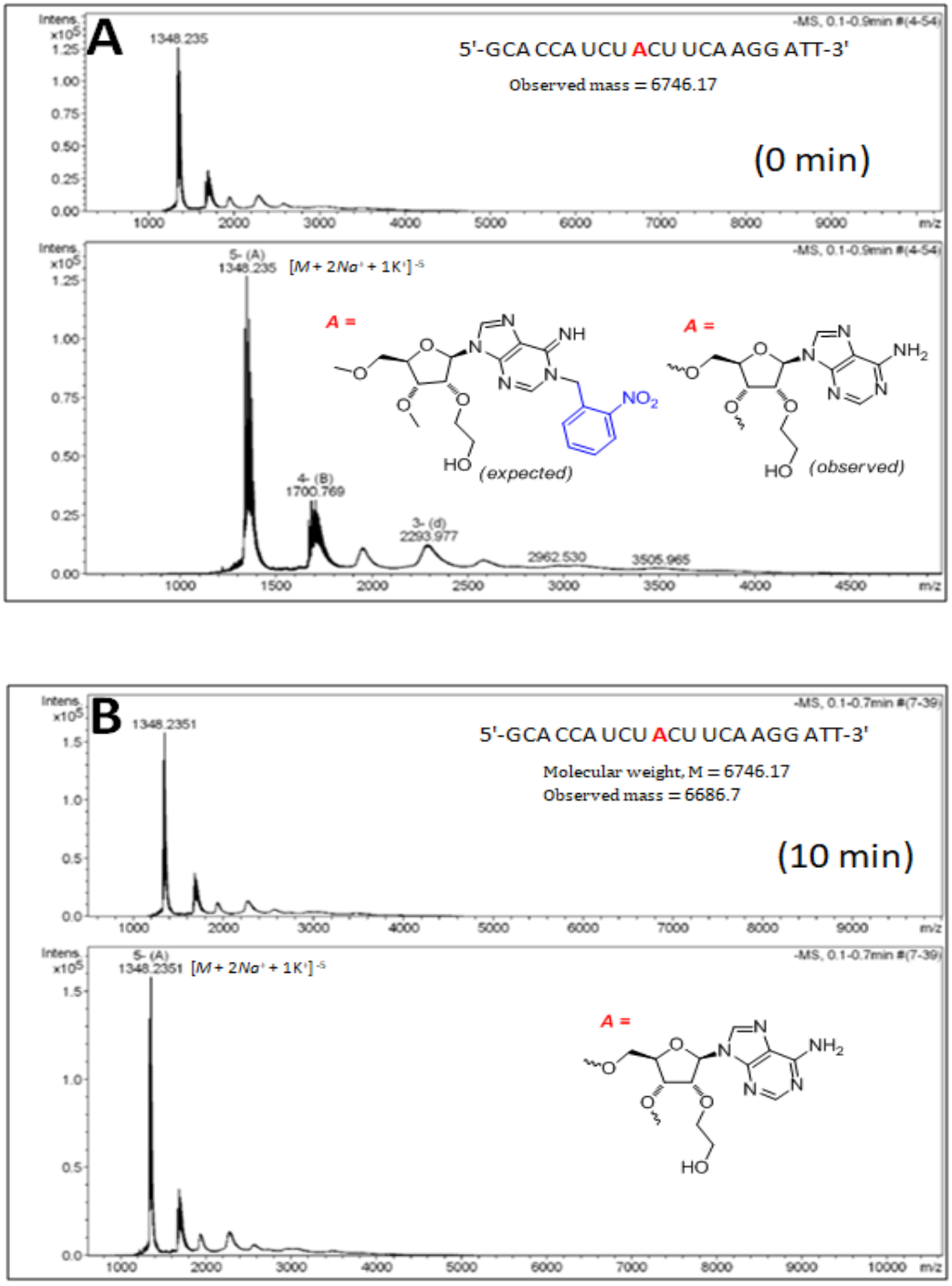


Figure S61. ESI-MS of NBn[1]A-RNA-6, before (A) and after (B) exposed at 365 nm for 10 min. No change in m/z after irradiation, which confirm the absence of 2-nitrobenzyl group.

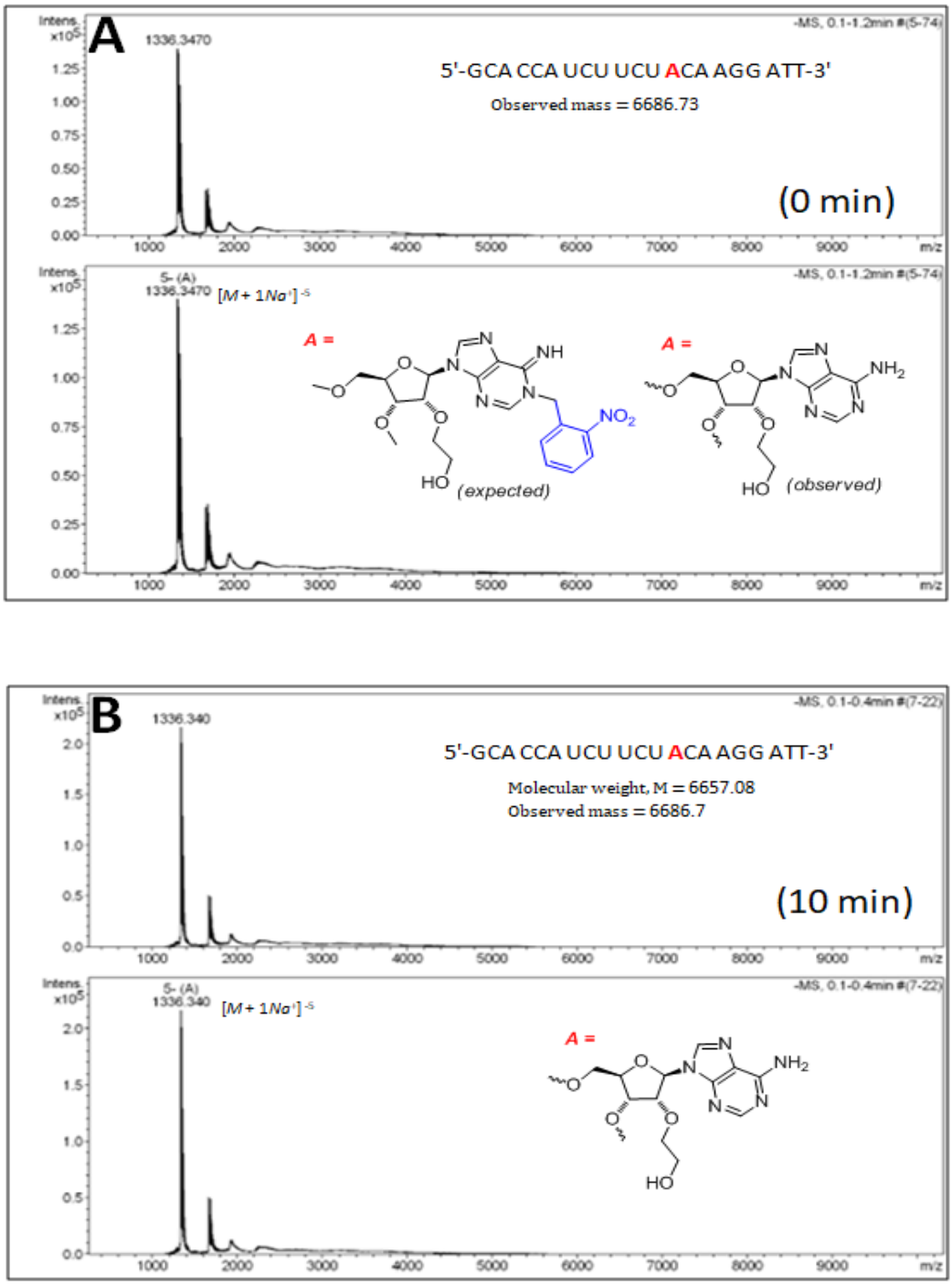


Figure S62. ESI-MS of NBn[1]A-RNA-7, before (A) and after (B) exposed at 365 nm for 10 min. No change in m/z after irradiation, which confirm the absence of 2-nitrobenzyl group.

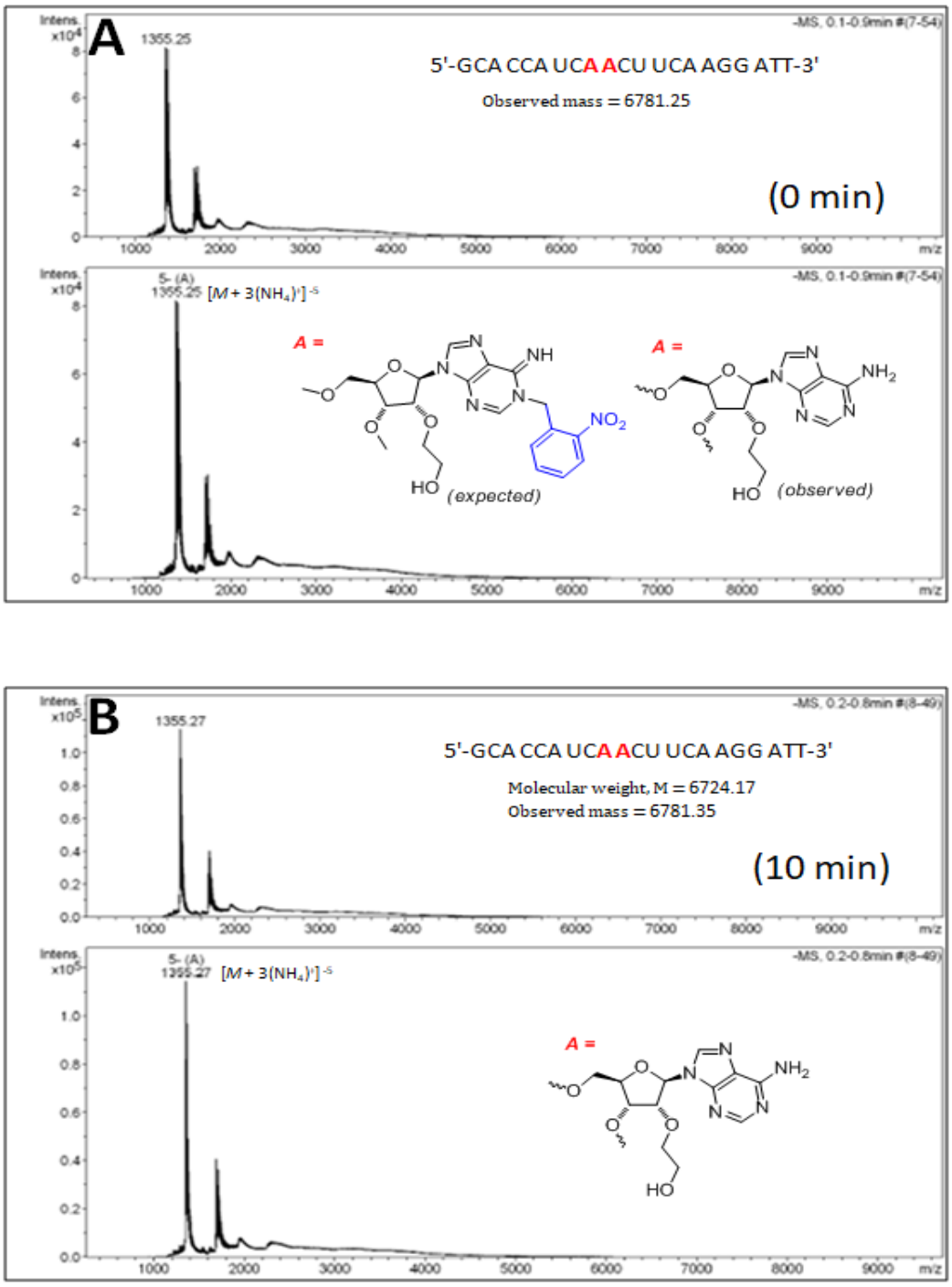


Figure S63. ESI-MS of NBn[1]A-RNA-8, before (A) and after (B) exposed at 365 nm for 10 min. No change in m/z after irradiation, which confirm the absence of 2-nitrobenzyl group.

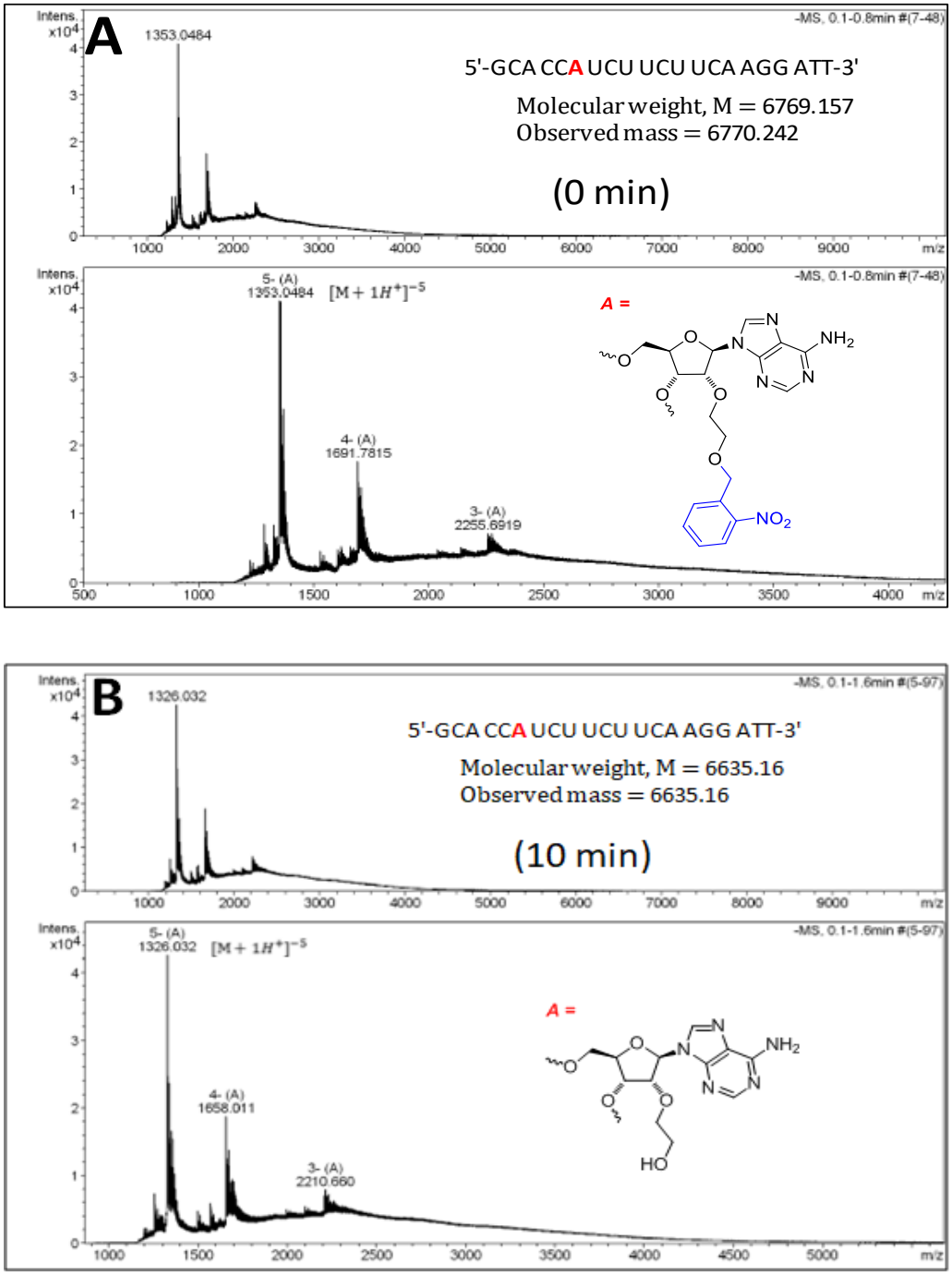


Figure S64. ESI-MS of A-NBnOE-RNA-1, before (A) and after (B) irradiating at 365 nm for 10 min. ((6770.24 -6635.16 = 135.08)/1 modifications = 135.06 mass difference per one *o*-nitrobenzyl group) indicating the complete photo cleavage of *o*-nitrobenzyl groups.

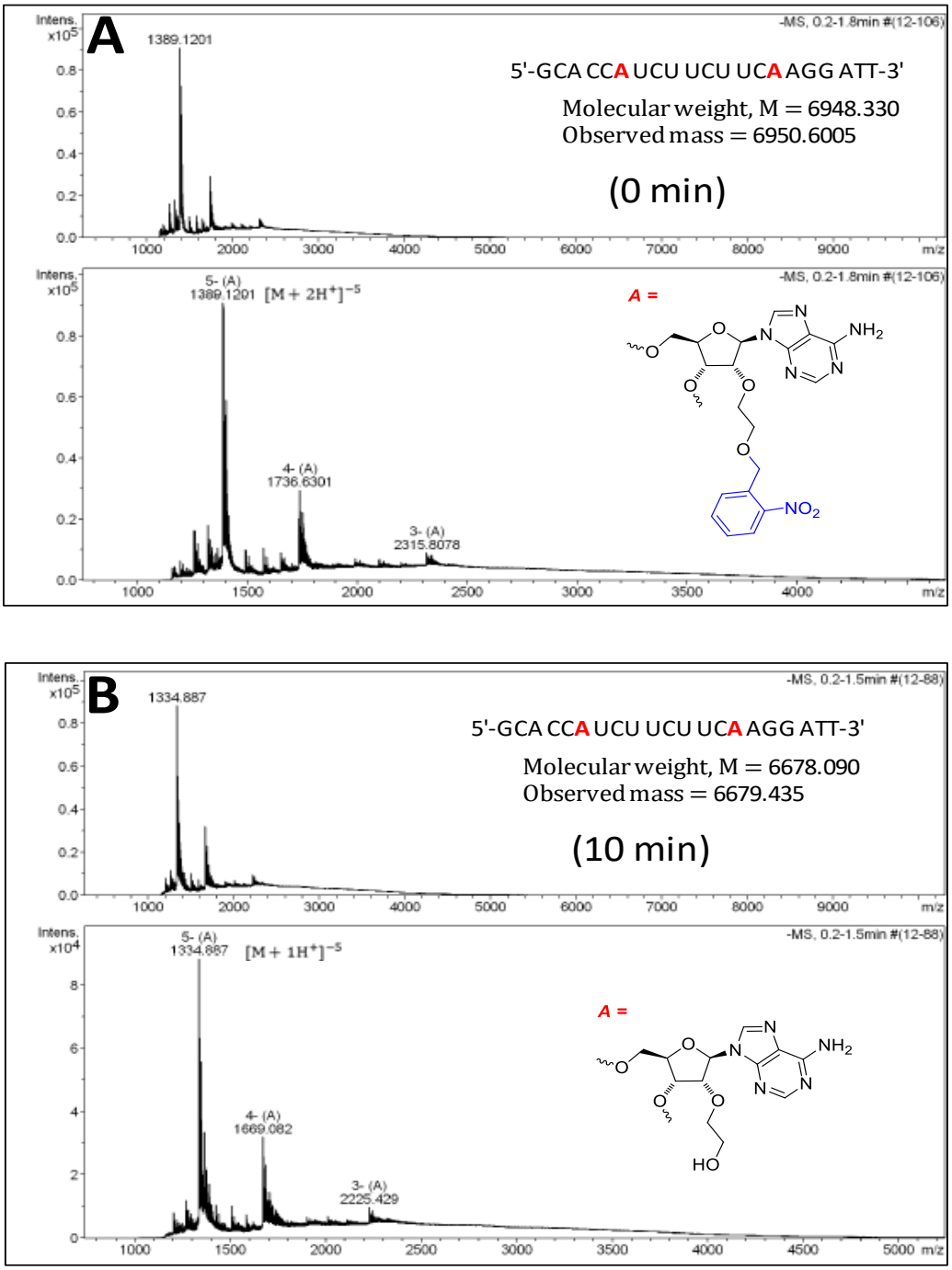


Figure S65. ESI-MS of A-NBnOE-RNA-2, before (A) and after (B) irradiating at 365 nm for 10 min. ((6950.60 -6679.43 = 271.17)/2 modifications = 135.58 mass difference per one *o*-nitrobenzyl group) indicating the complete photo cleavage of two *o*-nitrobenzyl groups.

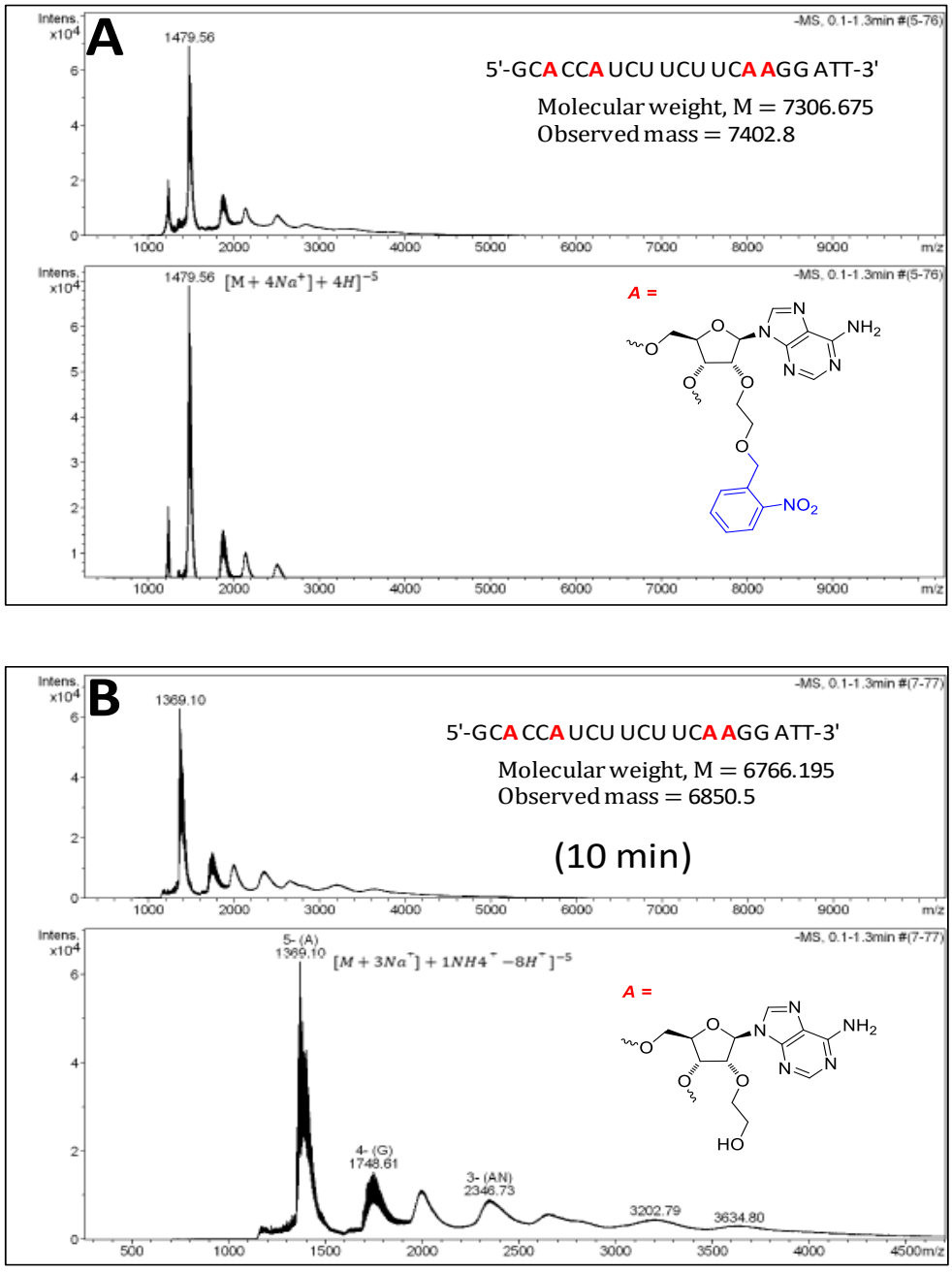


Figure S66. ESI-MS of A-NBnOE-RNA-2, before (A) and after (B) irradiating at 365 nm for 10 min. ((7402.8 -6850.5 = 552.3)/4 modifications = 138.07 mass difference per one *o*-nitrobenzyl group) indicating the complete photo cleavage of 4 *o*-nitrobenzyl groups.

CHAPTER 3

Part-A

Synthesis and Photophysical Studies of *Tropolone* Conjugated Nucleoside and DNA Oligonucleotides

Chapter 3A. Synthesis and Photophysical Studies of *Tropolone* Conjugated Nucleoside and DNA Oligonucletides

3A.1 Introduction

Tropolone is an α-hydroxy derivative of *Tropone* molecule which is a non-benzenoid seven-membered aromatic compound with a carbonyl functional group.[1-3] Tropolone derivatives, tropolonoids, occur in nature mostly in plants and fungi.[4] Tropolone and related derivatives possess biological activities, such as antibacterial,[5] antifungal,[6] antitumor,[7] and antiviral[8] activities. Recent research shows that tropolones could be potent and selective inhibitors for enzymes with zinc-cofactors.[9] Tropolone is also reported as distinctive metal chelating scaffold especially for $Cu^{2+}/Zn^{2+}/Ni^{2+}$, because of its α-hydroxy carbonyl functional group.[10, 11] The aromatic structure of tropolone exhibits interesting photophysical properties (absorptions & emission) and considered as a useful structural moiety for designing the small molecule derived chromophores. Absorption spectra of tropolone exhibits two peaks at wavelength ranges 200-260 nm and 260-380 nm due to two types of electronic transitions: two π-π* and n-π* transitions.[12, 13] Further, the pK_a of tropolone is reported as pK_a 6.5 and is reported as weak fluorescent molecule.[14] Its fluorescence properties are described in terms of the existence of the (n-π*) singlet state. Martin and co-workers show that the fluorescent characters of tropolone, at excitation wavelength (λ) 315 nm, are dependent upon the solvent polarity and pH conditions such as quantum yield of tropolone is measured as 0.005 in cyclohexane (neutral), 0.003 in acidic water (pH 3.0) and $0.005*10^{-2}$ in basic conditions (pH 12.0).[15] These variation are due to the different forms of tropolone structures as neutral or anion forms due to hydroxyl protonation and deprotonation process under different pH conditions. Thus we realized that tropolone could be an interesting scaffold to incorporate into nucleic acids for

tuning their physical and structural properties. Recently, our lab reported the tropolonyl bearing amino acid/peptides and established the role of tropolonyl carbonyl in conformational changes of peptides.[16] However, we have found the role of tropolone in an unusual cleavage of adjacent amide bond under mild pH conditions and reversible amidation in mild basic conditions.[16, 17] To explore the photophysical properties of tropolone related compounds, we have also explored the tropolone derived BODIPY analogues and demonstrated their photophysical properties with fluorescent quantum yield upto ~16%.[18] The incorporation of tropolone into nucleic acid would bring conjugated DNA/RNA analogues. However, the various chromophore conjugated nucleosides and their DNA are reported and some of them have shown distinctive photophysical properties as the enhancement in fluorescence after formation of duplex with native complimentary DNA oligonucleotide. Recent examples include, 2-aminopurine containing DNA as first potential fluorescent adenine surrogates,[19-21] Tor's isomorphic RNA alphabet,[22, 23] Wilhelmsson's internucleobase FRET pair,[24] Sasaki's fluorescent sensors for nucleobase damage,[25] Purse's tricyclic cytidine analogue,[26, 27] Luedtke's ^{DMA}T nucleoside,[28] and a number of modified nucleobases oriented in major groove of DNA duplex which interact with protein.[29] Their outcomes suggest that the photo physical behavior of chromophores are affected with alteration in microenvironment of DNA duplex such as polarity, hydrophobicity, and pH conditions. Importantly, the photophysical behavior as fluorescent character of tropolone decreases with increasing the polarity and pH conditions of environment.[15]

3A.2 Objective of our work

In repertoire of small chromophore conjugated DNA analogues, we hypothesized to conjugate tropolone moiety into 2‘-deoxyuridine (at C-5) via conjugated alkyne linker which might be beneficial to utilize the fluorescent character of tropolone moiety in the sensing of DNA duplex formation. Further the tropolone is good metal chelator and can be used in binding with metal. (Figure 3A.1)

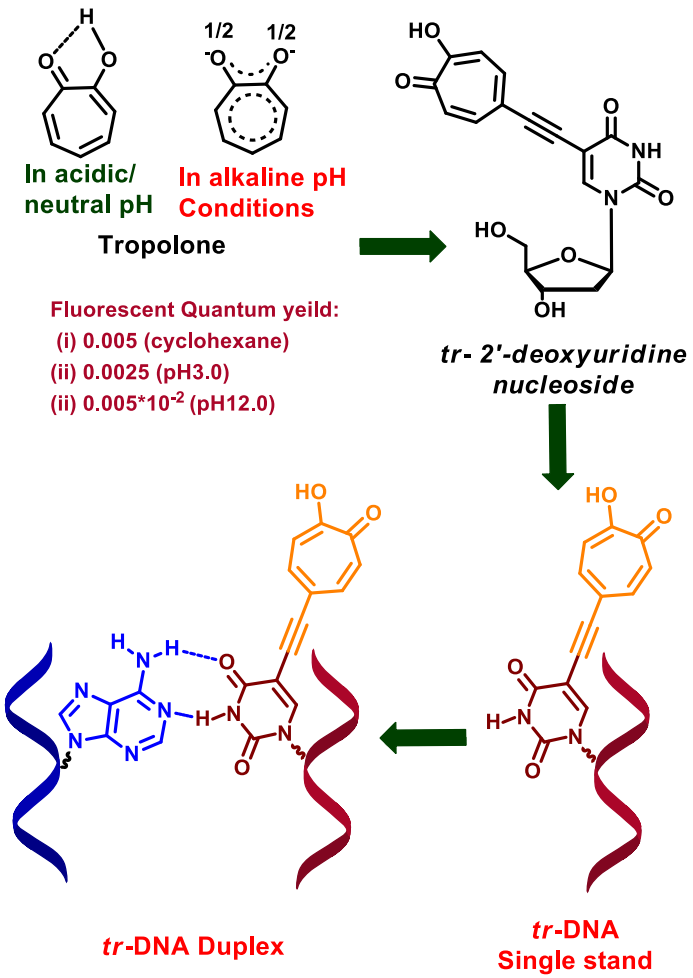
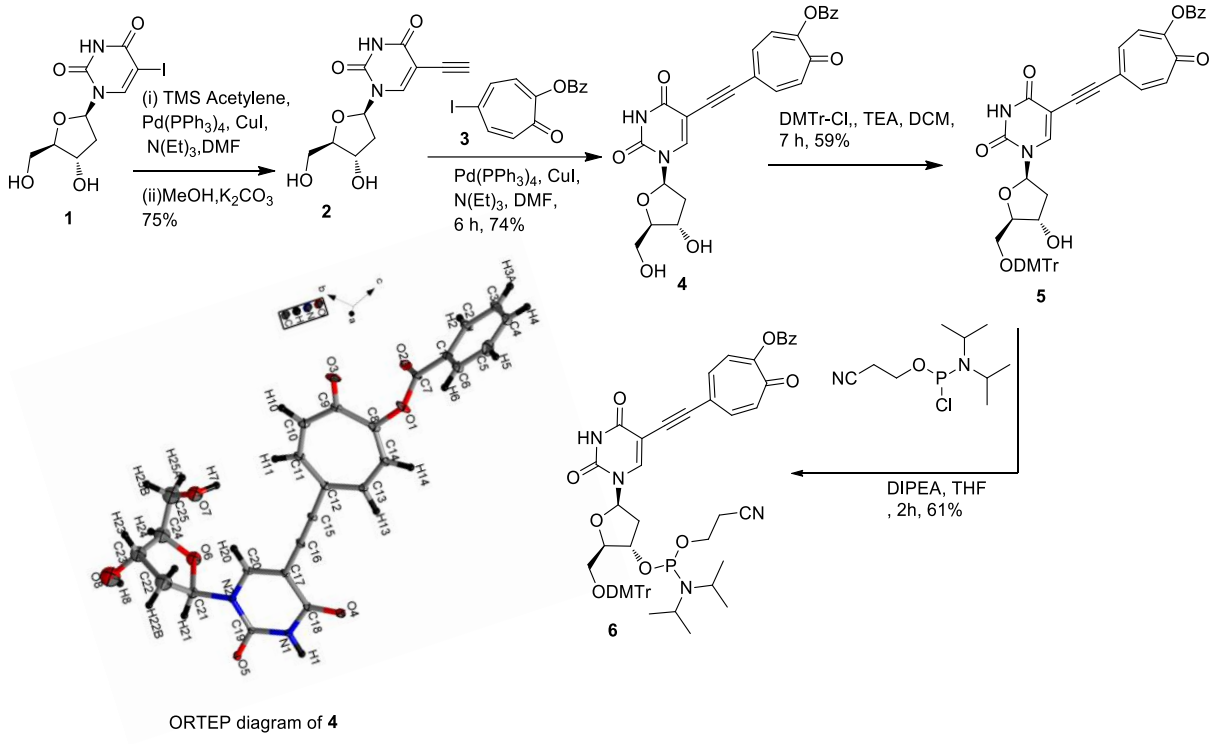


Figure 3A.1. (a) Tropolone; (b) tropolonylated nucleoside and DNA.

3A.3 Results and Discussion

We began the synthesis of the rationally designed tropolonylated thymidine nucleoside and its phophoramidites from commercially available 5-Iodo-2'-deoxyuridine (**1**) and Tropolone (Scheme 1). 5-Iodo-2'-deoxyuridine (**1**) was converted into 5-Ethynyl-2'-deoxyuridine (**2**) in two steps by following the literature procedure.[30] The alkyne group of nucleoside (**2**) was coupled with benzoyl protected 5-Iodotropolone (**3**) under sonogashira coupling reaction conditions using Pd-catalyst which produced 5-tropolonylated 2'-deoxyuridine nucleoside analogue (**4**) in a good yield. The 5-iodotropolone derivative (**3**), however, was prepared from commercially available Tropolone by following the reported procedures.[31] This tropolonylated 2'-deoxyuridine nucleoside (**4**) was characterized by ^{1}H-/^{13}C-NMR/ESI-HRMS analyses (See Appendix). Pleasantly we obtained the single crystal of tropolonylated 2'-deoxyuridine derivative (**4**) from 2% MeOH in DCM solvent. The X-ray analysis of that single crystal confirms the structure of new tropolonylated 2'-deoxyuridine nucleoside (**4**). The single crystal X-ray data are deposited to Cambridge Crystallographic Data Centre with CCDC 1830067. The ORTEP diagram of modified 2'-deoxyuridine derivative (**4**) is also depicted in Scheme 1, while its X-ray parameters are provided in the Appendix. For synthesis of phosphoramidities, 5'-OH of tropolonylated nucleoside (**4**) was protected with DMTr as compound **5**, and employed for phosphorylation, at 3'-OH, with phosphorylating agent (diisopropylchlorophosporamidites) under basic conditions which produced the desired tropolonylated thymidine phsophoramidite (**6**) in good yield. The characterization data (^{1}H-/^{13}C-NMR/^{31}P-NMR/ESI-HRMS) of compound **6** are provided in Appendix.

Scheme 3A.1. Synthesis of tropolonylated 2'-deoxyuridine phosphoramidites and ORTEP diagram of compound (**4**).



Further we attempted to incorporate the new tropolonylated phosphoramidite (**6**) into DNA oligonucleotides at different positions with DNA oligo synthesizer. Herein we synthesized duplex forming DNA from known oligonucleotides (**ON1 & ON2**) with commercially available DNA phosphoramidites for control studies (Table 3A.1, Entry 1-2).Tropolonylated phosphoramidite (**6**) was incorporated into DNA oligonucleotide **ON2** and prepared tropolonylated DNA (*tr*-DNA) oligonucleotides such as *mono*-tropolonylated DNA oligonucleotide (**ON3**)**,** *di*-tropolonylated *DNA* oligonucleotide (**ON4**) and *tri*-tropolonylated DNA oligonucleotides (**ON5)** (Entry 3-5, Table 3A.1).

Further, phosphoramidite (**6**) is also incorporated into G-rich, C-rich sequences (Entry 6-13, Table 3A.1) and also tested with self coupling upto 4-nucleotide long in 5-mer DNA (Entry 14, Table

3A.1). The oligonucleotides were purified by Sephadex column followed by HPLC and then characterized by ESI-MS which confirmed the sucesfull syntheses of tropolone incorporated oligonucleotides (Column 4-5, Table 3A.1), their ESI-Mass spectra are provided in the Appendix. Herein, **ON1-ON5** are exclusively studied while others (**ON6-ON14**) are in progress.

Table 3A.1. Synthesized DNA oligonucleotides and their mass

Entry	code	DNA Oligonucleotides (5'-3') [a]	Calculated Mass (M)	Observed Deconvoluted Mass
1	**ON1**	GCGTACACATGCG	3959.639	3956.521
2	**ON2**	CGCATGTGTACGC	3950.626	3947.636
3	**ON3**	CGCATG**T***GTACGC	4080.639	4076.615
4	**ON4**	CGCATG**T***G**T***ACGC	4210.639	4205.630
5	**ON5**	CGCA**T***G**T***G**T***ACGC	4340.639	4334.503
6	**ON6**	CCCTAACCCTAACCCTAACCCT	6504.303	6504.35
7	**ON7**	AGGGTTAGGGTTAGGGTTAGGG	6966.57	6966.6
8	**ON8**	CCCTAACCCTAACCC**T***AA CCCT	6634.303	6632.74
9	**ON9**	AGGGTTAGGGTTAGGG**T***TAGGG	7096.57	7096.58
10	**ON10**	**T***GG GGT	1993.27	1992.42
11	**ON11**	**T***CC CCC	1818.16	1817.37
12	**ON12**	TGG GGT	1863.27	1862.23
13	**ON13**	TCC CCC	1688.16	1687.38
14	**ON14**	**T* T* T* T***T	1979.40	1978.73

[a] **T*** is tropolonylated 2'-deoxyuridine residue

After successful syntheses of *tr*-DNA oligonucleotides we attempted to explore their biophysical and photophysical properties. We determined extinction coefficient (ε) of debenzolylated tropolonylated nucleoside as **4-OH** under aqueous solution in presence of little NH_4OH (~pH 10)

due to poor solubility in neutral pH, and found as 29123 $M^{-1}cm^{-1}$ wavelength (λ) 260 nm (see in Appendix).

CD Studies. Further we examined the formation of tropolonylated DNA hybrid duplex from *tr*-DNA (**ON3**-**ON5**) and unmodified complementary DNA oligonucleotide (**ON1**) near physiological conditions by Circular Dichroism (CD) studies. To avoid the photodecomposition of tropolone under acidic pH conditions (~pH3.0), we recorded the CD spectra of tropolonylated DNA complexes (**ON1:ON3, ON1**-**ON4, ON1:ON5**) at 10 ^{o}C under different pH 6.0/7.0/8.0 conditions. For comparative studies we recorded the CD spectra of control DNA complex (**ON1:ON2**) in pH 7.0. Their CD spectra at pH 7.0 are depicted in Figure 3A.2, CD spectra at other pH conditions are provided in Appendix. In Figure 3A.2A, the CD spectra of *tr*-DNA duplexes (**ON1:ON3/ON1:ON4/ON1:ON5)** exhibit as minima ($\lambda_{252\ nm,}$) and maxima ($\lambda_{284\ nm}$), which are characteristic of DNA B-form helices. In literature, DNA B-form helices have similar CD spectra with characteristic signals as minima (~λ250 nm) and maxima (~λ280 nm).[32, 33] Though, the CD spectra of *tri*-tropolonylated t*r*-DNA duplex (**ON1:ON5)** shows the destabilization in B-form helix under similar pH conditions. For comparative studies, we plotted CD ellipticity vs number of tropolonyl nucleoside unit at different pH conditions (Figure 3A.2B). This plot shows the depletion in the CD signals of *tr*-DNA duplex with increasing *tr*-T nucleoside in DNA strand. At pH 6.0, the CD spectra of *tr*-DNA duplexes (**ON1:ON3/ON1:ON4/ON1:ON5)** exhibit almost similar signals as B-form DNA helix (Figure 2B). At pH 8.0, The CD spectra of *mono*-tropolonylated DNA duplex (**ON1:ON3)** also exhibit almost similar as B-form DNA helix. Moreover, CD spectra of *di*/*tri*-tropolonylated DNA duplex show an extra signals as the minima at ~$\lambda_{370\ nm}$ which are more prominent with *tri*-tropolonylated DNA duplex (**ON1:ON5**) (Figure 2B).

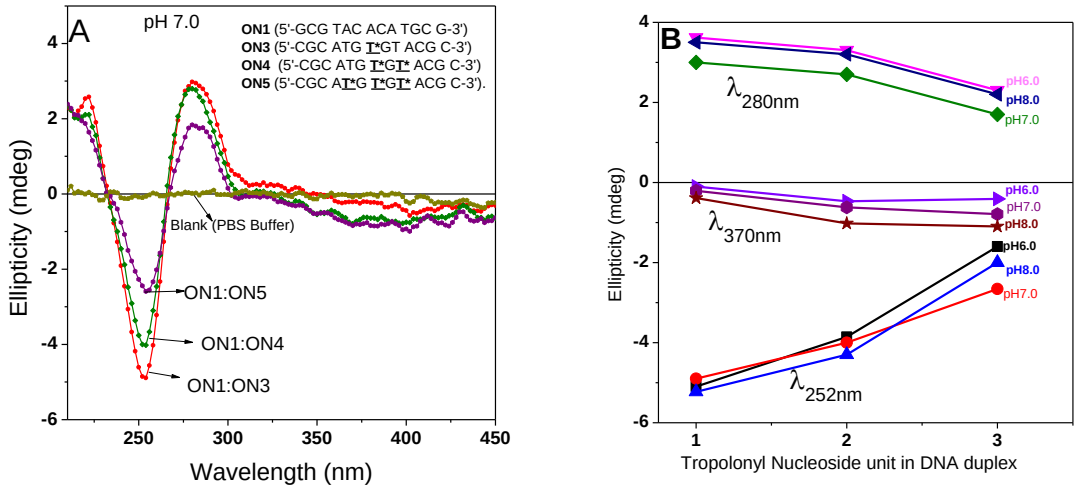


Figure 3A.2. CD spectra of DNA duplexes (2.0 μM) in sodium phosphate buffer (100 mM) at pH7.0 conditions (A); CD ellipticity vs number of tropolonyl nucleoside unit at different pH conditions (B).

UV-Thermal melting: Further the stability of tropolonylated DNA duplex structures were examined by UV-thermal melting or thermal denaturation experiments. We prepared tropolonylated DNA hybrid duplex structure as **ON1:ON3, ON1-ON4, and ON1:ON5** by annealing of equimolar DNA strands under different pH conditions. We then performed UV-Melting experiments at wavelength (λ) 260 nm. For comparative studies, the similar experiment was performed with unmodified DNA strands (**ON1:ON2)** near physiological conditions (pH 7.0) and considered as control experiments. Their melting profiles (Absorbance vs Temperature plots) at pH 7.0 are depicted in Figure 3A.3, while other melting profiles at different pH conditions are provided in Appendix. The typical melting profile of DNA duplex structure is characteristic positive sigmoid curve.[34] At pH 7.0, the UV melting profiles of *tr*-DNA duplex (**ON1:ON3**, **ON1:ON4**, &**ON1:ON5)** exhibit positive sigmoid curves as like control DNA duplex structure

and their first derivative curves are given as inset (Figure 3A.3). The melting profile of same *tr*-DNA duplex at different pH conditions are also exhibit as positive sigmoid curves (Appendix).

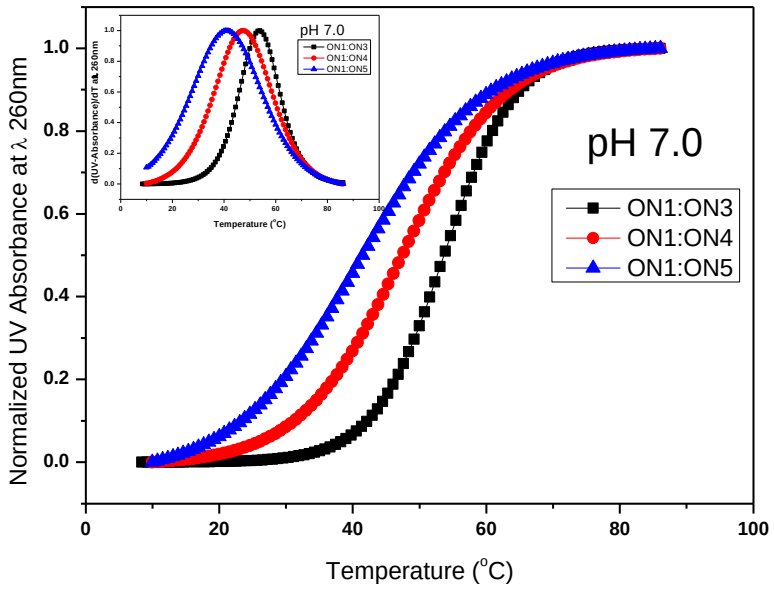


Figure 3A.3. UV thermal melting profiles of *tr*-DNA duplexes (2 µM) at pH 7.0 in sodium phosphate buffer.

The T_m values of *tr*-DNA duplexes are extracted from the first derivative curve of their respective melting profiles which are described in Table 3A.2. Herein we noticed that mono-tropolonylated *tr*-DNA duplex (**ON1:ON3**) was less stable as compared to the control DNA duplex (**ON1:ON2**) near physiological conditions (pH 7.0) by 9.5 ^{o}C. Further the stability of *di-/tri*-tropolonylated DNA duplex (**ON1:ON4**, &**ON1:ON5)** were decreased by 15 ^{o}C and 22.6 ^{o}C respectively at same pH conditions. In pH dependent comparative studies, we found that these *tr*-DNA duplex (**ON1:ON3**, **ON1:ON4**, & **ON1:ON5)** were marginally stabilized by increasing pH conditions to

pH 8.0, while destabilized by lowering the pH condition to pH 6.0. Thus stability of *tr*-duplex is dependent upon pH condition and number of incorporated tropolonyl residue.

Table 3A.2. Synthesized DNA oligonucleotides and their mass

Entry	Duplex	T_m (°C) of DNA duplex different pH conditions		
		pH 6.0	pH 7.0, (Δ_{Tm}*)	pH 8.0
1	**ON1:ON2**	-	63.2	-
2	**ON1:ON3**	50.8	53.7 (-9.5)	54.3
3	**ON1:ON4**	40.9	47.4 (-15.8)	49.9
4	**ON1:ON5**	37.5	40.6 (-22.6)	47.5

*Δ_{Tm}: Difference in T_m of control DNA duplex (pH 7.0) and Tr-DNA duplex

UV/Fluorescence studies

Tropolone can accept and donate protons from carbonyl group and hydroxyl groups respectively. Under strong acidic conditions, tropolone undergoes protonation at carbonyl group and the equilibrium between protonated and neutral forms of tropolone exist at pK_a = -0.86.[35] However tropolone can also undergo deprotonation, the equilibrium between neutral form and deprotonated forms exists at pK_a = 6.9. Further, the neutral form also exists in rapid equilibrium with its tautameric form due to rapid intermolecular proton transfer (Figure 3A.4). Due to the ability of protonation, deprotonation and rapid proton transfers of tropolone, their photophysical properties are varied depending on the microenvironment. Tropolone also have fluorescence properties, although its quantum yield is low. Importantly, the fluorescence properties of tropolone are

influenced by polarity and pH conditions of solvents.[15] Fluorescence quantum efficiency of tropolone are 0.005 (in cyclohexane/neutral pH), 0.0025 (low pH conditions), and negligible ($0.005*10^{-2}$) in alkaline conditions (pH 12).[15] The UV-absorption peaks (~$\lambda_{250\ nm}$ and ~$\lambda_{400\ nm}$) are mainly due to the electronic transitions (π-π* & n-π*) and intramolecular proton transfer, though the fluorescence properties owing to the n-π* and intramolecular proton transfer.[15] Hence the photophysical properties of tropolone conjugated nucleosides are certainly influenced with solvent polarity and microenvironment within the nucleoside and their DNA oligonucleotides. Herein we also performed UV/Fluorescence studies of *tr*-nucleoside (**4** & **4-OH**) and *tr*-DNA (**ON3-ON5**) near physiological pH conditions. For control studies we performed the similar studies with free tropolone.

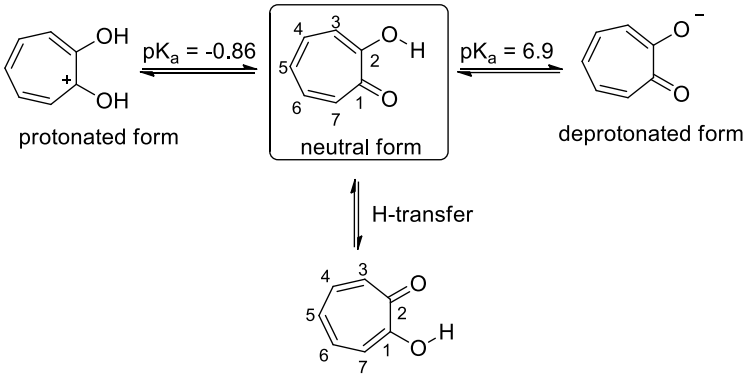


Figure 3A.4. Protonated, neutral and deprotonated forms of tropolone.

*pH dependent UV/Fluorescence studies of **4-OH***

The UV and emission spectra of tropolonyl-2'-deoxyuridine (**4-OH**) at different pH conditions are depicted in (Figure 3A.5), while spectra of free tropolone are provided in the Appendix. The UV spectra of nucleoside **4-OH** exhibit red shift and hyperchromic shift with increasing the pH from

3.7 to 10.8, whereas the emission spectra at λ_{ex} 395 nm exhibit the blue shift with remarkable depletion in the fluorescence intensity by increasing pH (3.7 to 10.8) conditions. The similar effects were noticed with control molecule (tropolone).

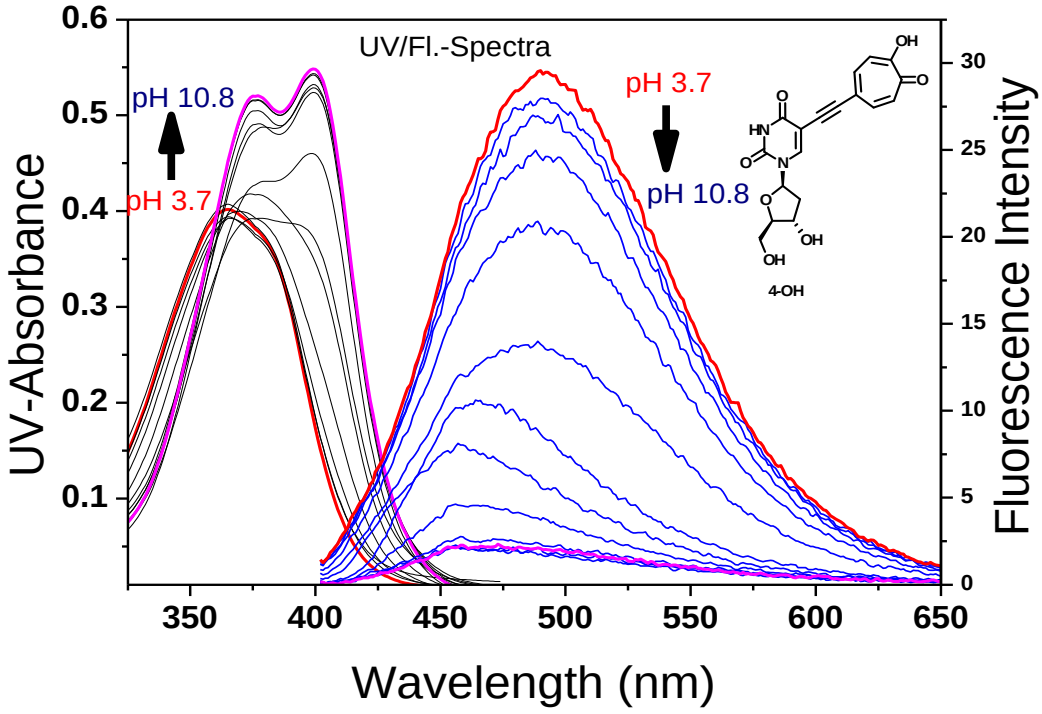


Figure 3A.5. The pH dependent UV absorbance and emission intensity of 22 µM tropolonyl thymidine (4-OH) in 100 mM buffer (pH 3.7-5.6, sodium acetate buffer; pH 6.0-8.0 phosphate buffer; pH 9.2-10.8 sodium carbonate buffer)

Here we extracted pH dependent UV-absorbance ($\lambda = 398$ nm) and fluorescence intensity ($\lambda = 490$ nm) values of nucleoside **4-OH**, then plotted pH profiles (Figure 3A.6) for determination of pK_a. The mid-points of sigmoidal pH profile is ~5.6 which is pK_a of nucleoside 4-OH. Similarly, we determined the pK_a of tropolone as ~6.9 for control and comparative studies (Appendix). Thus pK_a of nucleoside **4-OH** is lower than free tropolone molecule. This pK_a change in nucleoside (**4-OH**)

is attributed because of nucleobase residue which is known as acceptor for charge transfer process in fluorophore conjugated nucleosides.

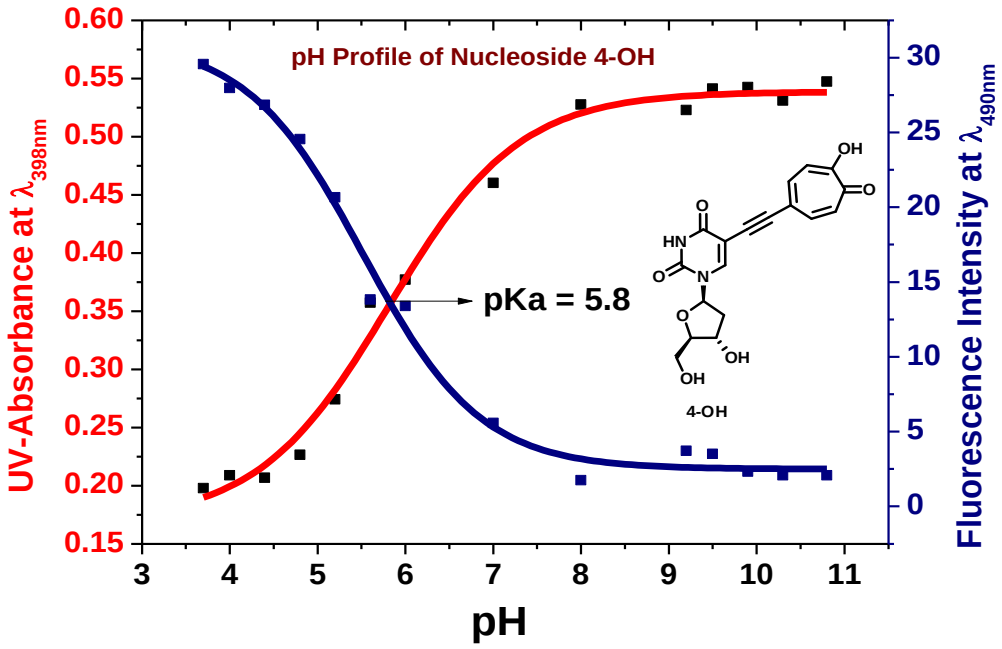


Figure 3A.6. pH dependent UV absorbance at 398 nm (red) and emission intensity at 490 nm of 22 µM tropolonyl thymidine (4-OH) in 100 mM buffer (pH 3.7-5.6, sodium acetate buffer; pH 6.0-8.0 phosphate buffer; pH 9.2-10.8 sodium carbonate buffer)

We recorded the UV spectra of single strand *tr*-DNA (**ON3**/**ON4**/**ON4**) and their duplex (**ON1:ON3**/**ON1:ON4**/**ON1:ON5)** with unmodified DNA (**ON1**) under different pH range (6.0-8.0). The UV spectra of *tr*-DNA duplex, at pH 7.0, are provided in Figure 3A.7, while the UV spectra of tr-DNA duplex at different pH are provided in Appendix. The UV-spectra of *tr*-DNA and its duplex exhibit two absorption peaks at λ_{250nm} and λ_{400nm} which are almost similar to the tropolone molecule near physiological pH range (6.0-8.0). The similar observations are noticed in UV spectra of other *tr*-DNA and their duplexes under different pH conditions (pH 6.0/8.0). Most

importantly, the UV spectra of *tr*-DNA duplex exhibit a little blue shift, ~$\lambda_{250\ nm}$ and ~$\lambda_{400\ nm}$, with increasing the *tr*-T nucleoside residue into *tr*-DNA duplex. This blue shift indicates some sort of π-π interactions with neighboring nucleobases.

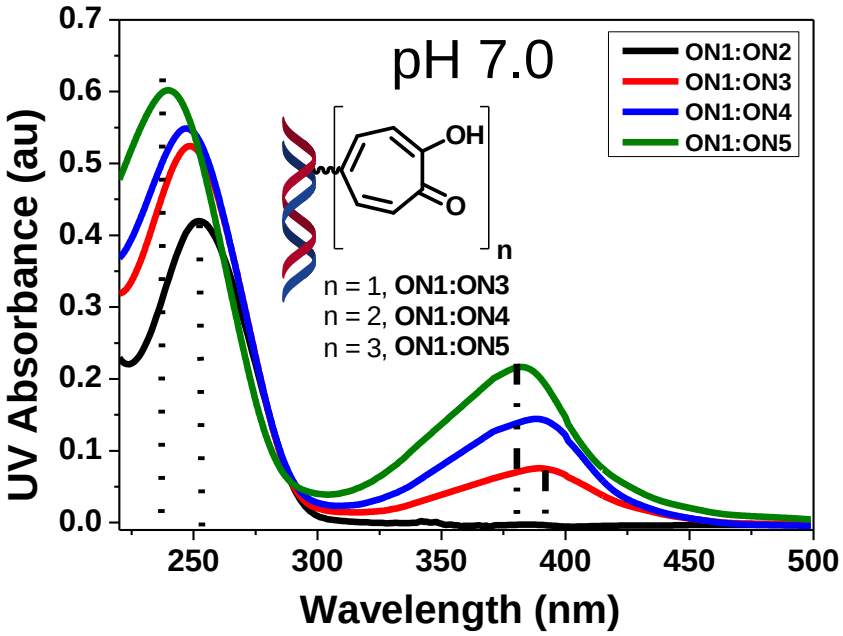


Figure 3A.7. UV spectra of *tr*-DNA duplex (**ON1:ON3**, **ON1:ON4**, **ON1:ON5**) and control under pH 6.0/7.0/8.0 conditions (sodium phosphate buffer) with 2.0 µM.

Further we determined the quantum yield of *mono*-tropolonylated DNA (**ON3**)**~**1.2% at excitation wavelength 400 nm ($\lambda_{ex.400\ nm}$) under aqueous pH 7.0 condition, using quinine sulphate as standard reference (0.1M H_2SO_4). The quantum yield (~1.2%) of *tr*-DNA (**ON3**) is significantly enhanced almost five folds as compare to tropolone molecule (~0.25%, at acidic pH 3.0). We recorded the fluorescence spectra of *tr*-DNA oligonucleotides (**ON3**/**ON4**/**ON5**) and their *tr*-DNA duplex (**ON1:ON3, ON1:ON4,** & **ON1:ON5**) near physiological pH range (6.0-8.0). The fluorescence spectra of *tr*-DNA (**ON3**) and its duplex **ON1:ON3** at different pH 6.0/7.0/8.0 are depicted in

Figure 3A.8, while fluorescence spectra of other *tr*-DNA (**ON4**/**ON5)** and their duplexes at different pH 6.0/7.0/8.0 are provided in the Appendix. In Figure 3A.8A, the fluorescence intensity of hybrid *tr*-DNA duplexes (**ON1**:**ON3, ON1:ON4,** & **ON1:ON5**) are remarkably enhanced as compare to single strand *tr*-DNA oligonucleotides (**ON3**/**ON4**/**ON5**). The fluorescence intensity of *tr*-DNA oligonucleotides (**ON3**/**ON4**/**ON5**) under same acidic pH (6.0) conditions is almost equal with little variations, especially with *tri*-tropolonylated DNA oligonucleotide (**ON5**). These results suggest that fluorescence efficiency of *tr*-DNA oligonucleotides is increased after formation of duplex. For comparative studies, we extracted fluorescence intensity value of *tr*-DNA oligonucleotides and their duplex under different pH conditions, and generated a bar-diagram as florescence intensity vs. *tr*-DNA (Figure 3A.8B). This bar diagram shows that the fluorescence efficiency of *tr*-DNA oligonucleotides (**ON3**/**ON4**/**ON5**) and their duplexes (**ON1**:**ON3, ON1:ON4,** & **ON1:ON5)** are enhanced at pH 6.0 conditions. These variations are more pronounced in *tr*-DNA duplexes as compare to single strand *tr*-DNA oligonucleotides. Importantly we also notice that fluorescence efficiency of *tr*-DNA oligonucleotides and their duplexes are marginally enhanced by increasing the *tr*-T residue in DNA.

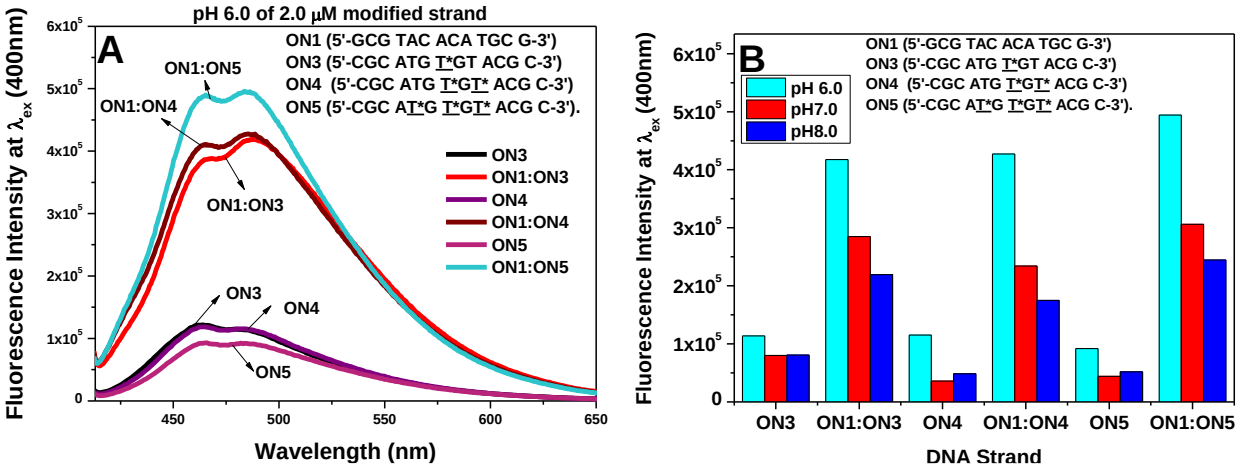


Figure 3A.8. Fluorescence spectra of single strand tr-DNA (**ON3/ON4/ON5**) and *tr*-DNA (**ON1:ON3/ON1:ON4/ON1:ON5**) under pH 6.0 (sodium phosphate buffer) conditions (A); Comparative bar diagram of fluorescence intensity vs *tr*-DNAs under different pH conditions.

Discussion

Since tropolone experiences intramolecular hydrogen bonding at acidic pH and neutral pH conditions;[15] and to explore the role of tropolonyl residue in the stability of DNA duplex structure, we, therefore, have synthesized tropolone conjugated thymidine nucleoside (**4-OH**) and its oligonucleotides (**ON3-ON5**). Since tropolone is excellent chromophore comprising weak fluorescence properties. So we also studied the photophysical properties of *tr*-*T* nucleoside (**4-OH**), *tr*-oligonucleotides (**ON3-ON5**) and *tr*-DNA duplex.

First, the CD spectral analyses suggest the formation of B-form type of *tr*-DNA duplex structures near physiological pH conditions as like control. Importantly, the perturbation of their CD signal are noticed with increasing *tr*-T nucleoside unit into DNA which reveal the destabilization of B-form type of DNA duplex structures. However, the new CD signal ($\lambda_{370\ nm}$) has noticed with *tri*-tropolonylated DNA duplex which is more prominent at pH 7.0/8.0. These CD studies suggest the conformational changes of tropolone in DNA duplex in different pH conditions.

Further the UV-melting results of *tr*-DNA duplex structures indicate that tropolone residue of *tr*-DNA duplex has critical role in destabilization of their duplex structures. Herein the stability of *tr*-DNA duplex increases with increasing the pH value i.e. more stable under pH 8.0 as compare to pH 6.0. The stability of *tr*-DNA duplex, however, decreases with increasing the number of *tr*-T nucleoside. In literature, nucleoside conjugated chromophores of DNA are involved in π-π

interactions with neighboring nucleobase residues.[36, 37] Moreover, the destabilization of DNA B-form duplex structures are reportedly known because of the π-π interactions between conjugated chromophore and neighboring nucleobases.[26] Tropolone residue of *tr*-DNA duplex, therefore, may have similar types of π-π interactions with neighboring G-nucleobase. The increment in *tr*-T nucleosides is probably enhanced the perturbation in base stacking structure of DNA duplex which are responsible for destabilization of *tr*-DNA duplex structure.

The fluorescence character of tropolone is appeared because of the intramolecular charge transfer (n-π* transition) which is sensitive to the pH and polarity of environments.[15] Firstly, we determined the pK_a of nucleoside (**4-OH**) and control tropolone. We noticed the pK_a of tropolone residue of *tr*-T nucleoside (**4-OH**) is marginally lower (~0.9) than free tropolone, possibly due the presence of conjugated thymine nucleobase which is known as acceptor in charge transfer molecular system. Thus charge transfer ability is enhanced in thymine conjugated tropolone nucleoside **4-OH**. Hence fluorescence property of *tr*-T (**4-OH**) is relatively higher than free tropolone.

The UV spectra of nucleoside (**4-OH**) exhibit red shift with increasing the pH value (pH 3.7-10.7), while its fluorescence spectra shows slight blue shift with remarkable depletion in fluorescence intensity with increasing the pH value (pH 3.7-10.7). The pK_a of nucleoside **4-OH** is also extracted from pH dependent fluorescence profile, which is same as from pH dependent UV absorbance profile (pK_a = 5.8±0.2). Since tropolone exists in different structural forms, as neutral form near neutral pH and anionic form near alkaline pH. The photophysical properties of such molecules are severely affected with pH and polarity of environments.

To support the fluorescence character of nucleoside *tr*-T (**4-OH**), we optimized the structure of nucleoside (**4-OH**) under neutral and anionic forms of tropolone residue by DFT (B3YLP-6, 31-6G*) calculation, and then pleasently extracted their frontier molecular orbitals (HOMO-LUMO) and their energy level (Figure 3A.9). In neutral form of nucleoside (**4-OH**), both HOMO and LUMO are localized on tropolone and extended up to thymine residue with an energy difference (ΔE) of 3.31 eV. In contrast, in anionic form of **4-OH**, the HOMO is localized on tropolone residue, while LUMO is localized on thymine residue but their energy difference (ΔE = 2.352 eV) is lower than its neutral form (ΔE = 3.31 eV). Generally delocalization of HOMO-LUMO decreases ΔE and enhances fluorescence character. It is also well known that when HOMO-LUMO are localized and decreases energy gap, then the decay of excited state to ground state dominates through radiationless pathways which results in lowering of fluorescence.[28, 38, 39] These calculations support the pH dependent UV/fluorescence behaviour of nucleoside **4-OH** under different pH conditions having different sets of HOMO-LUMO under different pH conditions.

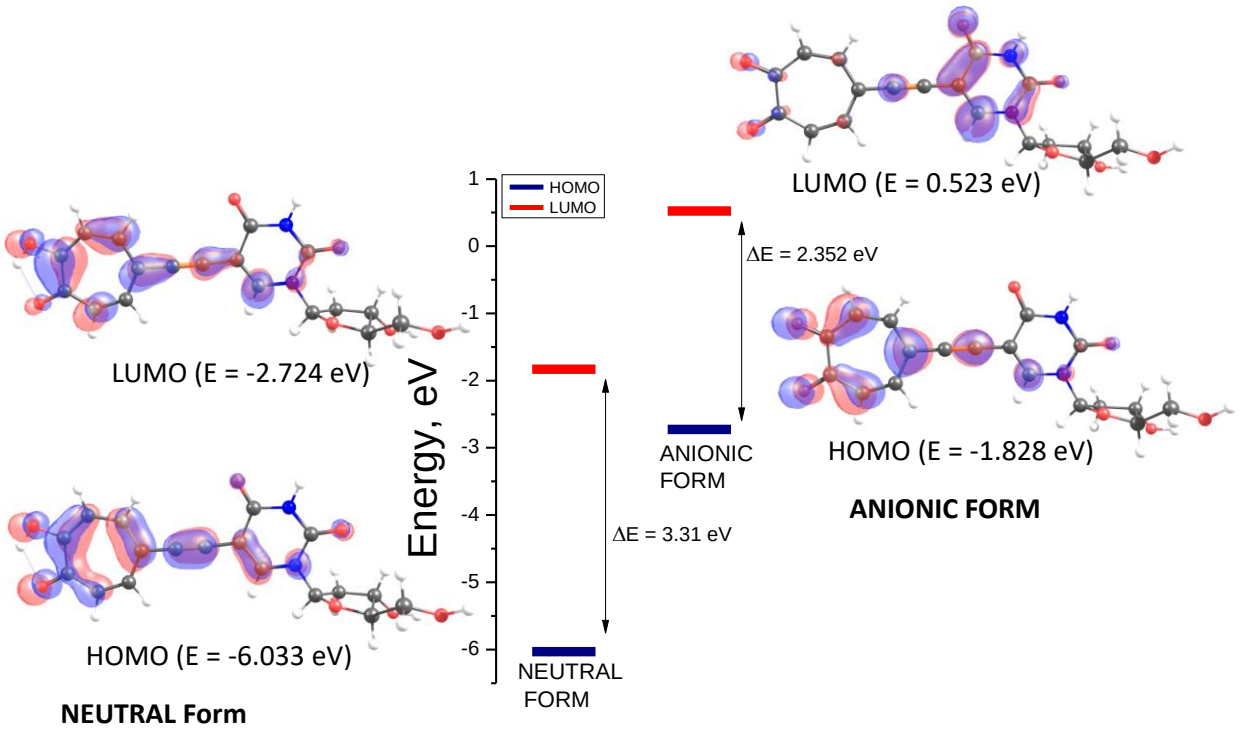


Figure 3A.9. HOMO-LUMO and their calculated energies in neutral form and anionic form of tropolonylated thymidine nucleoside (**4-OH**). Calculated using DFT-B3LYP-6 from their optimized structures, neutral form (left) anionic form (right).

The fluorescent characters of intramolecular charge transfer fluorophores are influenced with polarity of solvent and enhanced in non-polar solvent.[40] The intramolecular charge transfer carrying fluorophores also exhibit pH dependent fluorescence character and explained by that mechanism.[28] Further the fluorescence character of fluorophore due to intramolecular charge transfer, especially because of push-pull systems, is significantly higher in non-polar solvents as compare to polar solvents. Importantly the fluorescence quantum yield of tropolone is higher in hexane (non-polar solvent) as compare to aqueous (pH 7.0). Thus by changing the microenvironment of nucleoside (**4-OH**) could affect its fluorescence properties. Thus the quantum yield of *tr*-DNA oligoucleotide (**ON3**) is enhanced in aqueous medium near

physiological pH conditions as compared to free tropolone molecule, and further enhanced even by slight lowering the pH conditions possibly because of push-pull mechanism (Table S1). At pH 6.0, the intramolecular hydrogen bonded tropolone (neutral form) are relatively populated in *tr*-DNA oligonucleotide and experience non-polar microenvironment.

The DNA duplex (B-form) has non-polar microenvironment due to the presence of grooves (major & minor) in aqueous medium. CD results of *tr*-DNA duplex strongly support the formation of B-form helical structure. Thus the fluorescence character of *tr*-DNA duplex may affect in different pH conditions due to the presence of fluorophore nucleoside, *tr*-T in one of the strand. Presumably, tropolone ring of *tr*-DNA duplex (**ON1:ON3**) is occurred in non-polar microenvironment and experience π-π interactions with neighboring nucleobases guanosine. Such interactions affect Watson-Crick (W-C) hydrogen bonding between complementary nucleobases and play critical role in stability of DNA duplex structure.[26] Since π-π attractive interactions between neutral aromatic ring and nucleobase ring system suppress the W-C hydrogen bonding ability at pyrimidine nucleobases (T/U/C), while π-π repulsive interactions between anionic aromatic ring and nucleobase ring system enhance the W-C hydrogen bonding ability at pyrimidine nucleobases (T/U/C).[41] More W-C hydrogen bonding quenches fluorescence properties of ICT-fluophore conjugated DNA at pyrimidine nucleobases. At pH 6.0, the population of neutral form of tropolone structure in *tr*-DNA duplex is relatively higher. In contrast, at pH 8.0, the population of anionic form of tropolone is high. Thus π-π attractive interactions between neutral tropolone ring and neighboring nucleobases are relatively higher at low pH 6.0, and consequently decrease the W-C hydrogen bonding in *tr*-DNA duplex. As resultant, fluorescence character as quantum yield is enhanced at low pH but reduced the stability of DNA structures. In contrast, the π-π repulsive interactions between anionic tropolone ring and neighboring nucleobases are relatively high at pH

8.0, and consequently increase the W-C hydrogen bonding in *tr*-DNA duplex. As resultant, the fluorescence character as quantum yield is significantly reduced in *tr*-DNA at pH 8.0 and enhanced the stability of DNA structures. Further, the increasing *tr*-T nucleoside residue in *tr*-DNA duplex, the B-form helical structures significantly destabilize at pH 6.0 but their fluorescence properties exhibit almost same. However, at pH 8.0 fluorescence properties of di-/tri-incorporated *tr*-T nucleoside are remarkably reduced. The stability of three *tr*-T incorporated *tr*-DNA duplex structure is relative stable at pH 8.0 as compare to pH 6.0. Moreover, *tr*-DNA duplex structures destabilize with increasing the *tr*-T nucleoside and the similar trends are consistent under all the three pH conditions. The strength of W-C hydrogen bonding is relatively more in control DNA duplex as compared to *tr*-T incorporated *tr*-DNA duplex.

3A.4 Conclusion

We have successfully synthesized the tropolone conjugated *tr*-T nucleoside and DNA oligonucleotides. Since tropolone is excellent non-benzenoid chromophores comprising pH dependent weak fluorescent properties. We studied UV/fluorescence properties of *tr*-T and determined the pK_a 5.9 of nucleoside which is lower than control tropolone (pK_a 6.8). We have accomplished pH dependent UV/fluorescence properties of tropolone conjugated DNA nucleoside/oligonucleotides/DNA duplexes. Our results support the enhancement of quantum yield of oligonucleotides after the formation of DNA duplex structure (B-form helix) and that is more pronounced at low pH (6.0) as compare to pH 7.0/8.0. However, the stability of tr-DNA duplex structures is relatively higher at pH 8.0 as compare to pH 6.0. The increasing the number of *tr*-T incorporation in DNA oligonucleotides destabilize their DNA duplex B-form helical structure under all pH conditions. We demonstrate that *Tr*-T nucleoside and its oligonucleotides

comprise pH dependent UV-Fluorescence properties. Which are further supported by DFT calculation Hence *tr*-T DNA nucleoside and its oligonucleotides are new fluorescent DNA analogues which are pH sensitive and enabling the sensing of duplex formation.

3A.5 Experimental Section

All required materials were obtained from commercial suppliers and used without any further purification. Pyridine, methylene chloride, was distilled over calcium hydride and stored on 4Å molecular sieves. Tetrahydrofuran was freshly distilled over sodium /benzophenone. Reactions were monitored by thin layer chromatography, visualized by UV and ninhydrin. Purifications of compounds using column chromatography were performed with 230-400 silica gel. ESI-Mass spectra and HRMS were obtained from Bruker micrOTOF-Q II Spectrometer. ^{1}H NMR, ^{13}C NMR, and ^{31}P NMR spectra were recorded on Bruker AV-400 or 700 MHz at 298 K. ^{1}H and ^{13}C NMR chemical shifts were recorded in ppm downfield from tetramethylsilane or residual solvent peak. Splitting patterns are abbreviated as: s, singlet; d, doublet; dd, doublet of doublet; t, triplet; q, quartet; dq, doublet of quartet; m, multiplet.

5- (2-benzoyloxy tropone) ethynyl-2'-deoxyuridine (**4**): 5-iodo-2-benzoyloxytropone **3** (1.19 gm , 3.38 mmol) and $Pd(PPh_3)_4$ (0.195 gm, 0.169 mmol) was dissolved in mixture of Dry DMF (30 ml) and Et_3N (0.7ml, 0.05 mmol) then added CuI (0.289 gm, 1.52 mmol) and 5-ethynyl deoxyuridine **2** (0.938 gm 3.7 mmol). The reaction mixture stirred at room temperature for 6 h, then reaction mixture was filtered and the filtrate was concentrated to give a brown crude residue, which was purified by flash column chromatography (0.5 % Et_3N/10 % MeOH/$CHCl_3$) to give **4** as a yellow solid (1.193 gm, 74% yield). R_f 0.36 (10% MeOH/$CHCl_3$); ^{1}H NMR (400 MHz,

DMSO, & 1-2 drops Et_3N) δ 11.78 (s, 1H), 8.53 (s, 1H), 8.08 (d, *J* = 7.3 Hz, 2H), 7.77 (t, *J* = 7.4 Hz, 1H), 7.61 (t, *J* = 7.7 Hz, 2H), 7.41 (s, 3H), 6.12 (t, *J* = 6.3 Hz, 1H), 5.27 (d, *J* = 4.3 Hz, 1H), 5.20 (t, *J* = 4.7 Hz, 1H), 4.33 – 4.22 (m, 1H), 3.83 (d, *J* = 3.3 Hz, 1H), 3.75 – 3.54 (m, 2H).^{13}C NMR (101 MHz, DMSO) δ 163.25, 161.16, 149.33, 145.42, 134.30, 129.94, 129.04, 128.56, 128.26, 97.12, 93.50, 87.84, 87.64, 85.12, 69.68, 60.66, 40.29. HRMS (ESI-TOF) m/z: $[M + Na]^+$ Calcd for $C_{22}H_{20}N_2O_8Na$ 499.1112; Found 499.1135.

Tropolonylated 2'-deoxyuridine ***(4-OH)***. Compound **4** (15 mg, 0.031 mmol) was treated with 1,4-dioxane and ammonium hydroxide (1:1), 0.6 ml and stirred at room temperature for 1.5 h. The solvents were evaporated and resulting solid is washed with ether and EtOAc each time and dissolved in water and adjusted to pH 6-7 with 1N HCl and dried under high vacuum to get title compound. ^{1}H-NMR (400 MHz, DMSO) δ 12.19 – 10.88 (b, 1H), 8.42 (s, 1H), 7.10 (s,b, 2H), 6.10 (d, J = 2.9 Hz, 1H), 5.32 (d, J = 30.0 Hz, 2H), 4.26 (s, 2H), 3.81 (d, J = 2.8 Hz, 1H), 3.58 (dd, J = 18.9, 7.7 Hz, 2H), 2.14 (t, 2H). ^{13}C-NMR (176 MHz, DMSO) δ 172.59, 161.33, 149.40, 144.11, 139.01, 124.14, 97.99, 87.66, 84.94, 69.89, 60.74, 40.26. HRMS (ESI-TOF) m/z: $[M - H]^-$ Calcd for $C_{18}H_{15}N_2O_7$, 371.0871; Found 371.1204.

5'-DMTr-(*O*-benzoyl-tropolonylated)-2'-deoxyuridine (**5**). Compound 4 (0.94 gm, 1.96 mmol) was co-evaporated with anhydrous pyridine for 3 times and then dissolved in dry CH_2Cl_2 (20 ml). To the resultant solution anhydrous Et_3N (1.36 ml, 9.84 mmol) followed by DMTr-Cl (1.0 g, 2.95 mmol) was added and stirred at room temperature for 7 h. The crude reaction mixture evaporated to dryness and then subjected to column chromatography (0.5% Et_3N/ 2.5% MeOH/CH_2Cl_2) to yield 0.9 gm (59%) as yellow solid. Rf 0.38 (2.5% MeOH/ CH_2Cl_2); ^{1}H NMR (400 MHz, $CDCl_3$& 1-2 drops Et_3N) δ 8.55 (s, 1H), 8.13 (d, J = 7.5 Hz, 2H), 7.61 (t, J = 7.3 Hz, 1H), 7.51 – 7.43 (m, 4H), 7.38 – 7.26 (m, 6H), 7.18 (d, J = 7.2 Hz, 1H), 6.81 (dd, J = 8.8, 2.5 Hz, 6H), 6.42 (t, J = 8.5

Hz, 3H), 4.57 (s, 1H), 4.10 (s, 1H), 3.81 – 3.50 (m, 8H), 3.19 (d, J = 8.7 Hz, 1H), 2.54 (dd, J = 8.1, 2.9 Hz, 1H), 2.45 – 2.31 (m, 1H).^{13}C NMR (101 MHz, $CDCl_3$) δ 163.83, 161.01, 158.75, 158.74 149.04, 144.61, 143.94, 135.78, 135.34, 133.84, 130.48, 129.94, 129.83, 128.69, 128.57, 128.47, 128.19, 127.85, 127.25, 113.46, 113.43, 99.53, 94.90, 87.18, 86.94, 86.01, 84.94, 71.91, 63.2, 55.30, 42.04. HRMS (ESI-TOF) m/z: [M + Na], Calcd for $C_{46}H_{38}N_2O_{10}Na$, 801.2419; Found 801.2394.

tropolonylated-2'-deoxyuridine phosphoramidites (**6**): To a stirred solution of compound **5** (0.2 g, 0.25 mmol) in anhydrous THF (1.5 ml) and N,N-Diisopropyl ethylamine (0.27 ml, 15.4 mmol) was treated with 2-cyanoethyl diisopropylchlorophosphoramidite (0.15 ml, 5.13 mmol) at room temperature. After stirred for 3.5 h subjected to column chromatography (1% Et_3N /60% EtOAc/ n-Hexanes) to yield 0.150 gm (61%) as pale yellow foam. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.61 (d, J = 13.0 Hz, 1H), 8.14 (d, J = 7.3 Hz, 2H), 7.62 (t, J = 7.4 Hz, 1H), 7.48 (t, J = 8.0 Hz, 4H), 7.40 – 7.27 (m, 5H), 7.24 – 7.13 (m, 1H), 6.81 (dt, J = 8.8, 3.6 Hz, 6H), 6.42 – 6.29 (m, 3H), 4.81 – 4.47 (m, 1H), 4.31 – 4.08 (m, 1H), 3.92 – 3.44 (m, 11H), 3.18 (dd, J = 8.3, 3.2 Hz, 1H), 2.61 (dd, J = 11.8, 5.5 Hz, 2H), 2.53 – 2.32 (m, 2H), 1.15 (dd, J = 9.0, 3.8 Hz, 8H), 1.02 (d, J = 6.8 Hz, 3H), 0.92 – 0.81 (m, 3H). ^{31}P NMR (162 MHz, $CDCl_3$) δ 149.67 (s), 149.25 (s). HRMS (ESI-TOF) m/z: [M + Na] Calcd for $C_{55}H_{55}N_4O_{11}PNa$ 1001.3497; Found 1001.3522.

3A.6 Reference

1. Burnham, D.; Cook, M., HMO calculations on [4, 5c]-furotropone and related systems. *Tetrahedron Lett.* **1968,** *9*, 3771-3774.

2. Liu, N.; Song, W.; Schienebeck, C. M.; Zhang, M.; Tang, W., Synthesis of naturally occurring tropones and tropolones. *Tetrahedron* **2014,** *70*, 9281-9305.

3. Bentley, R., A fresh look at natural tropolonoids. *Nat. Prod. Rep.* **2008,** *25*, 118-38.

4. Cox, R. J.; Al-Fahad, A., Chemical mechanisms involved during the biosynthesis of tropolones. *Curr. Opin. Chem. Biol.* **2013,** *17*, 532-536.

5. Saleh, N.; Zwiefak, A.; Mordarski, M.; Pulverer, G., Antibacterial activity of selected tropones and tropolones. *Zentralbl Bakteriol Mikrobiol. Hyg. A* **1988,** *270*, 160-170.

6. Angawi, R. F.; Swenson, D. C.; Gloer, J. B.; Wicklow, D. T., Malettinin A: a new antifungal tropolone from an unidentified fungal colonist of Hypoxylon stromata (NRRL 29110). *Tetrahedron Lett.* **2003,** *44*, 7593-7596.

7. Yang, P. S.; Wang, M. J.; Jayakumar, T.; Chou, D. S.; Ko, C. Y.; Hsu, M. J.; Hsieh, C. Y., Antiproliferative Activity of Hinokitiol, a Tropolone Derivative, Is Mediated via the Inductions of p-JNK and p-PLCgamma1 Signaling in PDGF-BB-Stimulated Vascular Smooth Muscle Cells. *Molecules* **2015,** *20*, 8198-8212.

8. Lomonosova, E.; Daw, J.; Garimallaprabhakaran, A. K.; Agyemang, N. B.; Ashani, Y.; Murelli, R. P.; Tavis, J. E., Efficacy and cytotoxicity in cell culture of novel alpha-hydroxytropolone inhibitors of hepatitis B virus ribonuclease H. *Antivir. Res.* **2017,** *144*, 164-172.

9. Fullagar, J. L.; Garner, A. L.; Struss, A. K.; Day, J. A.; Martin, D. P.; Yu, J.; Cai, X.; Janda, K. D.; Cohen, S. M., Antagonism of a zinc metalloprotease using a unique metal-chelating scaffold: tropolones as inhibitors of P. aeruginosa elastase. *Chem. Commun.* **2013,** *49*, 3197-3199.

10. Bryant, B. E.; Fernelius, W. C.; Douglas, B. E., Formation Constants of some Metal-Tropolone Complexes. *Nature* **1952,** *170*, 247-248.

11. Muetterties, E.; Roesky, H.; Wright, C., Chelate Chemistry. V. Metal Chelates Based on Tropolone and Its Derivatives. *J. Am. Chem. Soc.* **1966,** *88*, 4856-4861.

12. Hosoya, H.; Tanaka, J.; Nagakura, S., Ultra-violet absorption spectra of tropone, troponium ion, tropolone and 2, 4, 6-octatrienal. *Tetrahedron* **1962,** *18*, 859-874.

13. Yamaguchi, H.; Amako, Y.; Azumi, H., The electronic spectrum of the tropone molecule. *Tetrahedron* **1968,** *24*, 267-277.

14. Doering, W. v. E.; Knox, L. H., Tropolone. *J. Am. Chem. Soc.* **1951,** *73*, 828-838.

15. Breheret, E.; Martin, M., Electronic relaxation of troponoids: tropolone fluorescence. *J. Lumin.* **1978,** *17*, 49-60.

16. Balachandra, C.; Sharma, N. K., Synthesis and conformational analysis of new troponyl aromatic amino acid *Tetrahedron* **2014,** *70*, 7464.

17. Balachandra, C.; Sharma, N. K., Instability of Amide Bond Comprising the 2-Aminotropone Moiety: Cleavable under Mild Acidic Conditions. *Org. Lett.* **2015,** *17*, 3948-51.

18. Balachandra, C.; Sharma, N. K., Novel fluorophores: Syntheses and photophysical studies of boron-aminotroponimines. *Dyes Pigm.* **2017,** *137*, 532-538.

19. Jean, J. M.; Hall, K. B., 2-Aminopurine fluorescence quenching and lifetimes: role of base stacking. *Proc. Natl. Acad. Sci. U S A* **2001,** *98*, 37-41.

20. Ward, D.; Reich, E.; Stryer, L., Fluorescence studies of nucleotides and polynucleotides I. Formycin, 2-aminopurine riboside, 2, 6-diaminopurine riboside, and their derivatives. *J. Biol. Chem.* **1969,** *244*, 1228-1237.

21. Datta, K.; Johnson, N. P.; von Hippel, P. H., DNA conformational changes at the primer-template junction regulate the fidelity of replication by DNA polymerase. *Proc. Natl. Acad. Sci. U S A* **2010,** *107*, 17980-17985.

22. Shin, D.; Sinkeldam, R. W.; Tor, Y., Emissive RNA alphabet. *J. Am. Chem. Soc.* **2011,** *133*, 14912-14915.

23. Rovira, A. R.; Fin, A.; Tor, Y., Chemical mutagenesis of an emissive RNA alphabet. *J. Am. Chem. Soc.* **2015,** *137*, 14602-14605.

24. Shi, Y.; Dierckx, A.; Wanrooij, P. H.; Wanrooij, S.; Larsson, N.-G.; Wilhelmsson, L. M.; Falkenberg, M.; Gustafsson, C. M., Mammalian transcription factor A is a core component of the mitochondrial transcription machinery. *Proc. Natl. Acad. Sci. U S A* **2012,** *109*, 16510-16515.

25. Koga, Y.; Fuchi, Y.; Nakagawa, O.; Sasaki, S., Optimization of fluorescence property of the 8-oxodGclamp derivative for better selectivity for 8-oxo-2′-deoxyguanosine. *Tetrahedron* **2011,** *67*, 6746-6752.

26. Burns, D. D.; Teppang, K. L.; Lee, R. W.; Lokensgard, M. E.; Purse, B. W., Fluorescence turn-on sensing of DNA duplex formation by a tricyclic cytidine analogue. *J. Am. Chem. Soc.* **2017,** *139*, 1372-1375.

27. Rodgers, B. J.; Elsharif, N. A.; Vashisht, N.; Mingus, M. M.; Mulvahill, M. A.; Stengel, G.; Kuchta, R. D.; Purse, B. W., Functionalized tricyclic cytosine analogues provide nucleoside fluorophores with improved photophysical properties and a range of solvent sensitivities. *Chem. Eur. J.* **2014,** *20*, 2010-2015.

28. Mata, G.; Schmidt, O. P.; Luedtke, N. W., A fluorescent surrogate of thymidine in duplex DNA. *Chem. Commun.* **2016,** *52*, 4718-4721.

29. Herdewijn, P., Heterocyclic modifications of oligonucleotides and antisense technology. *Antisense. Nucleic. Acid. Drug. Dev.* **2000,** *10*, 297-310.

30. Seo, S.; Onizuka, K.; Nishioka, C.; Takahashi, E.; Tsuneda, S.; Abe, H.; Ito, Y., Phosphorylated 5-ethynyl-2'-deoxyuridine for advanced DNA labeling. *Org. Biomol. Chem.* **2015,** *13*, 4589-4595.

31. Potenziano, J.; Spitale, R.; Janik, M. E., Improved and Highly Versatile Synthesis of 5-Aryltropones. *Syn. Commun.* **2005,** *35*, 2005-2016.

32. Hwang, G. T.; Hari, Y.; Romesberg, F. E., The effects of unnatural base pairs and mispairs on DNA duplex stability and solvation. *Nucleic Acids Res.* **2009,** *37*, 4757-4763.

33. Kypr, J.; Vorlíčková, M., Circular dichroism spectroscopy reveals invariant conformation of guanine runs in DNA. *Biopolymers* **2002,** *67*, 275-277.

34. Schroeder, S. J.; Turner, D. H., Optical melting measurements of nucleic acid thermodynamics. *Methods Enzymol.* **2009,** *468*, 371-387.

35. Hosoya, H.; Nagakura, S., The Basicities of Tropone and Tropolone. *Bull. Chem. Soc. Jpn.* **1966,** *39*, 1414-1418.

36. Lewis, F. D.; Liu, X.; Wu, Y.; Miller, S. E.; Wasielewski, M. R.; Letsinger, R. L.; Sanishvili, R.; Joachimiak, A.; Tereshko, V.; Egli, M., Structure and photoinduced electron transfer in exceptionally stable synthetic DNA hairpins with stilbenediether linkers. *J. Am. Chem. Soc.* **1999,** *121*, 9905-9906.

37. Egli, M.; Tereshko, V.; Mushudov, G. N.; Sanishvili, R.; Liu, X.; Lewis, F. D., Face-to-Face and Edge-to-Face π− π Interactions in a Synthetic DNA Hairpin with a Stilbenediether Linker. *J. Am. Chem. Soc.* **2003,** *125*, 10842-10849.

38. De Ornellas, S.; Slattery, J. M.; Edkins, R. M.; Beeby, A.; Baumann, C. G.; Fairlamb, I. J., Design and synthesis of fluorescent 7-deazaadenosine nucleosides containing π-extended diarylacetylene motifs. *Org. Biomol. Chem.* **2015,** *13*, 68-72.

39. Zilbershtein-Shklanovsky, L.; Weitman, M.; Major, D. T.; Fischer, B., Rules for the design of highly fluorescent nucleoside probes: 8-(substituted cinnamyl)-adenosine analogues. *J. Org. Chem.* **2013,** *78*, 11999-12008.

40. Suzuki, A.; Nemoto, N.; Saito, I.; Saito, Y., Design of an environmentally sensitive fluorescent 8-aza-7-deaza-2′-deoxyadenosine derivative with dual fluorescence for the specific detection of thymine. *Org. Biomol. Chem.* **2014,** *12*, 660-666.

41. Mignon, P.; Loverix, S.; Steyaert, J.; Geerlings, P., Influence of the π–π interaction on the hydrogen bonding capacity of stacked DNA/RNA bases. *Nucleic Acids Res.* **2005,** *33*, 1779-1789.

3A.7 Appendix

1. $^{1}H/^{13}C$-NMR & HRMS spectra of 2'-benzoyl-tropolonylated-2-deoxyuridine (4)

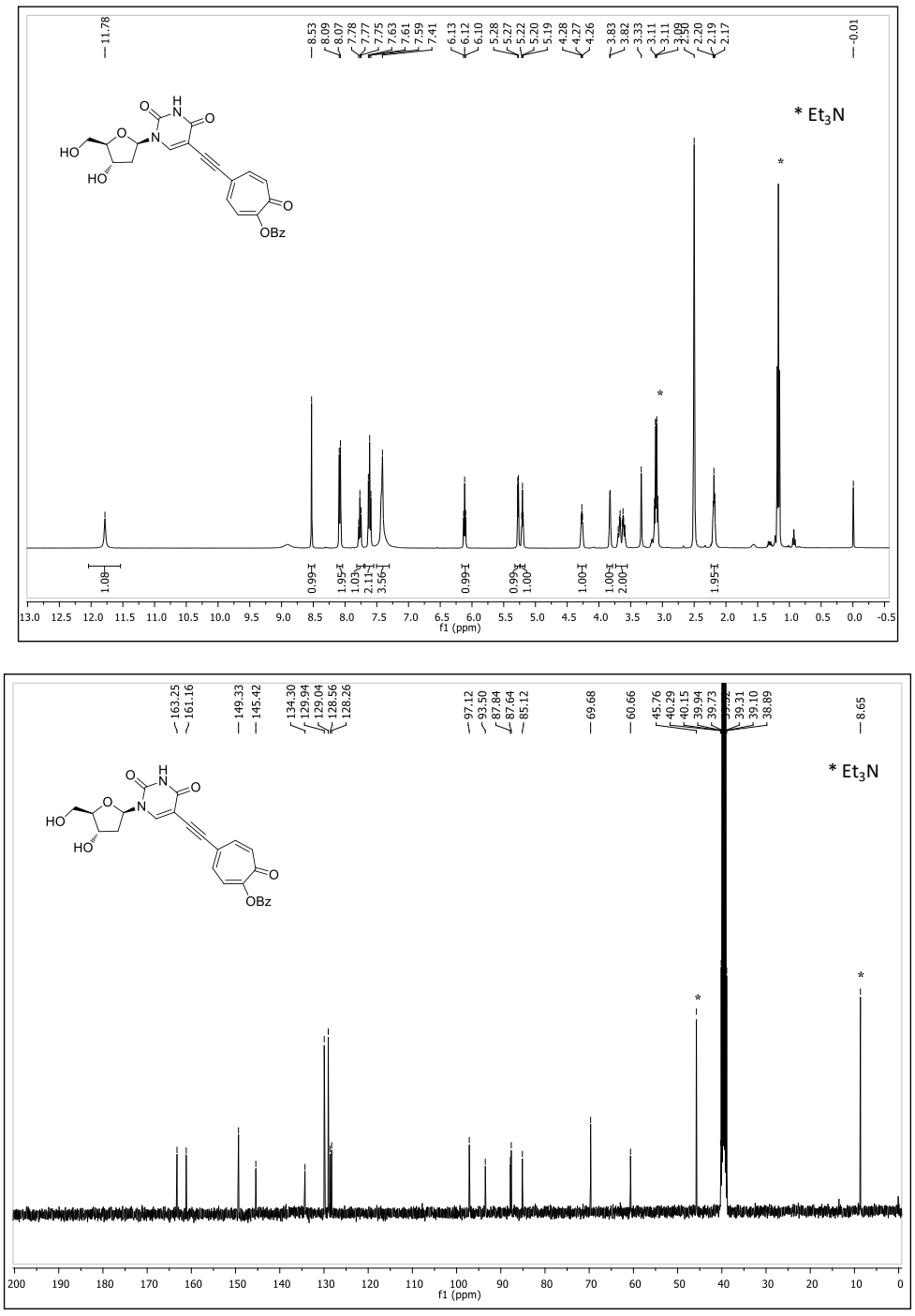


Figure S1. ^{1}H-/^{13}C-NMR of 2-benzoyl-tropolonylated-2'-deoxyuridine (**4**) in DMSO-d_6

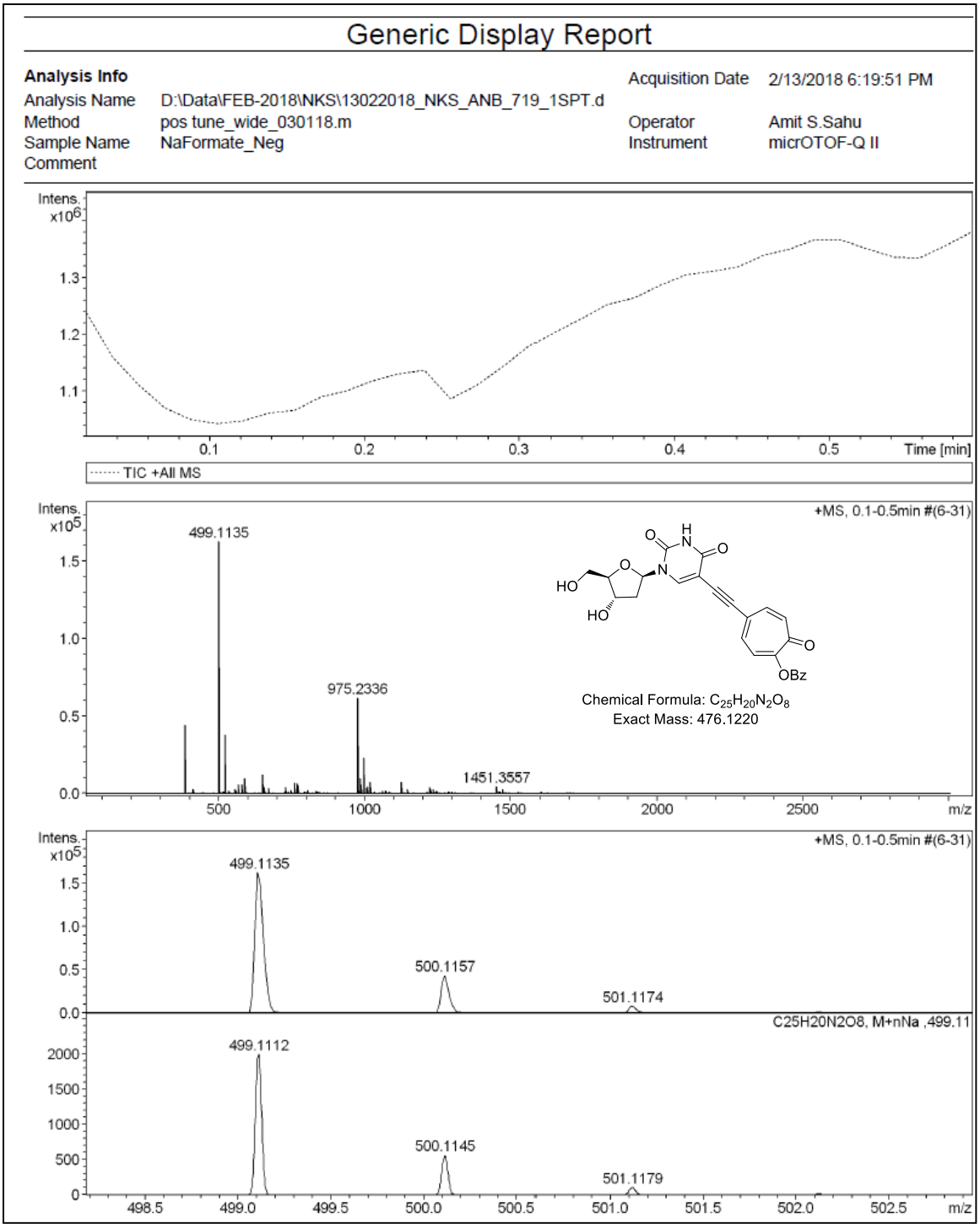


Figure S2. ESI-HRMS of 2-benzoyl-tropolonylated-2'-deoxyuridine (**4**)

2.X-ray data analyses of Single crystal **4**

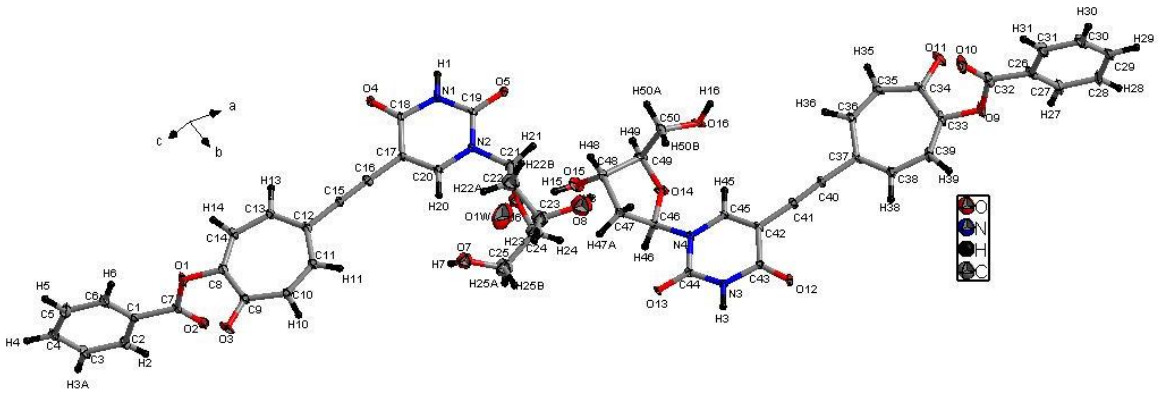


Figure S3. ORTEP diagram of compound **4** (Z =1)

The crystals quality were poor even after growing several times. The data did not improve even collecting it at liquid nitrogen temperatures. The sugar ring is disordered but could be solved in this space group to satisfaction. Therefore, the ring carbon atoms were isotopically refined. A peak of 1.10 eA-3 was found near the ring carbon atom could not be assigned to any entity and left as such

Single crystal data analyses of compound **4***:*

Parameter	**Compound 4**
Formula	C25 H20 N2 O8
*M*r	289.33
crystal system	Triclinic
space group	*P 1*
Cell Length/Å	*a*5.8443(2), *b*14.3718(6), *c*15.3063(6),
Cell angles/^{O}	α117.262(2), β92.910(3)γ99.549(2)
V/Å^3	1115.14(8)
Z	1
ρ_{calc} (gcm^{-3})	1.446
μ/ mm^{-1} μ(cm^{-1})	0.110
θ range (deg)	2.67 to 25.50°.
reflections collected	20099
Unique reflections[R(int)]	7524 [0.0376]
data/restraints/ parameters	7524 / 11 / 604
GOF on F^2	1.294
*R*1 and *R*2 [$I>2\sigma(I)$]	0.0873, 0.2168
*R*1 and *R*2 (all data)	0.1035, 0.2302
largest residual peaks (e.Å^{-3})	1.095 and -0.597
CCDC no	CCDC 1830067

3.^{1}H/^{13}C-/^{1}H-COSY/ D_2O exchange and suppression HRMS spectra of *tr-2'-deoxyuridine* **(4-OH)**

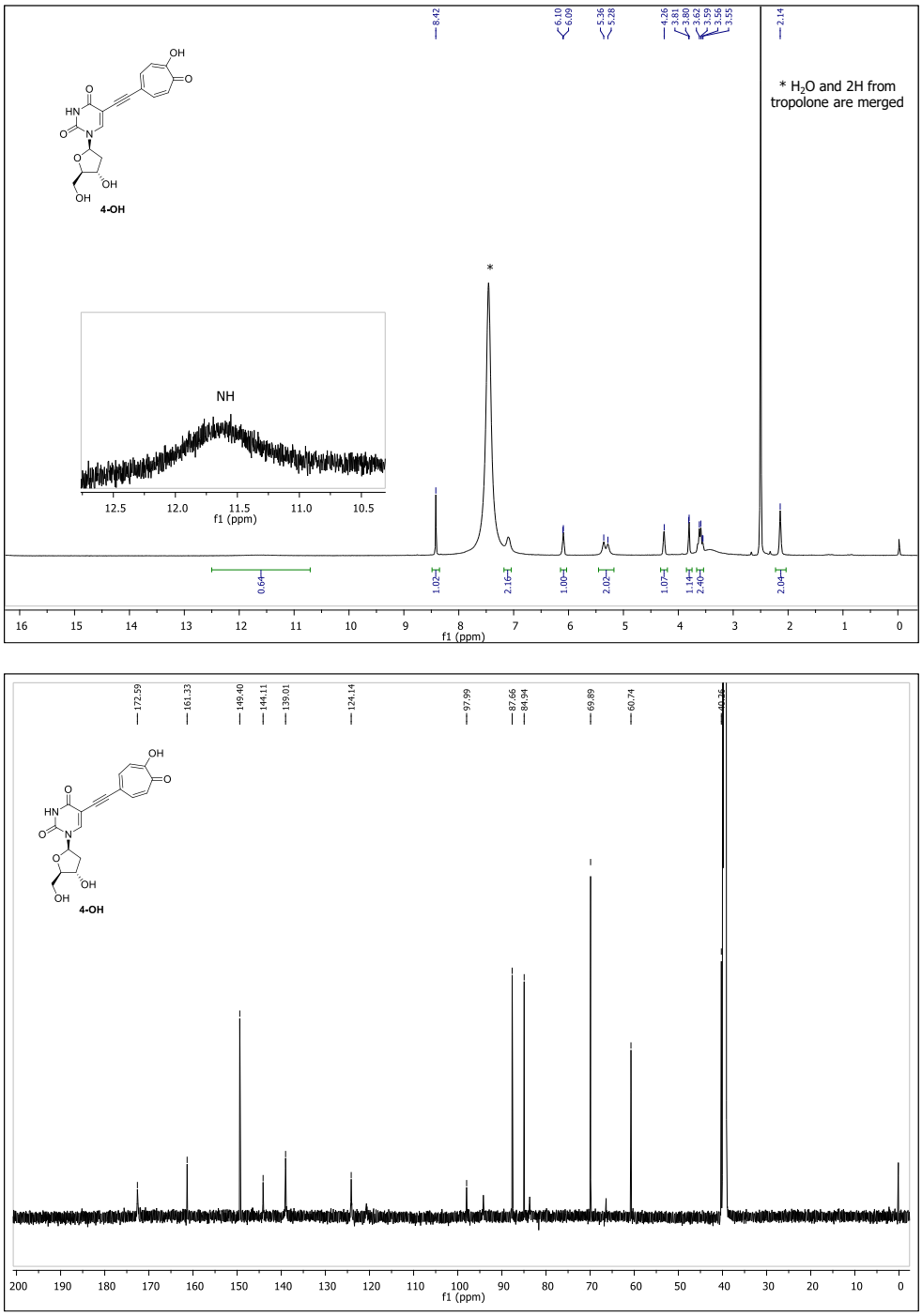


Figure S4. ^{1}H-/ ^{13}C-NMR of tropolonylated-2'-deoxyuridine (**4-OH**) in DMSO-d6

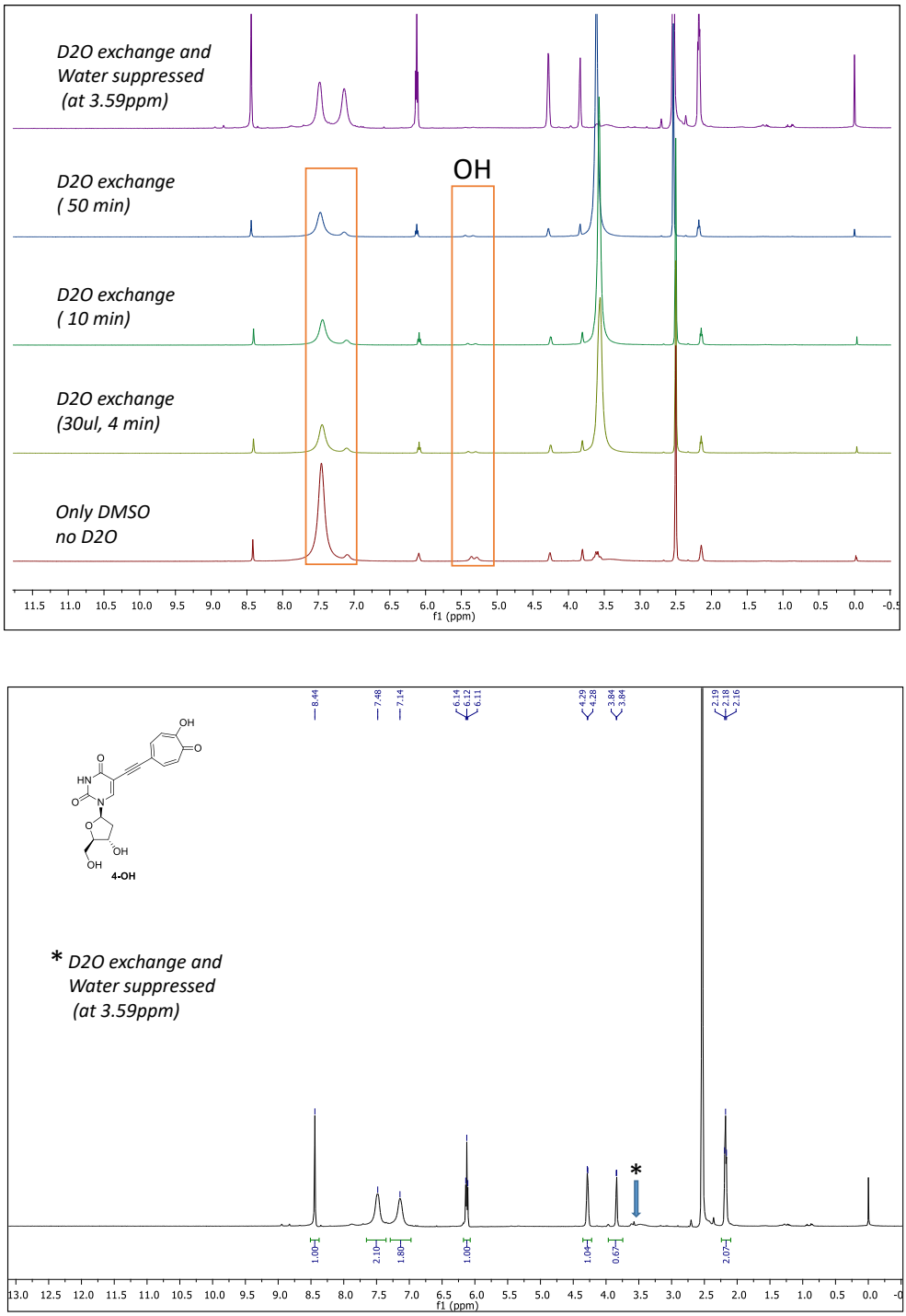


Figure 5. D_2O exchange and water suppression ^{1}H-NMR of Tr-2'-deoxyuridine (**4-OH**) in DMSO-d6 (in water suppression, the 5'methylne protons also suppressed due to close chemical shift value)

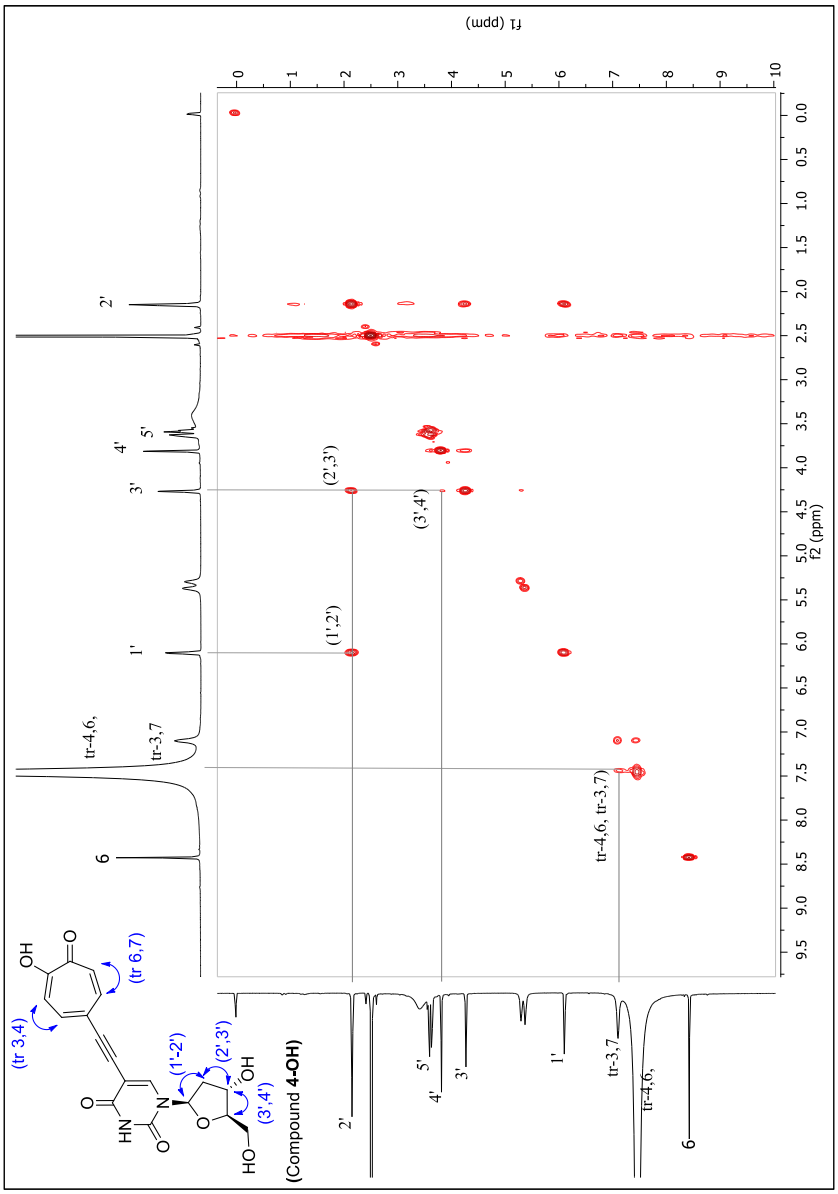


Figure 6. ^{1}H-COSY of Tr-2'-deoxyuridine (**4-OH**) in DMSO-d6 (30 mM)

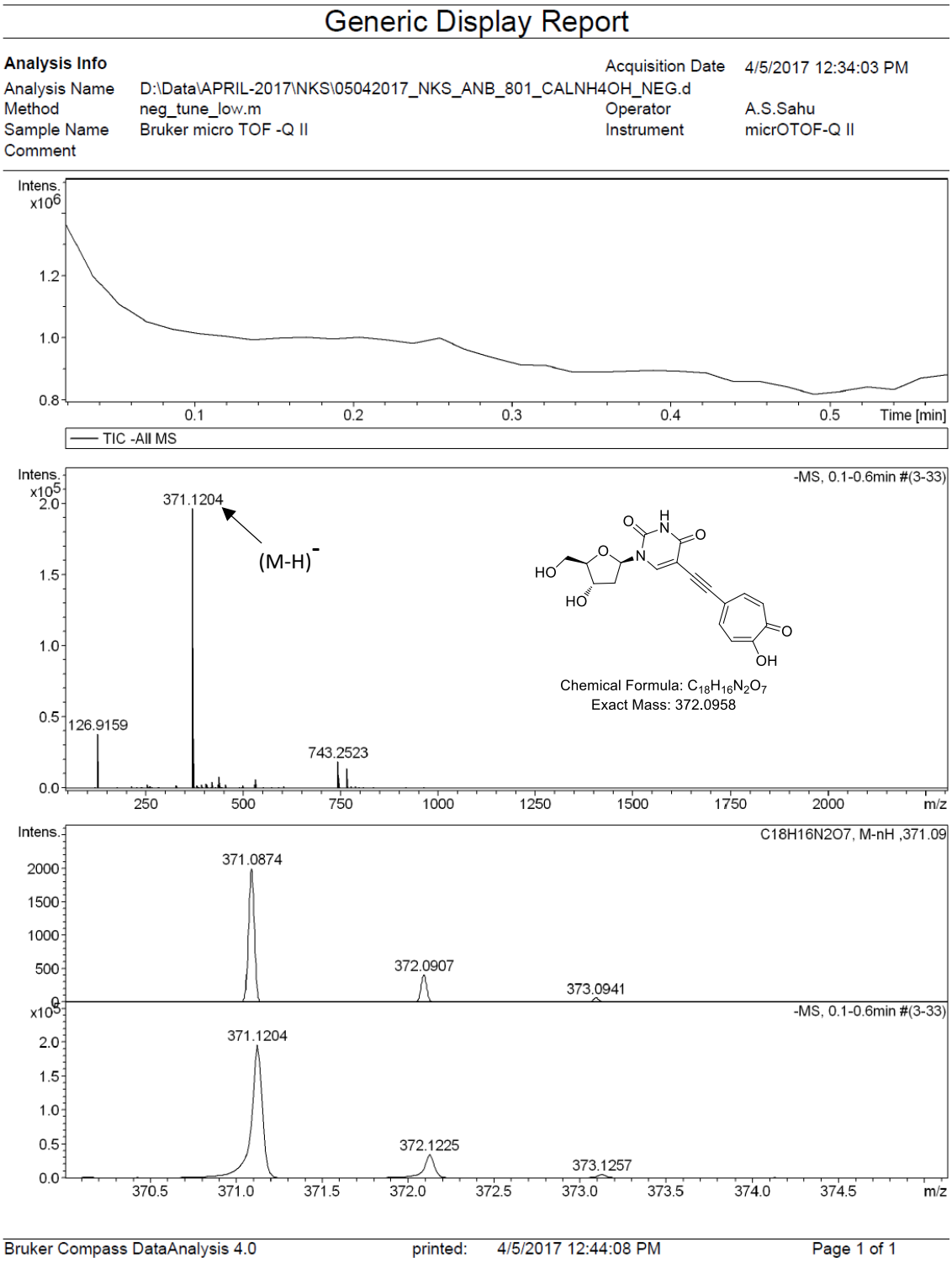


Figure S7. HRMS of *tr*- 2'-deoxyuridine (**4-OH**)

4.^{1}H/^{13}C/ HRMS spectra of 5’-DMTr-(O-benzoyl-tropolonylated)-2’-deoxyuridine **(5)**

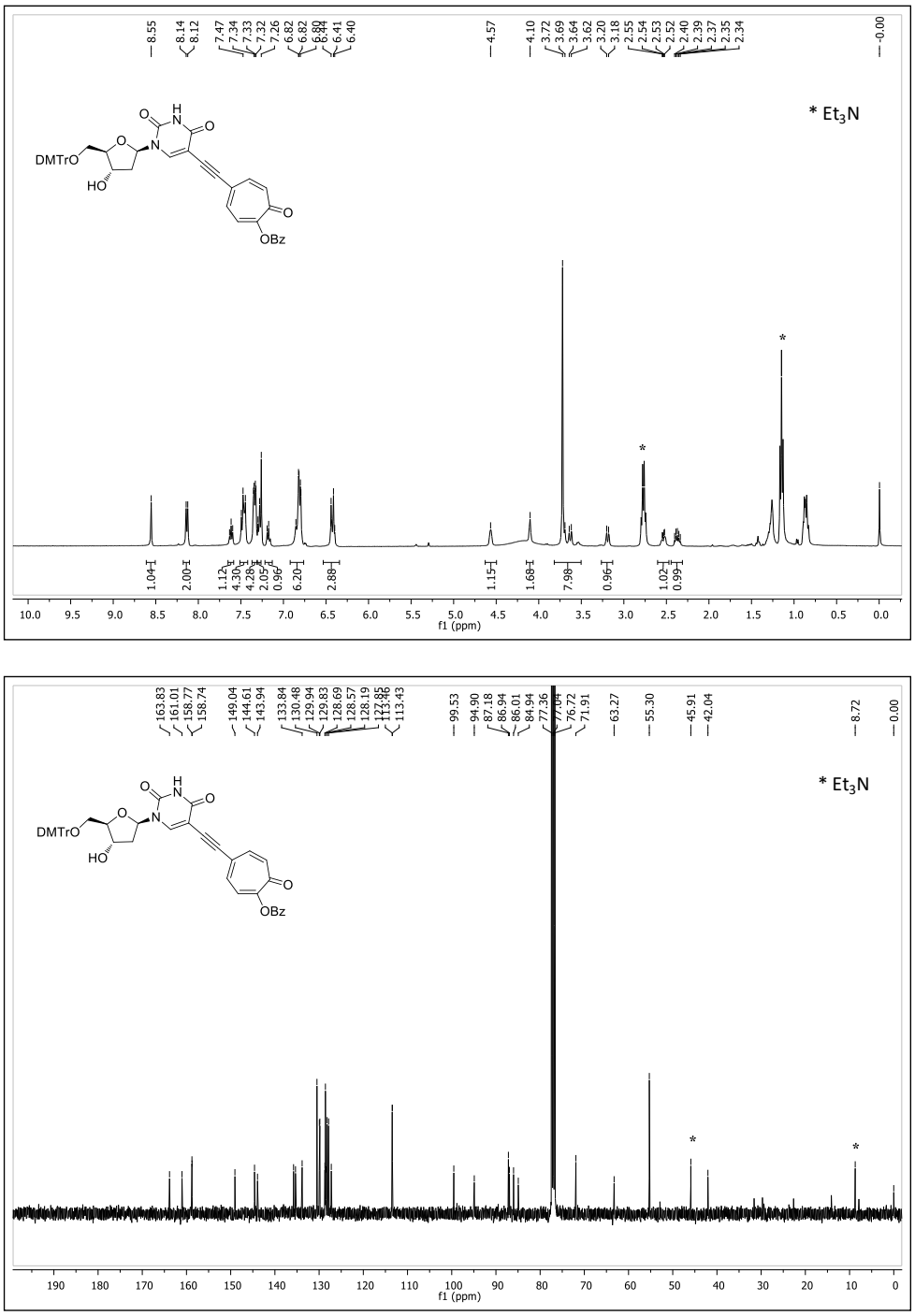


Figure S8. ^{1}H-/^{13}C-NMR of 5’-DMTr-(O-benzoyl-tropolonylated)- 2’-deoxyuridine (**5**)

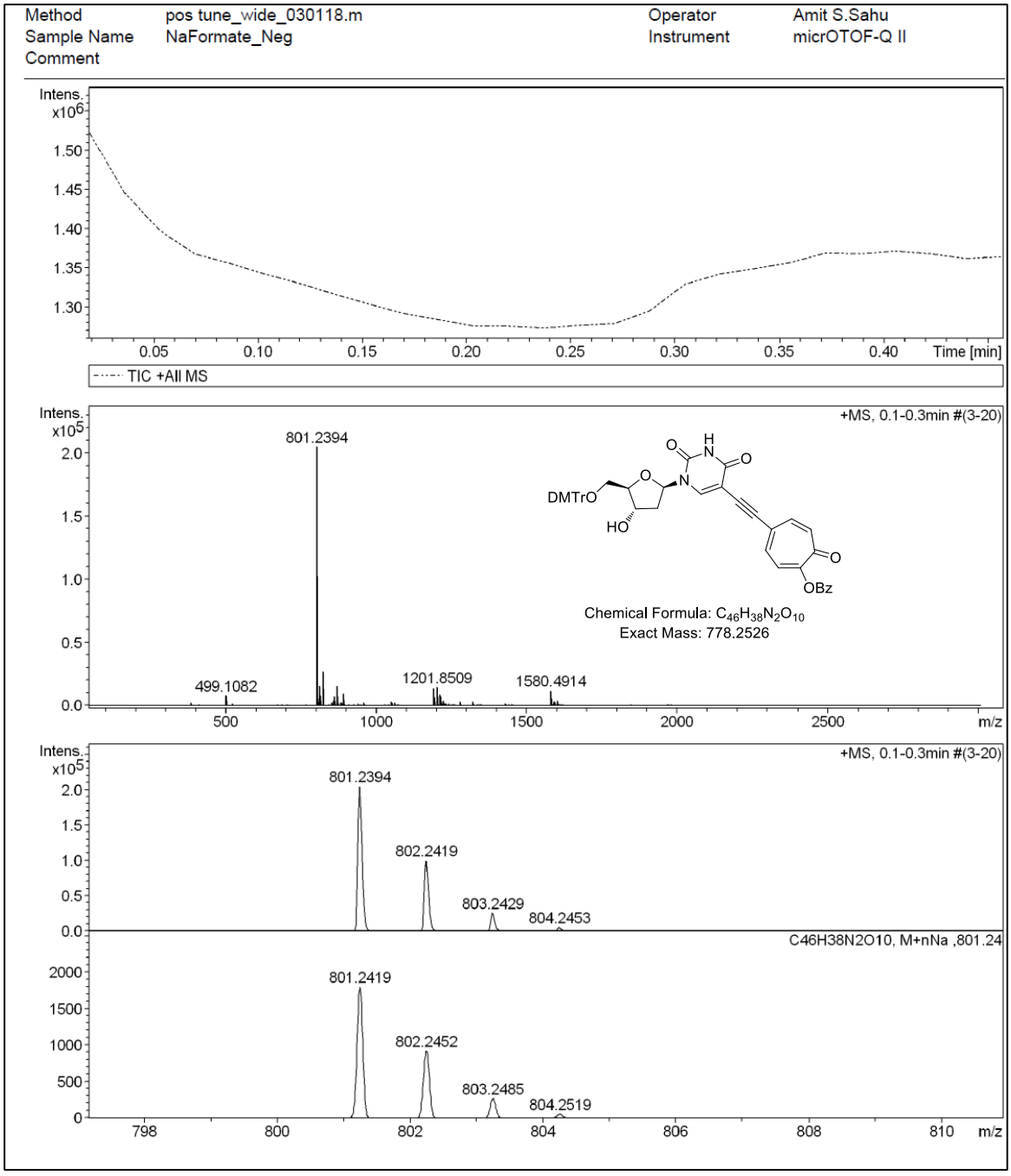


Figure S9. ESI-HRMS of 5'-DMTr-(O-benzoyl-tropolonylated)- 2'-deoxyuridine (**5**)

5.^{1}H/^{13}C/^{31}P-NMR & HRMS spectra of 2-benzoyloxy Tr-2'-deoxyuridine phosphoramidites **(6)**

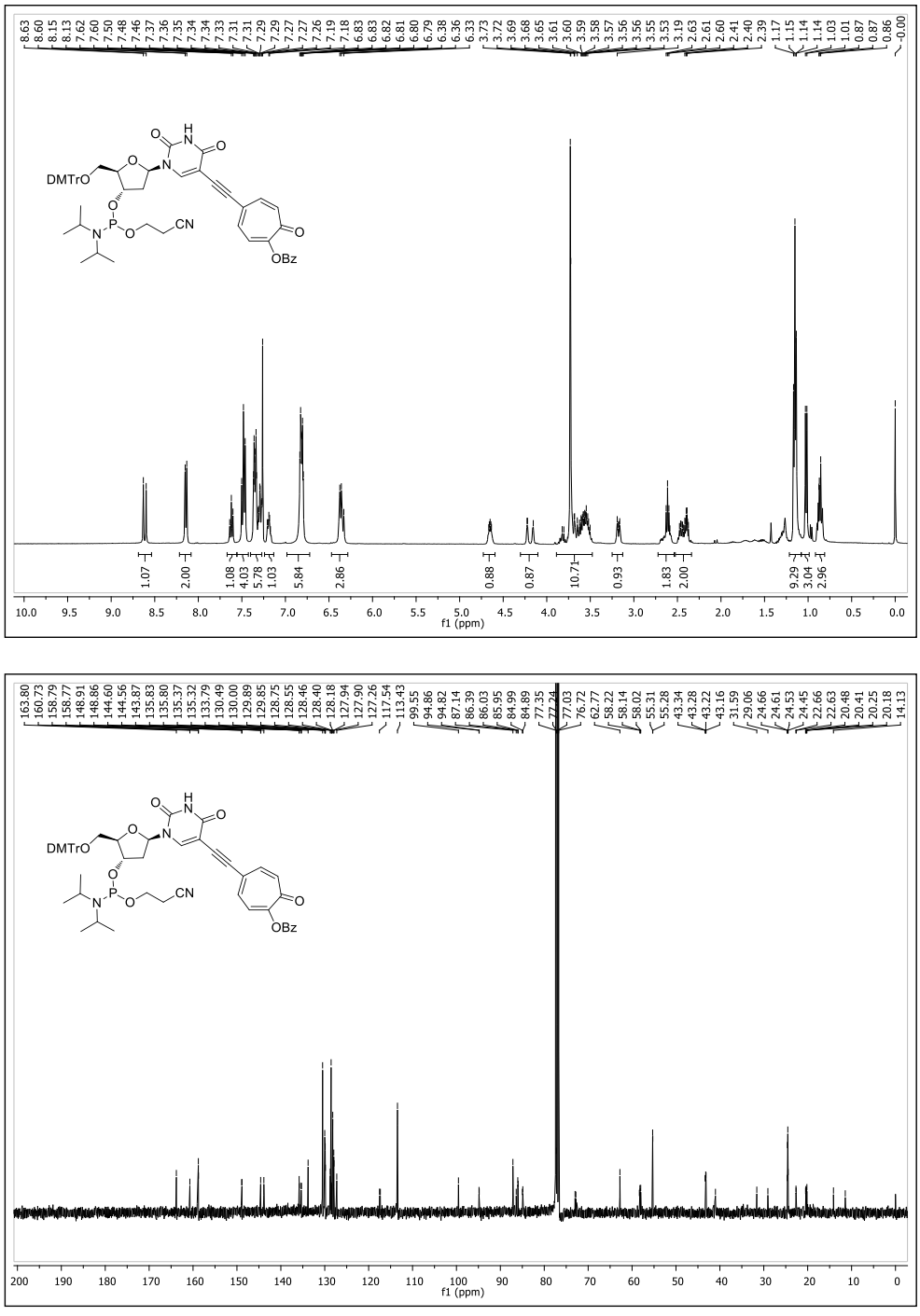


Figure S10. ^{1}H-/^{13}C-NMR of tropolonylated 2'-deoxyuridine phosphoramidites (**6**)

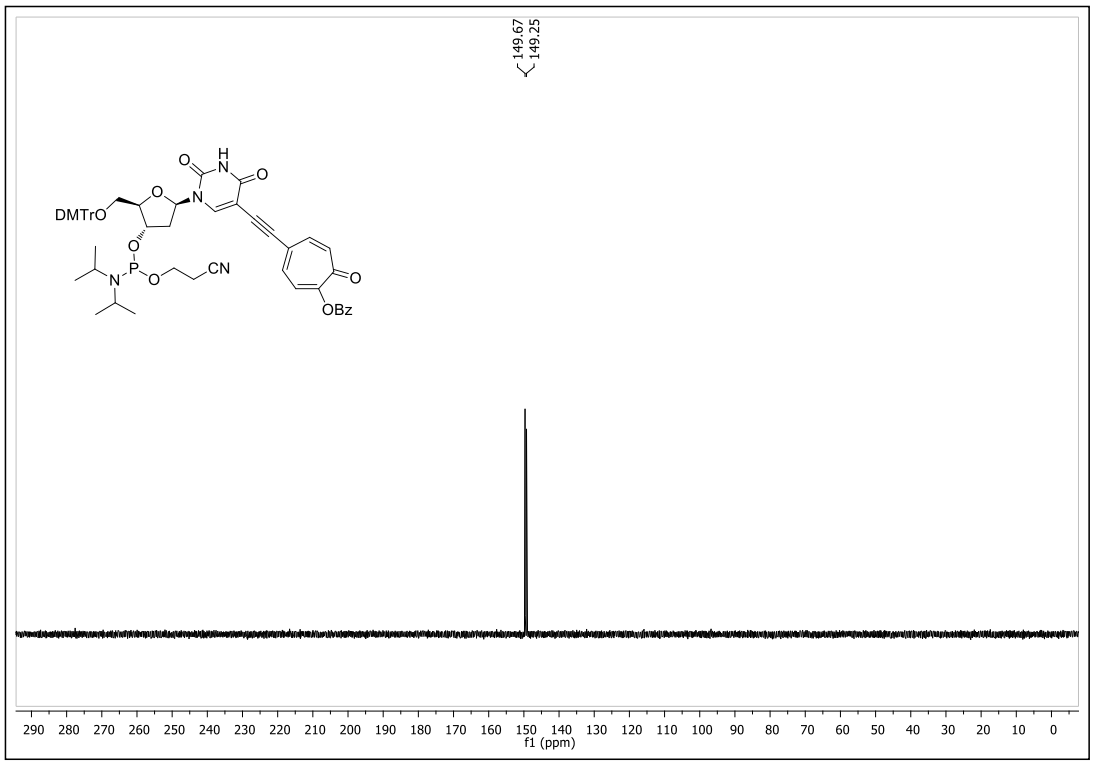


Figure S11. ^{31}P-NMR of tropolonylated 2'-deoxyuridine phosphoramidites (**6**)

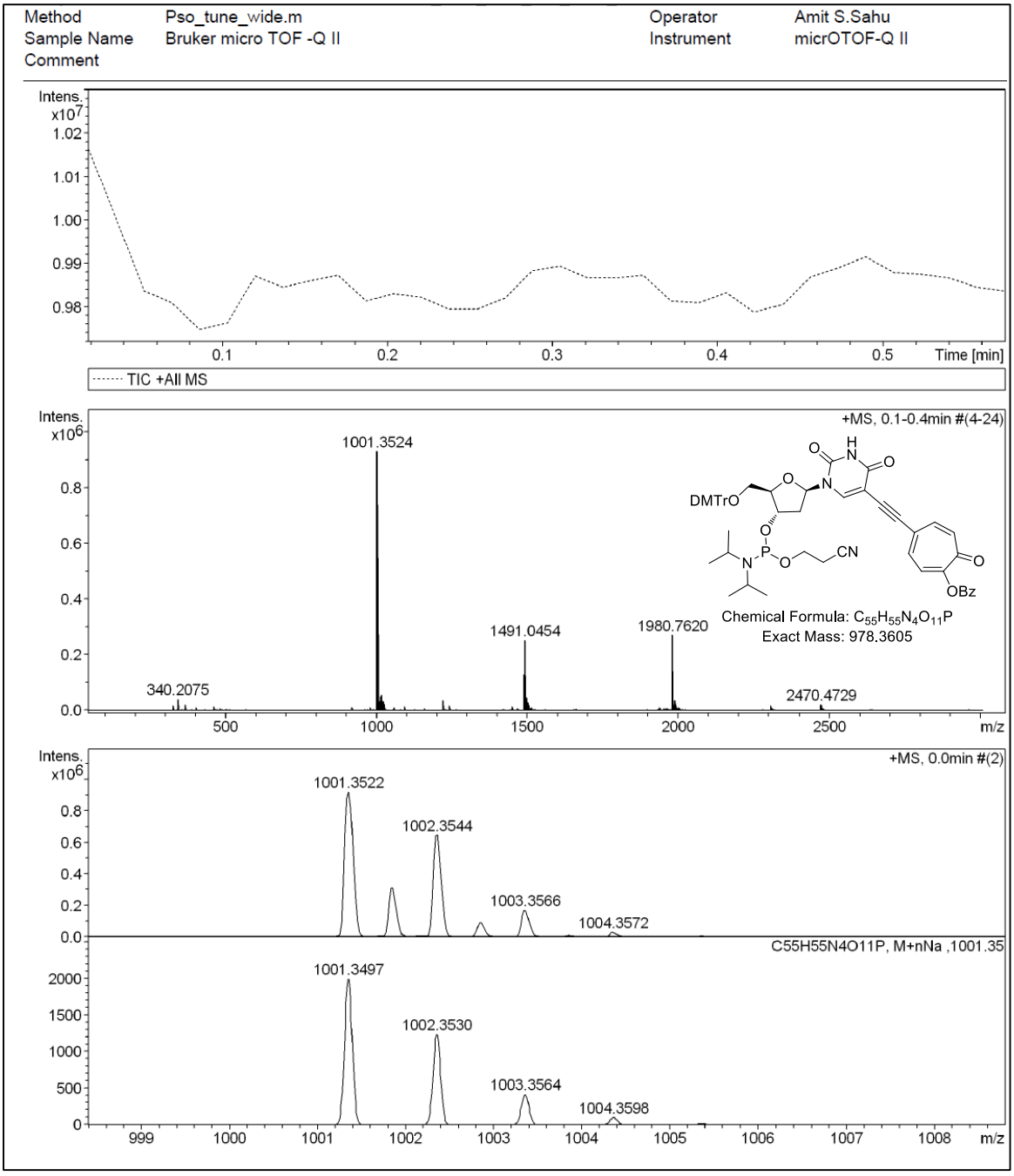


Figure S12. ESI-HRMS of tropolonylated 2'-deoxyuridine phosphoramidites (**6**)

HPLC Conditions: BUFFER A- 5% ACN in 0.1M TEA, Buffer B- 20% ACN in 0.1M TEA; Column: Sunfire C18, 5µm, 4.6x150mm,

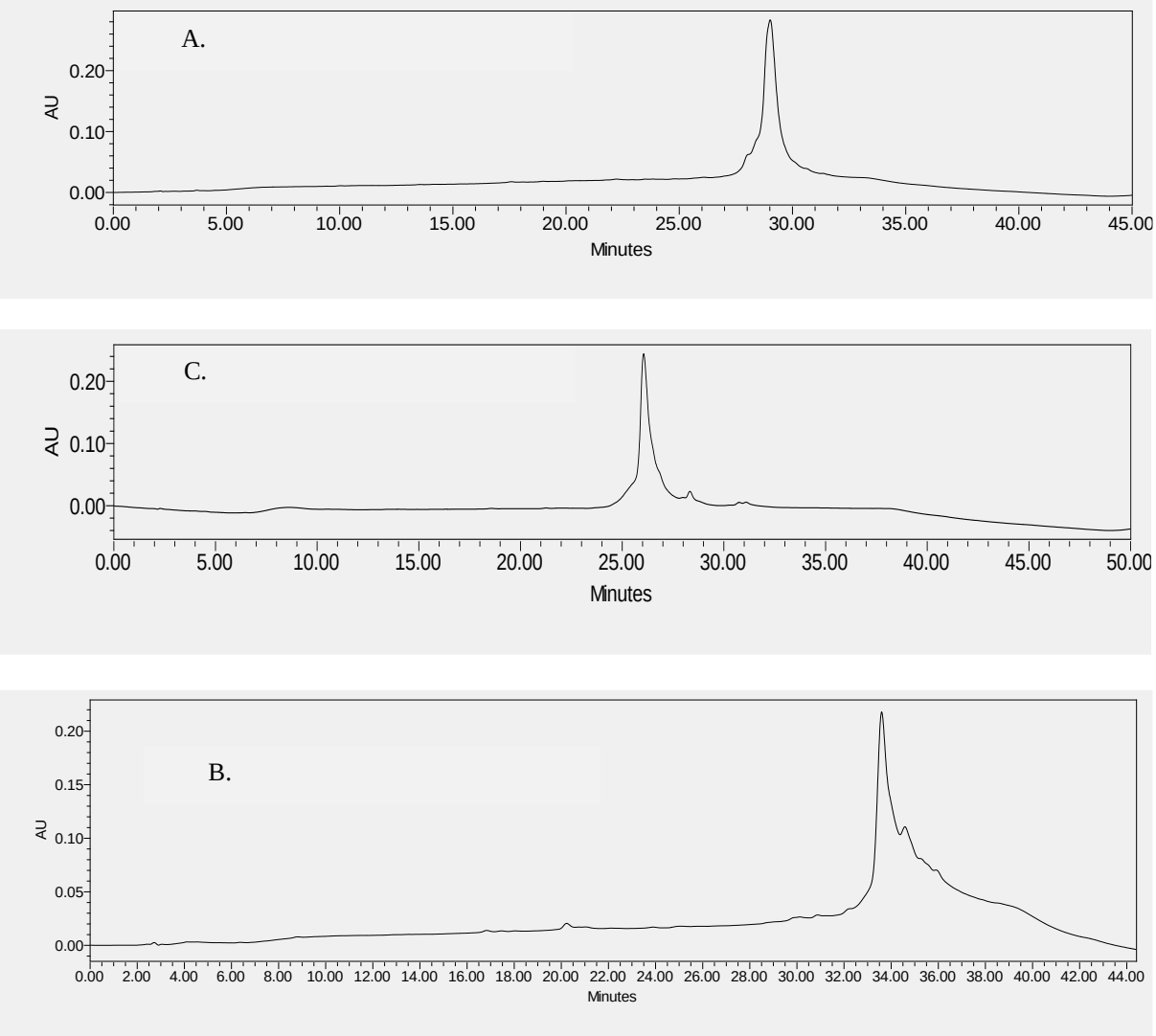


Figure S13. HPLC of modified oligonucleotides of ON1, ON2, ON3, similarly ON4 and ON6 were also purified.

6.ESI-Mass of Oligonucleotides **(ON1-ON5)**

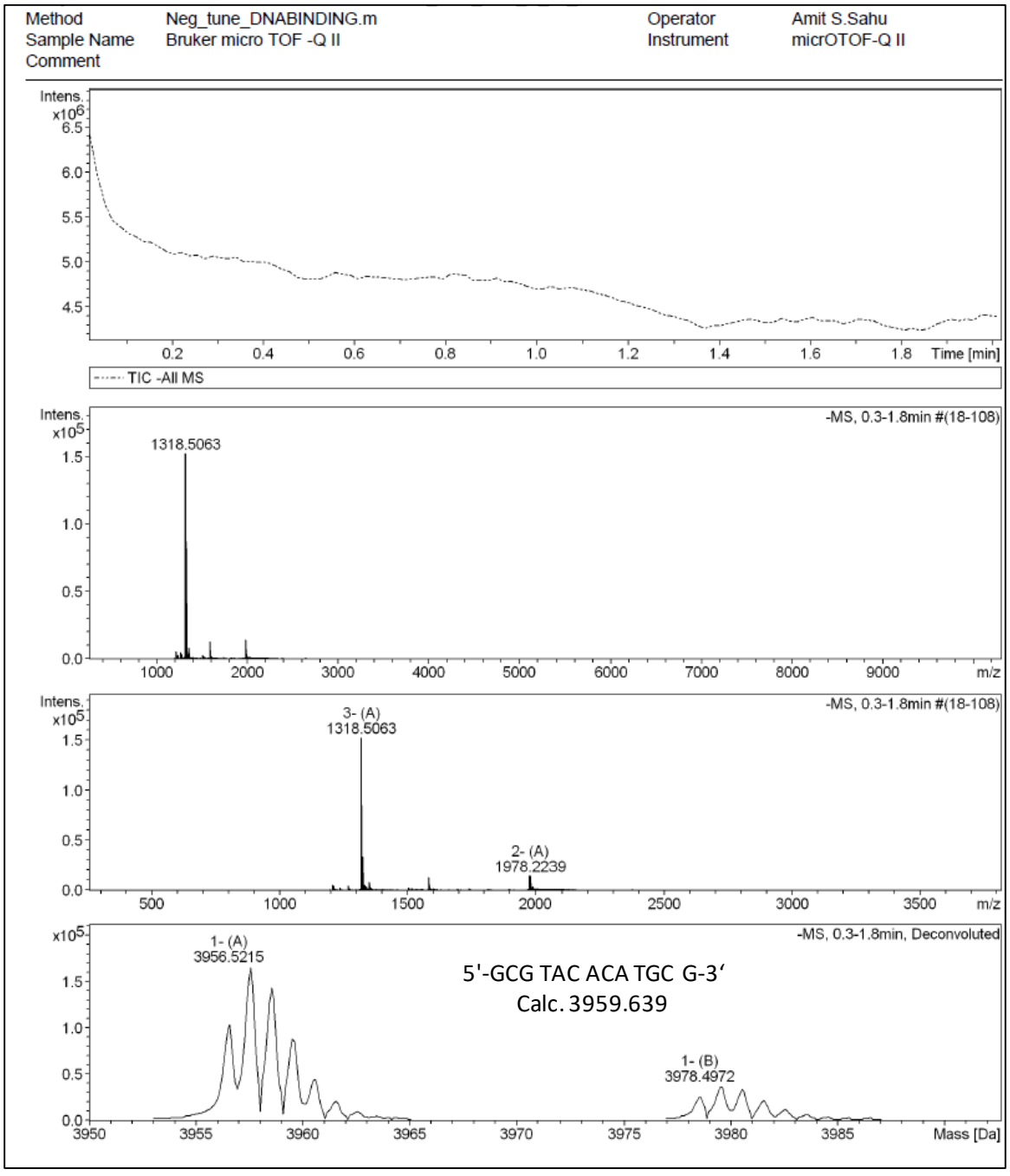


Figure S14. ESI-Mass spectra of modified Oligonucleotides (**ON1**)

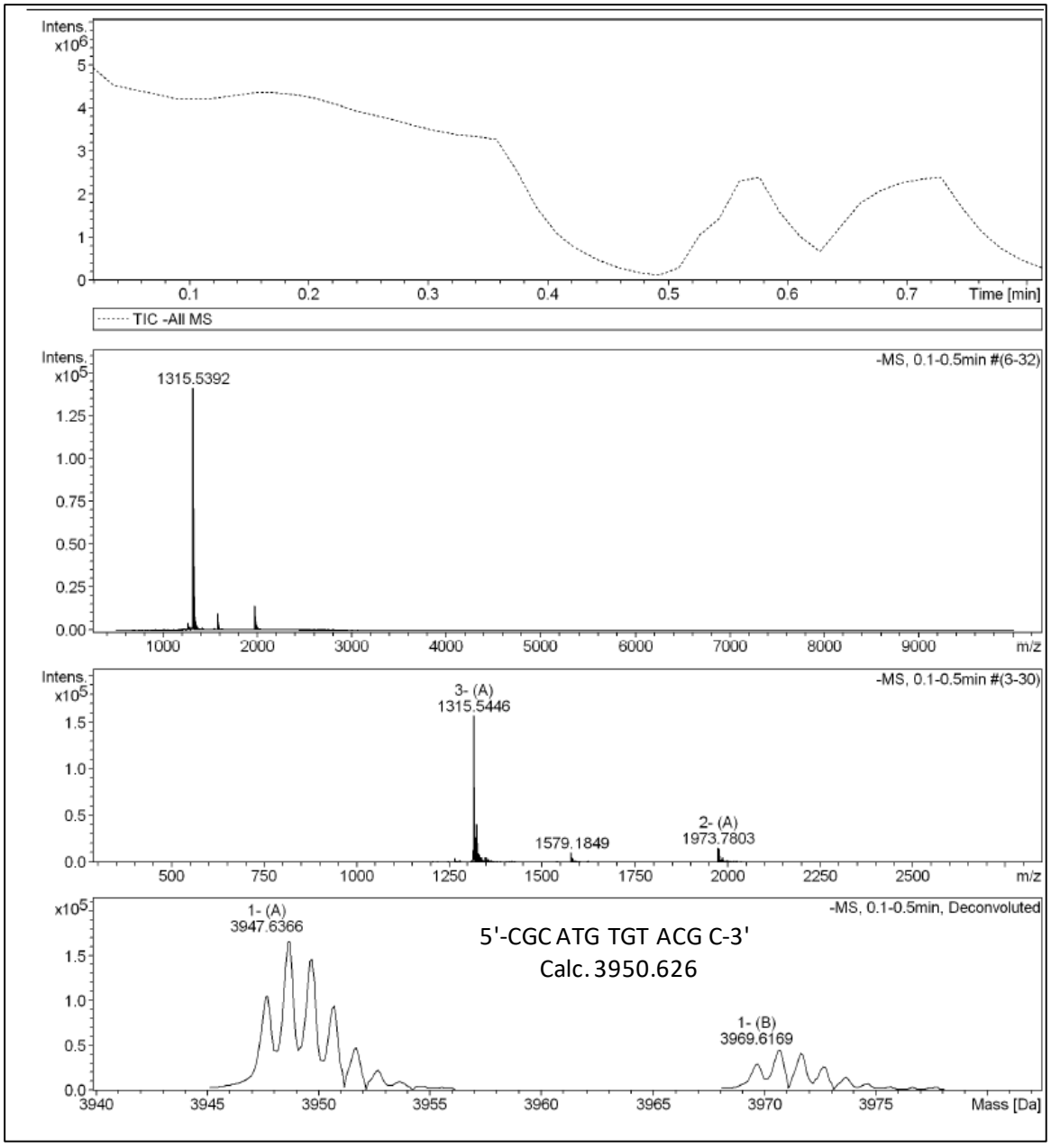


Figure S15. ESI-Mass spectra of modified Oligonucleotides (**ON2**)

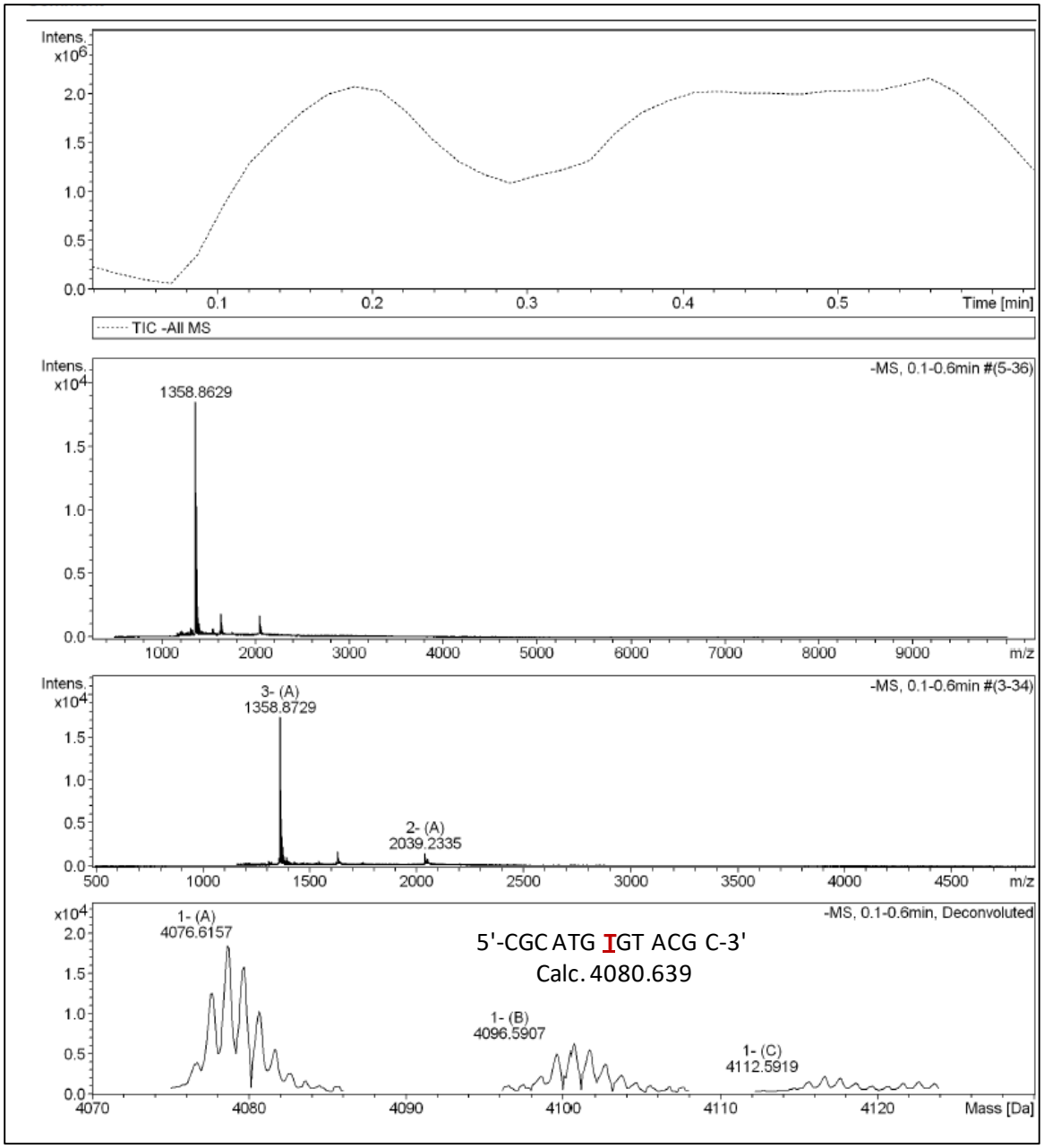


Figure S16. ESI-Mass spectra of modified Oligonucleotides (**ON3**)

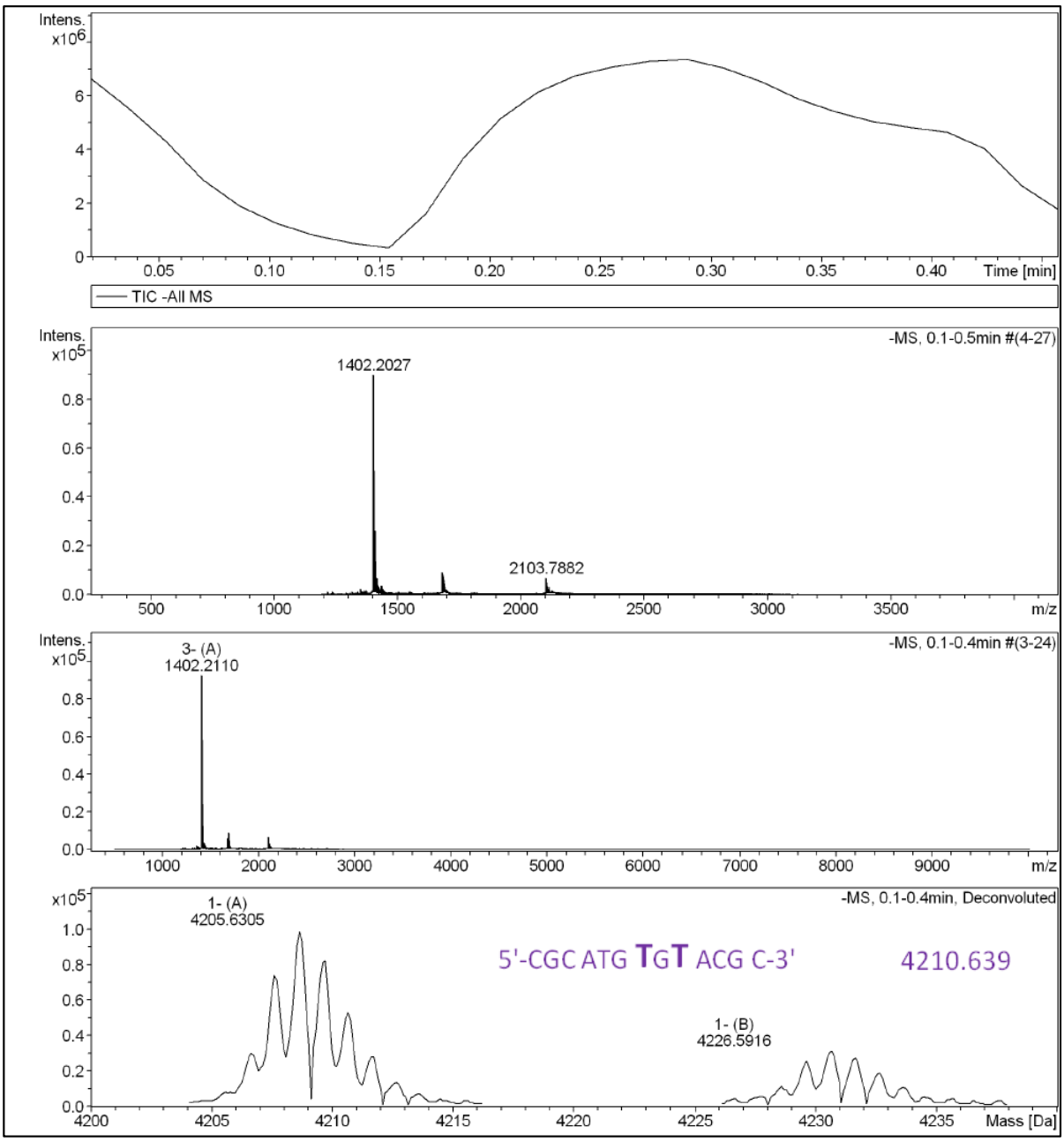


Figure S17. ESI-Mass spectra of modified Oligonucleotides (**ON4**)

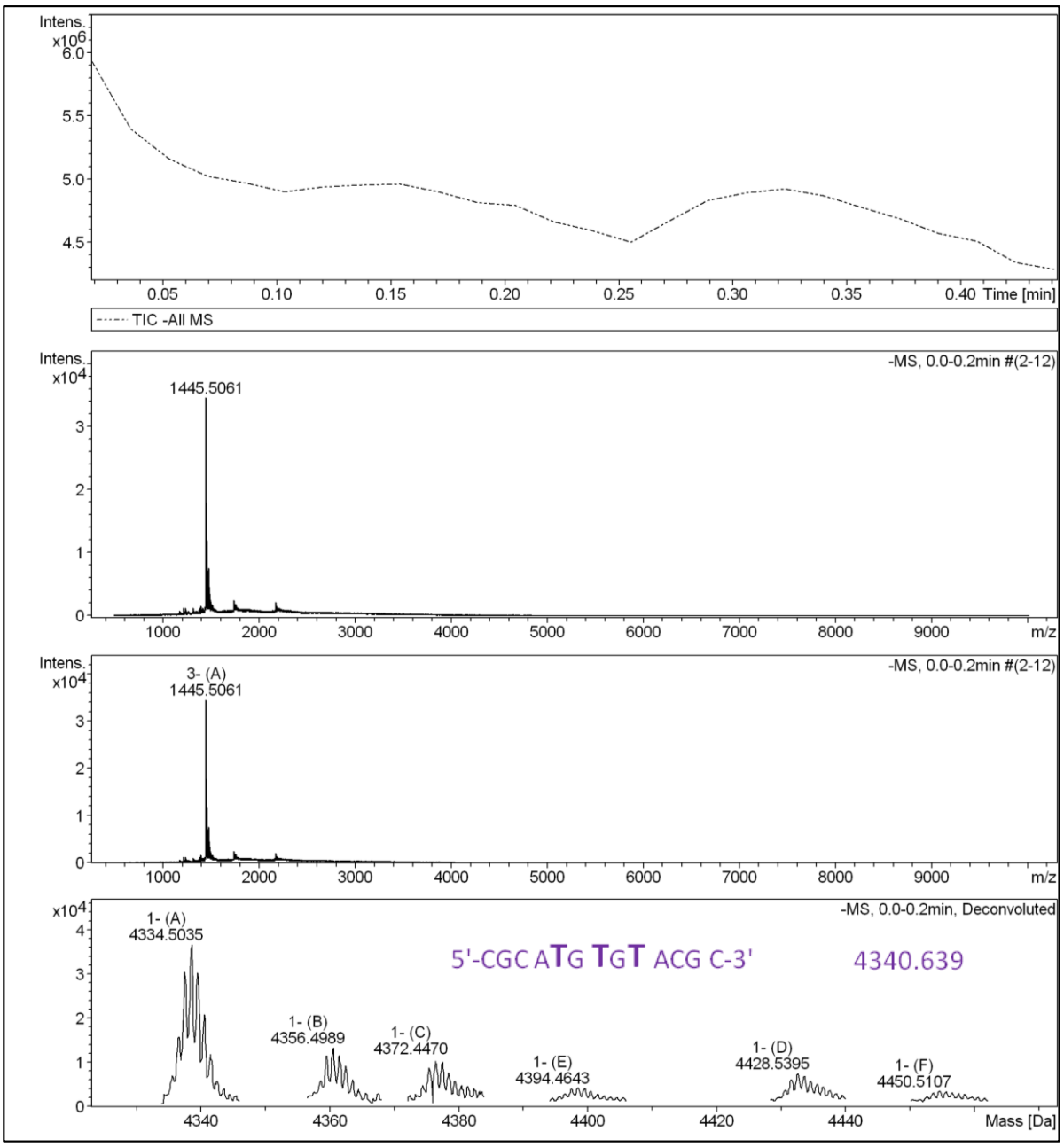


Figure S18. ESI-Mass spectra of modified Oligonucleotides (**ON5**)

7.Extinction coefficient determination of compound **4-OH**

Molar extinction coefficient (ε_{260}): Stock solutions were prepared from fully deprotected tr-thymidine (**4-OH**), weighing accurate to 2 mg in triplicates and was dissolved in 1ml of water and ammonium hydroxide (100/1, v/v, pH 10.6) for solubility. 5.0 µl of stock solutions were diluted to 3.0 ml and recorded the absorbance as triplicates. The Molar extinction coefficient of tr-thymidine (**4-OH**) found to be ε_{260}=29123 M^{-1} cm^{-1} at pH 10.61; ε_{260}=25006 M^{-1} cm^{-1} at pH 7.0; ε_{260}= 28891 M^{-1} cm^{-1} at pH 8.0. During this study we used ε_{260}=29123 M^{-1} cm^{-1} at pH 10.61 for further calculations.

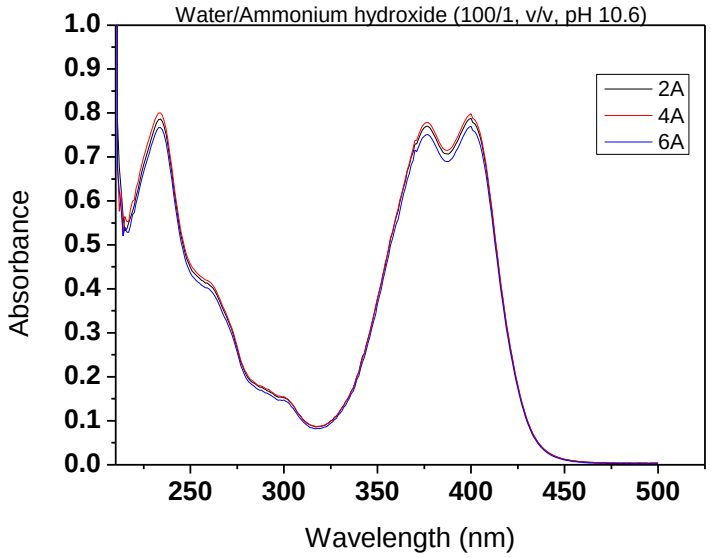


Figure S19. UV spectra of tropolonylated 2'-deoxyuridine (*tr*-T) **4-OH** in H_2O:NH_4OH (100/1, v/v, pH 10.6)

8.pK_a of control **(Tropolone)**

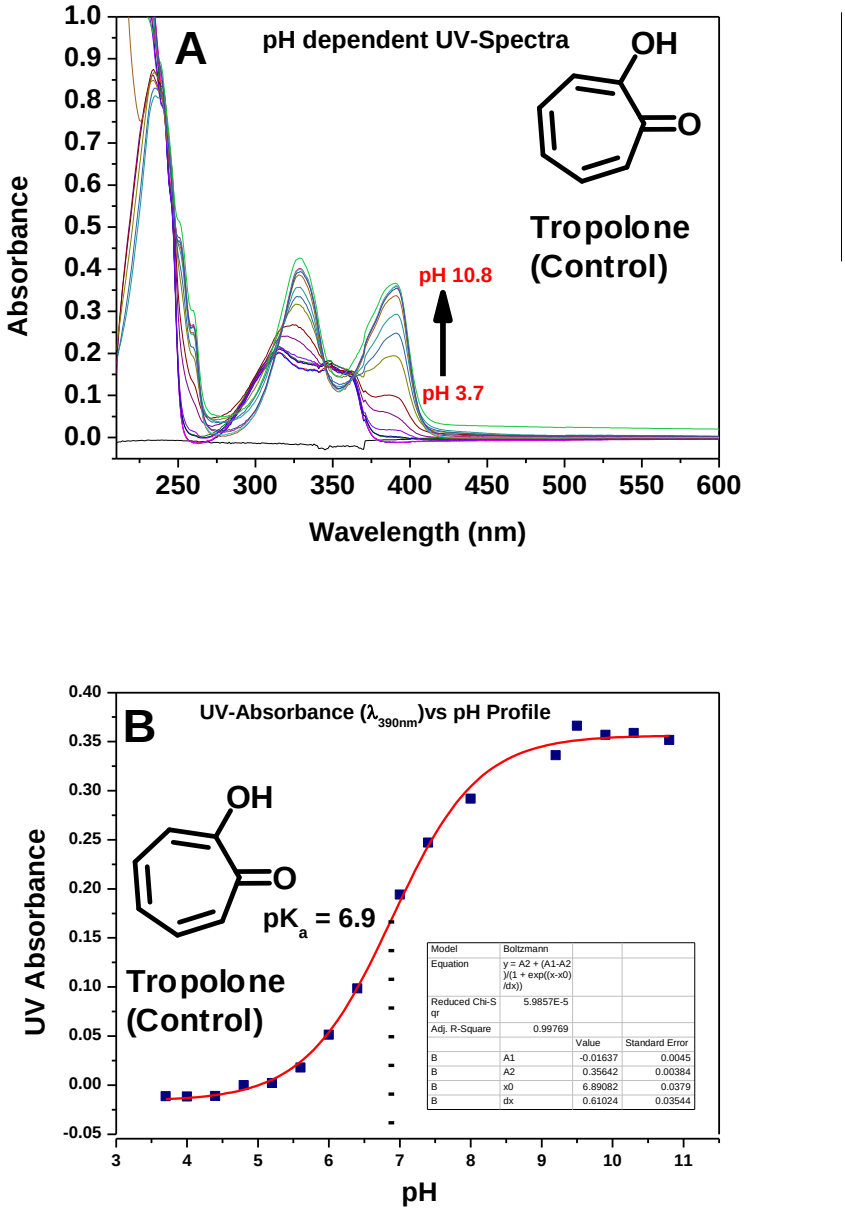


Figure 20. (A) pH dependent UV spectra of tropolone for control studies; (B) Determination of pK_a of Tropolone

9.UV and Fluorescence properties of tropolonyl nucleoside (**4/4-OH**)

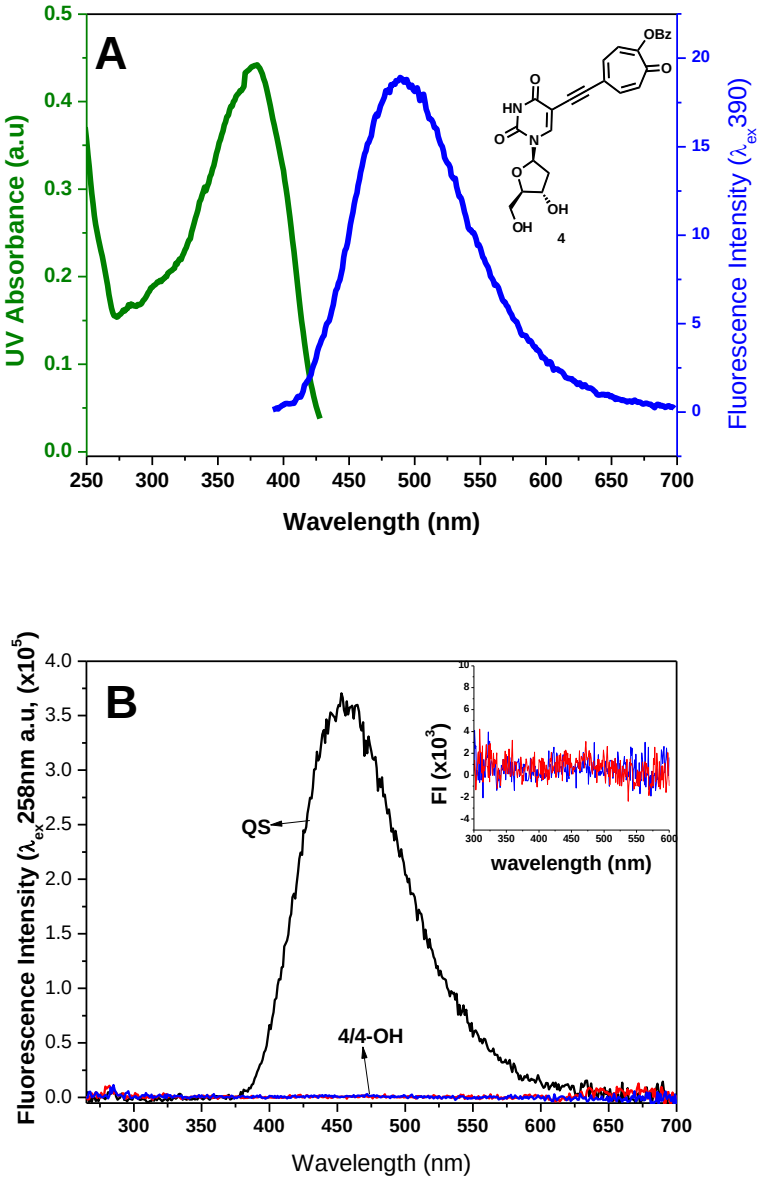


Figure 21. (A) UV (left) and Emission (right) spectra (bandwidth of 5) of nucleoside 4 in methanol, (22µm); (B) Fluorescence (emission) spectra of nucleoside **4** (in MeOH), **4-OH** (pH 7.0 PBS buffer) and Quine sulphate for relative quantum yield. (bandwidth of 2). Inset showing expanded portion

10.CD spectra and UV-Melting profiles of control DNA duplex

*CD Spectra***:** The samples were annealed by heating to 95 °C, followed by cooling to room temperature overnight and stored at 4 ºC. CD spectra were recorded in PBS (pH 7.02) at 10 ºC from 450-200 nm with Oligonucleotide concentrations of 2 µM. CD parameters : Data pitch 2 nm, DIT 2 sec, bandwidth 2 nm, scanning speed 100 nm/min, and accumulation 2. For pH dependent studies freshly prepared 100mM sodium phosphate buffers of 6.0, 7.02, and 8.1 was used. (doi:10.1101*/pdb.rec8303*, *Cold Spring harb protoc, 2006)*

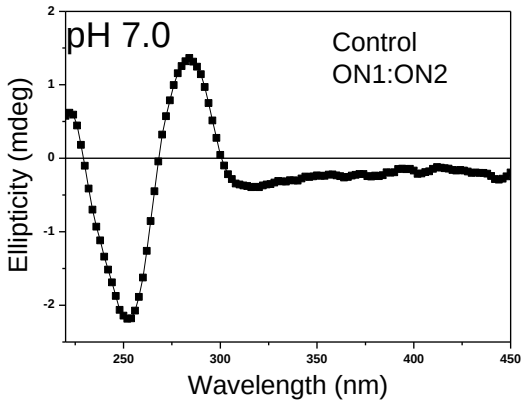


Figure S22. CD spectra of 2 µM DNA duplex (**ON1:ON2**) in sodium phosphate buffer (100 mM) pH 7.0

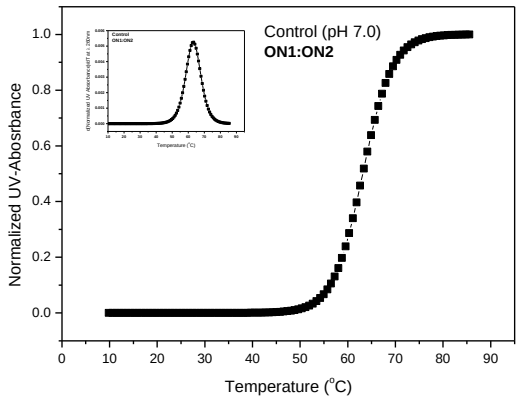


Figure S23. UV- melting of 2 µM DNA duplex (**ON1:ON2**) in sodium phosphate buffer (100 mM) pH 7.0

11.pH dependent CD-Spectra of *tr*-DNA duplexes

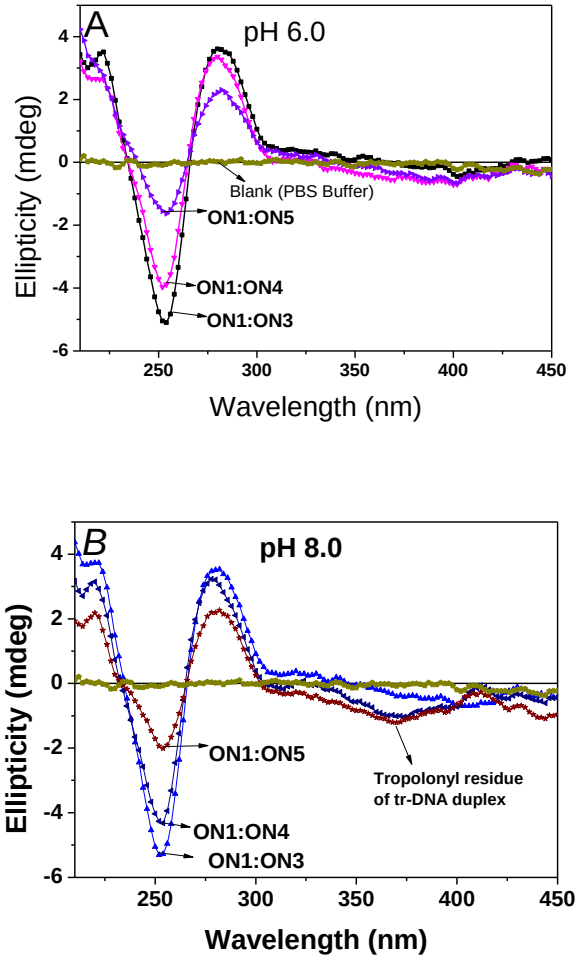


Figure S24. CD spectra of 2 µM DNA duplexes in sodium phosphate buffer (100 mM), a) at pH 6.0; b) at pH 8.0

12.pH dependent UV-Thermal Melting Profiles of *tr*-DNA duplex

UV-melting (T_m determination) and UV-Vis spectra: Oligonucleotide concentrations were calculated using extinction coefficient (ε_{260}) of *tr*- thymidine (*tr*-T) (ε_{260}=29123 M^{-1} cm^{-1}) using nearest-neighbor model (*Biophys. Chem.*, **2008**; *133*, 66-70). Equimolar volumes of complementary strands were thoroughly mixed in PBS and denatured by heating to 95 °C,

followed by cooling to room temperature overnight and stored at 4 °C. Prior to experiments samples were degassed for at least 5min with argon at 4 °C. Quartz cuvettes with a path length of 1.0 cm were used. Thermal denaturation curves of duplexes (2 μM final concentration of each strand in 1.5 ml) were recorded on a Peltier-controlled UV/VIS spectrophotometer (JASCO). A temperature ramp of 0.5 °C/min was used in all experiments. Experiments were performed in duplicates.

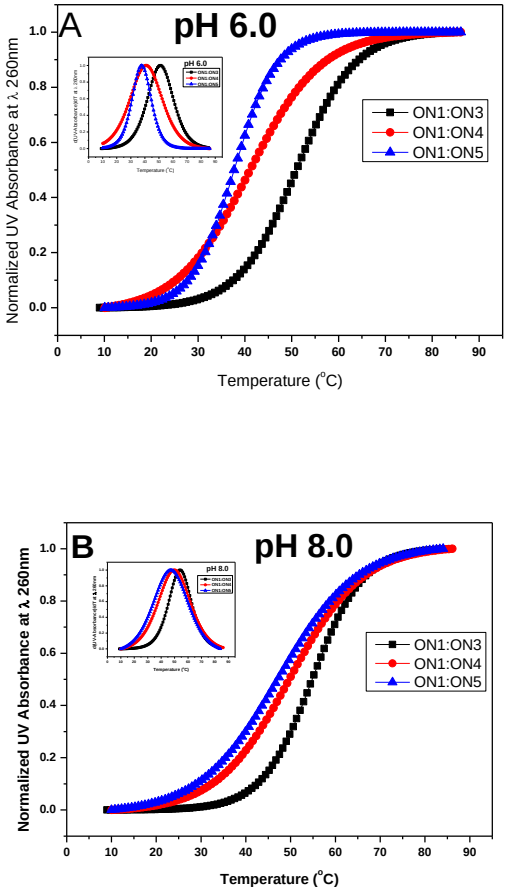


Figure S25. UV- melting of 2 μM DNA duplexes in sodium phosphate buffer (100 mM). a) At pH 6.0; b) at pH 8.0

13.UV-Visible Spectra of *tr*-DNAs and their duplex

(a) Number of *tr*-T nucleoside units dependent UV spectra of tr-DNA and their duplex

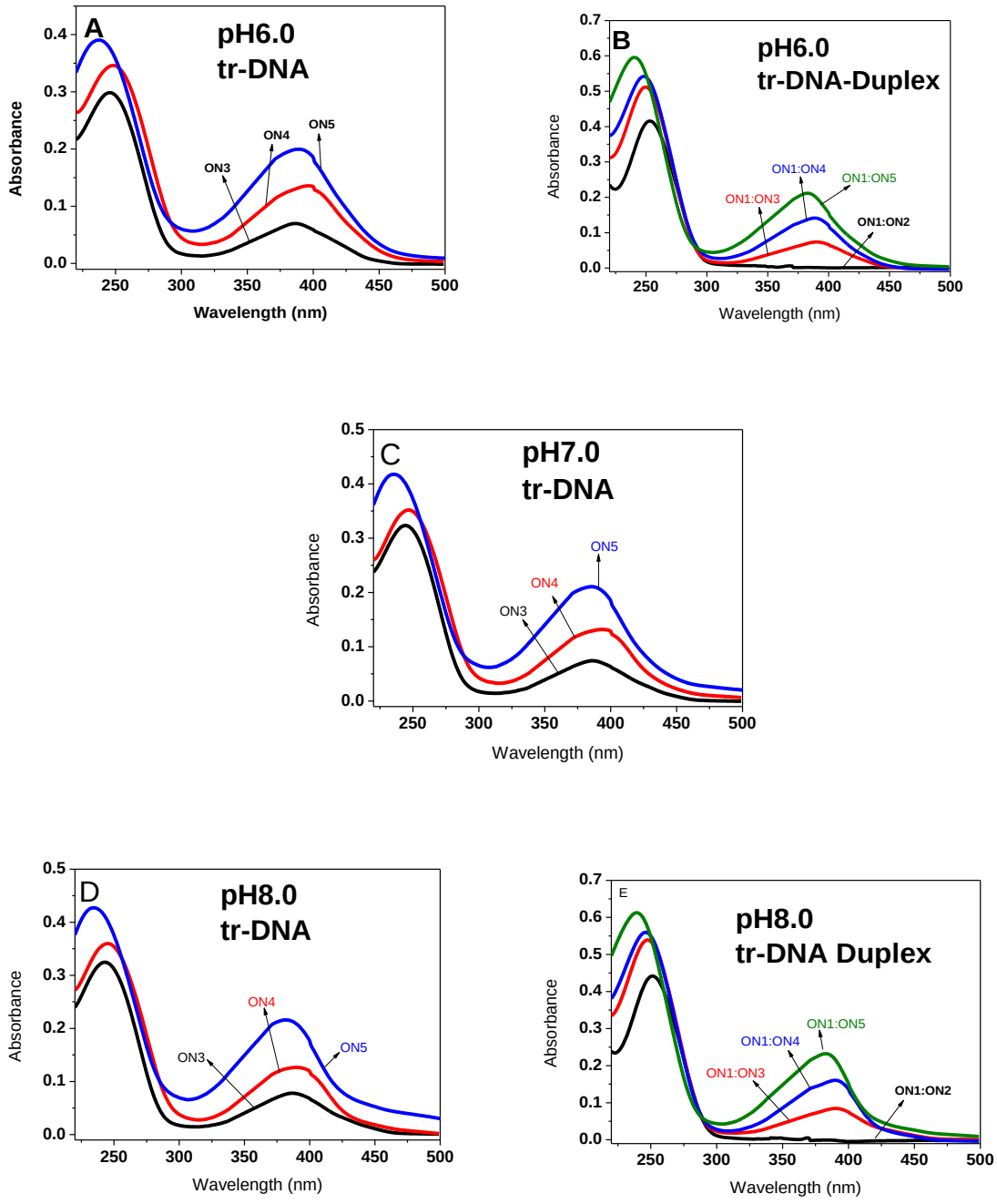


Figure S26. UV- Visible spectra of 2 μM DNA duplexes in sodium phosphate buffer(100 mM) a) & b) at pH 6.0 single *tr*-T modification, single strand and their duplex respectively; c) at pH 7.0 two *tr*-T

modification, single strand; d) & e) at pH 8.0 three *tr*-T modification, single strand and their duplex respectively.

(b) pH dependent UV spectra in tr-DNA and their duplex

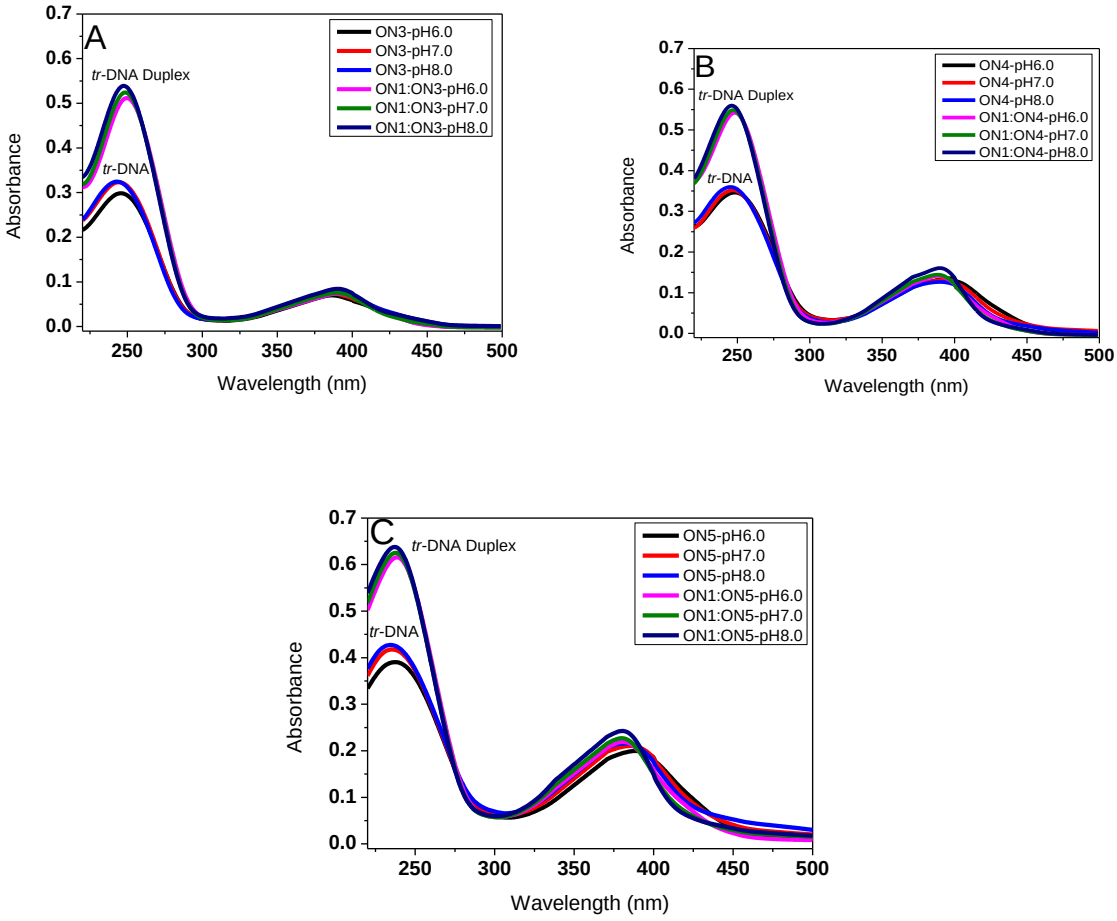


Figure S27. UV- Visible spectra of 2 μM *tr*-DNAs and their duplexes in sodium phosphate buffer (100 mM): (A) tr-DNA (**ON3**) and its duplex (**ON1:ON3**) at pH 6.0/7.0/8.0. (B) *tr*-DNA (**ON4**) and its duplex (**ON1:ON4**) at pH 6.0/7.0/8.0; (C) *tr*-DNA (**ON5**) and its duplex (**ON1:ON5**) at pH 6.0/7.0/8.0.

14.pH dependent Fluorescence Spectra of *tr*-DNAs and their duplex

Oligonucleotides of concentration 1.0 µM were annealed by heating to 95 °C, followed by cooling to room temperature overnight and stored at 4 °C. Prior to experiments samples were degassed for at least 5min with argon at 4 °C. The spectra were recorded with cuvettes having a path length of 2 mm with 700 µl of volume. For pH dependent studies freshly prepared 100 mM sodium phosphate buffers of 6.0, 7.02, and 8.1 was used. Fluorescence spectra were recorded on FS5 Fluorescence Spectrometer (Edinburgh Instruments). The fluorescence quantum yields were determined using standard quinine sulphate with 0.1N H_2SO_4.

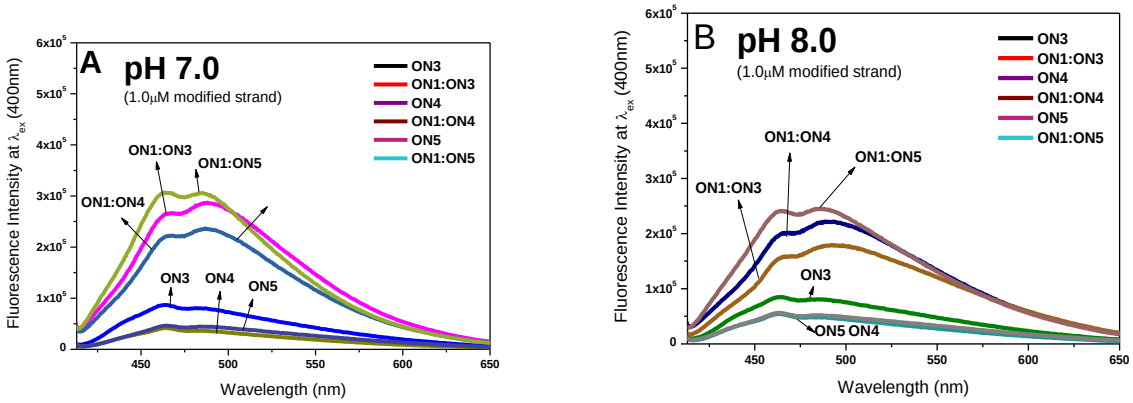


Figure S28. Fluorescence spectra of 1.0 µM *tr*-DNAs and their duplexes in sodium phosphate buffer (100 mM) a) at pH 6.0; b) at pH 8.0

15.Quantum yeild of tr-DNA/tr-DNA duplex

Fluorescence spectra were recorded on FS5 Fluorescence Spectrometer (Edinburgh Instruments). Prior to experiments samples were degassed for at least 5min with argon. The spectra were recorded with cuvettes having a path length of 2 mm with 700 µl of volume. The fluorescence quantum yields were determined using standard quinine sulphate with 0.1N H_2SO_4.

$$QY_{cpd} = QY_{ref} \times \frac{A_{cpd}}{A_{ref}} \times \frac{OD}{OD} \times \frac{R_{ref}^2}{R_{cpd}^2}$$

Where QY_{cpd}=Quantum yield of **ON3**, QY_{ref}=Quantum yield of reference (quinine sulphate with 0.1N H_2SO_4) A_{cpd} =area under the curve of **ON3**, A_{ref}= area under the curve of quinine sulphate, OD_{ref}, OD_{cpd} = absorption at excitation wavelength of reference and **ON3** respectively, R^2_{ref} **and** R^2_{cpd} =refractive index of Phosphate buffer (water, 1.33) and 0.1M H_2SO_4 (in water, 1.33) respectively.

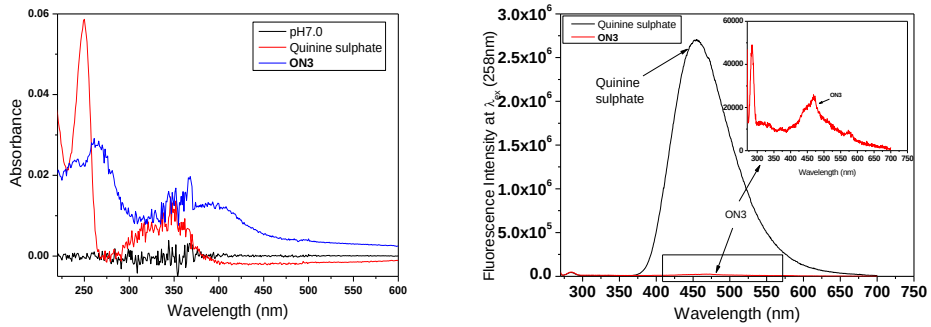


Figure S29. UV (left) and Emission (right) spectra of **ON3** and quinine sulphate for recording Quantum yield in sodium phosphate buffer pH 7.0 (Inset: fluorescence of **ON3**).

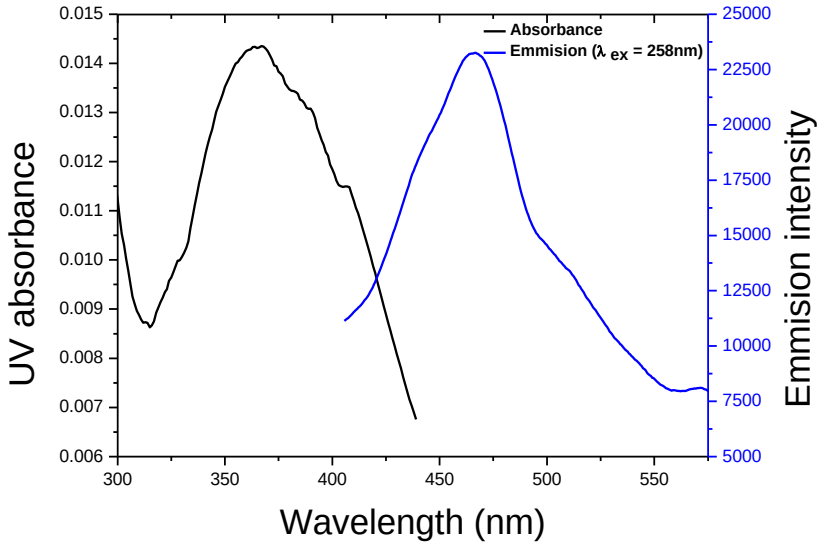


Figure S30. Smoothened UV and Emission spectra of **ON3** for Quantum yield in sodium phosphate buffer at pH 7.0

Quantum yields of *tr*-DNAs and their duplexes in sodium phosphate buffer (100 mM) at pH 7.0 using standard quinine sulphate with 0.1N H_2SO_4.

Entry	tr-DNA	pH6.0 (%)	pH7.0 (%)	pH8.0 (%)
1	ON3	1.2	0.9	0.8
2	ON1:ON3	3.1	2.3	1.6
3	ON4	1.1	0.4	0.5
4	ON1:ON4	3.8	1.9	1.2
5	ON5	0.8	0.4	0.5
6	ON1:ON5	4.1	2.8	2.3

16.DFT-B3LYP-6 calculations of Tropolonylated 2'-deoxyuridine (**4-OH**)

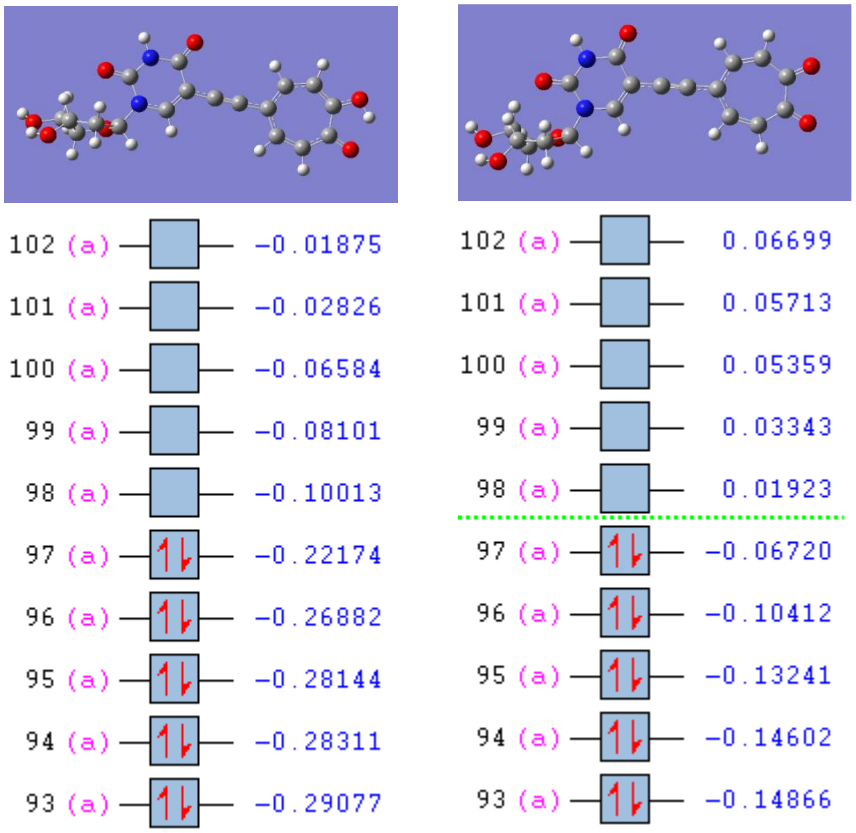


Figure 37. Optimized structure and HOMO-LUMO energies (atomic units in hartree) of Tropolonylated thymidine (**4-OH**)

A. Neutral form

TITLE tr-TTP_B3LYP-6

REMARK 4

REMARK 4 COMPLIES WITH FORMAT V. 2.2, 16-DEC-1996

REMARK 6 created by ArgusLab version 4.0.1

REMARK 6 http://www.arguslab.com

REMARK 6

ATOM	1	C	MOL	A	0	-5.658	-0.657	0.744	C
ATOM	2	C	MOL	A	0	-4.355	-0.819	1.533	C
ATOM	3	O	MOL	A	0	-3.886	-1.064	-0.802	O
ATOM	4	C	MOL	A	0	-5.301	-1.245	-0.635	C
ATOM	5	O	MOL	A	0	-6.709	-1.366	1.387	O
ATOM	6	C	MOL	A	0	-6.020	-0.572	-1.794	C
ATOM	7	O	MOL	A	0	-7.419	-0.749	-1.529	O
ATOM	8	C	MOL	A	0	-3.253	-0.941	0.459	C
ATOM	9	C	MOL	A	0	0.063	0.795	0.194	C
ATOM	10	C	MOL	A	0	-0.930	-0.135	0.331	C
ATOM	11	N	MOL	A	0	-2.271	0.164	0.396	N
ATOM	12	C	MOL	A	0	-2.720	1.486	0.265	C
ATOM	13	N	MOL	A	0	-1.705	2.418	0.142	N
ATOM	14	C	MOL	A	0	-0.316	2.216	0.089	C
ATOM	15	O	MOL	A	0	0.445	3.161	-0.035	O
ATOM	16	O	MOL	A	0	-3.900	1.800	0.270	O
ATOM	17	C	MOL	A	0	1.432	0.441	0.141	C
ATOM	18	C	MOL	A	0	2.624	0.198	0.088	C
ATOM	19	C	MOL	A	0	4.024	-0.063	0.020	C
ATOM	20	C	MOL	A	0	4.869	1.034	-0.156	C
ATOM	21	C	MOL	A	0	6.270	1.084	-0.261	C
ATOM	22	C	MOL	A	0	7.197	0.061	-0.218	C
ATOM	23	C	MOL	A	0	6.991	-1.394	-0.048	C
ATOM	24	C	MOL	A	0	5.690	-1.990	0.109	C
ATOM	25	C	MOL	A	0	4.439	-1.428	0.137	C
ATOM	26	O	MOL	A	0	8.490	0.372	-0.345	O

ATOM 27 O MOL A 0 8.038 -2.083 -0.049 O

ATOM 28 H MOL A 0 -4.174 0.002 2.229 H

ATOM 29 H MOL A 0 -4.407 -1.751 2.106 H

ATOM 30 H MOL A 0 -5.914 0.402 0.644 H

ATOM 31 H MOL A 0 -2.637 -1.829 0.638 H

ATOM 32 H MOL A 0 -5.535 -2.320 -0.629 H

ATOM 33 H MOL A 0 8.955 -0.506 -0.284 H

ATOM 34 H MOL A 0 -7.505 -1.252 0.839 H

ATOM 35 H MOL A 0 -5.753 0.492 -1.821 H

ATOM 36 H MOL A 0 -5.738 -1.039 -2.748 H

ATOM 37 H MOL A 0 -7.938 -0.320 -2.226 H

ATOM 38 H MOL A 0 -0.690 -1.190 0.400 H

ATOM 39 H MOL A 0 -2.016 3.381 0.058 H

ATOM 40 H MOL A 0 4.371 1.997 -0.223 H

ATOM 41 H MOL A 0 6.702 2.073 -0.397 H

ATOM 42 H MOL A 0 5.744 -3.071 0.221 H

ATOM 43 H MOL A 0 3.621 -2.132 0.270 H

B. Anionic form

TITLE tr-TTP-Neg_B3LYP-6

REMARK 4

REMARK 4 COMPLIES WITH FORMAT V. 2.2, 16-DEC-1996

REMARK 6 created by ArgusLab version 4.0.1

REMARK 6 http://www.arguslab.com

REMARK 6

ATOM 1 C MOL A 0 5.538 -0.788 -0.779 C

ATOM 2 C MOL A 0 4.197 -1.007 -1.484 C

ATOM 3 O MOL A 0 3.839 -1.085 0.881 O

ATOM 4 C MOL A 0 5.247 -1.233 0.669 C

ATOM 5 O MOL A 0 6.558 -1.574 -1.396 O

ATOM 6 C MOL A 0 5.985 -0.412 1.716 C

ATOM 7 O MOL A 0 7.387 -0.598 1.446 O

ATOM 8 C MOL A 0 3.140 -1.000 -0.357 C

ATOM 9 C MOL A 0 -0.115 0.874 -0.158 C

ATOM 10 C MOL A 0 0.845 -0.102 -0.246 C

ATOM 11 N MOL A 0 2.218 0.144 -0.330 N

ATOM 12 C MOL A 0 2.719 1.432 -0.283 C

ATOM 13 N MOL A 0 1.746 2.411 -0.197 N

ATOM 14 C MOL A 0 0.346 2.278 -0.128 C

ATOM 15 O MOL A 0 -0.351 3.278 -0.050 O

ATOM 16 O MOL A 0 3.919 1.709 -0.321 O

ATOM 17 C MOL A 0 -1.491 0.591 -0.094 C

ATOM 18 C MOL A 0 -2.688 0.325 -0.045 C

ATOM 19 C MOL A 0 -4.062 0.009 0.011 C

ATOM 20 C MOL A 0 -4.989 1.083 0.123 C

ATOM 21 C MOL A 0 -6.364 1.082 0.194 C

ATOM 22 C MOL A 0 -7.366 0.031 0.181 C

ATOM 23 C MOL A 0 -7.024 -1.467 0.077 C

ATOM 24 C MOL A 0 -5.666 -1.971 -0.022 C

ATOM 25 C MOL A 0 -4.430 -1.363 -0.051 C

ATOM 26 O MOL A 0 -8.566 0.337 0.258 O

ATOM 27 O MOL A 0 -7.969 -2.271 0.080 O

ATOM 28 H MOL A 0 3.992 -0.256 -2.249 H

ATOM 29 H MOL A 0 4.210 -1.993 -1.961 H

ATOM 30 H MOL A 0 5.811 0.270 -0.800 H

ATOM 31 H MOL A 0 2.484 -1.871 -0.442 H

ATOM 32 H MOL A 0 5.525 -2.294 0.770 H

ATOM 33 H MOL A 0 7.374 -1.410 -0.894 H

ATOM 34 H MOL A 0 5.698 0.642 1.614 H

ATOM 35 H MOL A 0 5.734 -0.761 2.726 H

ATOM 36 H MOL A 0 7.899 -0.000 2.010 H

ATOM 37 H MOL A 0 0.568 -1.149 -0.260 H

ATOM 38 H MOL A 0 2.096 3.363 -0.169 H

ATOM 39 H MOL A 0 -4.527 2.069 0.158 H

ATOM 40 H MOL A 0 -6.832 2.062 0.277 H

ATOM 41 H MOL A 0 -5.659 -3.058 -0.086 H

ATOM 42 H MOL A 0 -3.586 -2.048 -0.135 H

CHAPTER 3

Part-B

Synthesis of *Tropo-aminoiminyl* Conjugated and Fused 2'-deoxypyrimidine Nucleoside Phosphoramidites and their Incorporation into DNA-oligonucleotides

Chapter 3B. Synthesis of *Tropo-aminoiminyl* Conjugated and Fused 2'-deoxypyrimidine Nucleoside Phosphoramidites and their Incorporation into DNA-oligonucleotides

3B.1 Introduction

An approach to regulate specific-gene expression is explored by using chemically modified single strand antisense oligonucleotides (ASOs) such as DNA/RNA/chimeric forms.[1] These oligonucleotides form duplex with complimentary sequence of target messenger RNA (mRNA) and result in mRNA degradation mediated by enzyme machinery to accomplish target-gene expression.[2] Most of the enzyme functions depend on metal ions such as zinc, magnesium, copper, iron etc., and these metal ions play a vital role in expression of functional proteins, hence, the metallozymes have an important role in gene expression. In recent times, to understand the interactive microenvironment of the duplexes, many modified DNA oligonucleotides with small fluorescent nucleosides have been synthesized and probed the microenvironment[3-5] When the synthetic oligonucleotides interact with enzymes having metals (metalloenzymes) with high selectivity and specificity, then the function of metallozymes can affect, which further influence the gene expression. Thus modified oligonucleotides containing metal-chelating residues which interact with metalloenzymes with high selectivity/specificity can provide a powerful research tool to control the protein expression. Hence, synthesis of oligonucleotides containing metal chelating and environmental sensitive chromophores/fluorophores in controlling the gene expression would be of considerable interest.

3B.2 Objective of our work

Tropolone is a non-benzenoid seven-member non-classical aromatic compound. Tropolone and its naturally occurring derivatives (troponoids) exhibit broad spectrum of biological activities[6] These are demonstrated as metalloprotein inhibitors such as Zinc dependent metalloprotiens including carboxypeptidase A, thermolysin, matrix metalloproteases).[7-10] The mechanism of zinc dependent metalloprotein, thermolysin, inhibition is proposed via metal-coordination.[10] Tropolone has also been reported to be a distinctive metal-chelating scaffold, especially for $Cu^{2+}/Zn^{2+}/Ni^{2+}$, because of the α-hydroxy carbonyl functional group.[11, 12] The tropolone aminoimine (tropoaminoimine) derivatives are members of another class of compounds which exhibit unique aromatic properties which are different from tropolone[13] Like tropolone, the tropo-aminoimine derivatives are also versatile metal-chelators.[14] With these versatile metal-chelating properties of tropolone and tropo-aminoimine derivatives, we hypothesize, these derivatives could be suitable for site-specific incorporation into DNA oligonucleotides to interact with specific metallozymes. Herein, we planned to conjugate tropo-dibenzylaminoimine derivative to pyrimidine nucleoside and incorporate in DNA (Figure 3B.1).

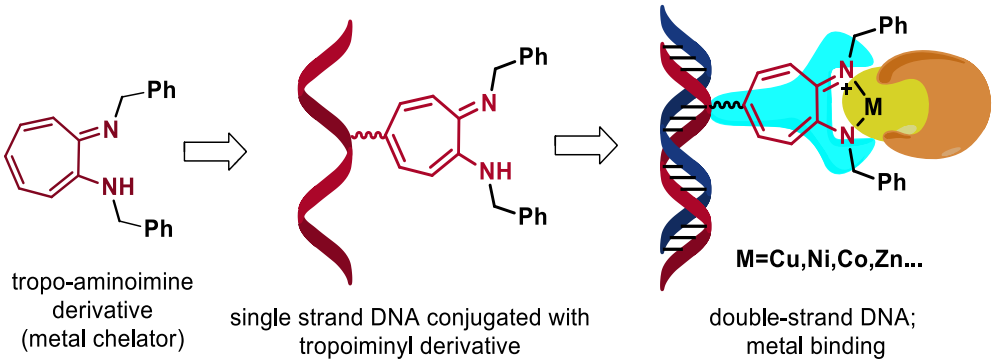
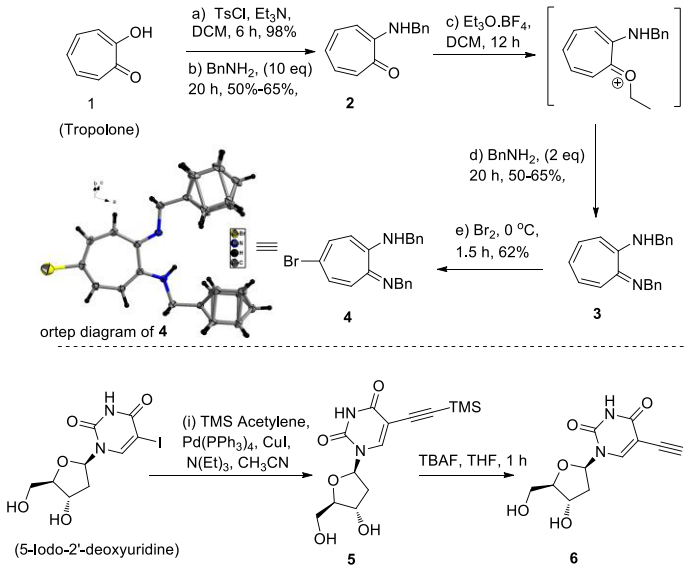


Figure 3B.1. Proposed hypothesis for tropo-aminoimine modified DNA oligonucleotides to interact with metallozymes.

3B.3 Results and Discussion

Our approach to the synthesis of 5-(N,N-dibenzyl 2-aminotropoiminyl) thymidine nucleoside involves sonogashira coupling of 5-bromo-N,N-dibenzyl-2-aminotropoimine intermediate **5** with 2'-deoxy-5-ethynyluridine **6** nucleoside. Nucleoside **6** was synthesized from commercially available 5-Iodo-2'-deoxyuridine **1** in two steps following previous reports. The synthesis of intermediate **4** involves five synthetic steps starting from commercially available tropolone. Compound **3** was synthesized by following the literature procedure. Bromination of Compound **3** with molecular bromine at low temperature exclusively brominated at 4th position to obtain required intermediate **4** with 62% yield. The structure of intermediate **5** is also confirmed by its single crystal X-Ray studies (Scheme 3B.1).

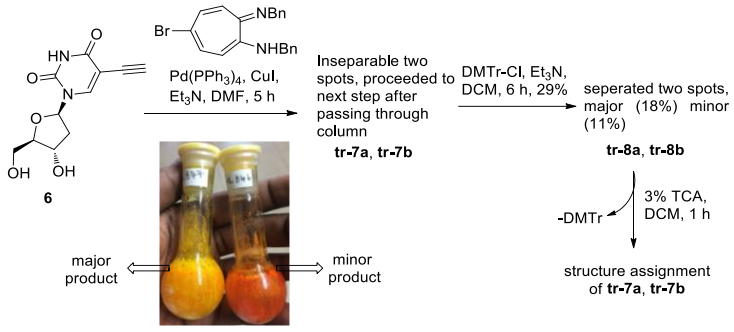
Scheme 3B.1. Synthesis of precursors for tropo-aminoiminyl 2'-deoxyuridine nucleoside.



In chapter 3A, 5-iodotropolone derivative, a close analogue of intermediate **4**, is successfully coupled with nucleoside **6**. Under similar conditions, the coupling of 5-bromo-N,N-dibenzyl-2-aminotropoimine intermediate **4** with 2'-deoxy-5-ethynyluridine **6** nucleoside resulted in two

products ***tr*-7a** and ***tr*-7b** which have very close retention values on TLC. These two spots are inseparable in our hands even under various silica gel column purification conditions in different solvent systems. However, we proceeded to the next step just by passing through silica gel column to remove metal catalysts and salts. Eventually, these two inseparable spots were treated with DMTr-Cl for 5'-OH protection. The reaction proceeded with 29% isolated yield and separated two clean products, ***tr*-8a** as major product (18%) with high retention value and ***tr*-8b** as minor product (11%) with low retention value (Scheme 3B.2).

Scheme 3B.2. Synthesis of tropo-aminoiminyl 2'-deoxyuridine derivatives.



Surprisingly, from our initial mass spectral assessment of ***tr*-8a** and ***tr*-8b**, the mass of both major (***tr*-8a**) and minor (***tr*-8a**) products are same as expected products. Moreover, structure assessment of ***tr*-8a** and ***tr*-8b** from their respective NMR spectra made difficult in aromatic region due to the presence of DMTr group. However, ***tr*-7a** and ***tr*-7b** were not characterized due to very close retention values and later steps made difficult in structure assignment from NMR spectroscopy. Therefore, ***tr*-8a** and ***tr*-8b** are treated with 3% TCA in two separate reactions and removed their respective DMTr protecting groups to obtain ***tr*-7a** and ***tr*-7b**. These products were purified through column chromatography and preceded for structure assignment.

At first, the mass spectra of ***tr*-7a** and ***tr*-7b** is depicted in Figure 3B.2, the mass (M) of ***tr*-7a** and ***tr*-7b** appeared to be same as expected product M = 550. The mass of ***tr*-7a** appeared as sodium adduct, m/z = 573.209 $[M+Na]^+$ (Figure 3B.2A), while ***tr*-7b** appeared as m/z = 551.230 $[M+H]^+$ (Figure 3B.2B). Importantly, in mass spectra of ***tr*-7a** and ***tr*-7b**, the peak at m/z = 435 and m/z = 457 (Figure 3B.2) is analyzed as nucleobase, B, fragment ion of as $BHNa^+$ (m/z = 457) and BH_2^+ (m/z = 435) respectively, this peak is resulted from the cleavage of gylcosidic bond (C1'-N1 bond cleavage), this fragment ion is more prominent in ***tr*-7b**. This indicates, sugar ring in both ***tr*-7a** and ***tr*-7b** are preserved after coupling reaction. Moreover, these strongly suggest both ***tr*-7a** and ***tr*-7b** are structural isomers and they differ in nucleobase structure.

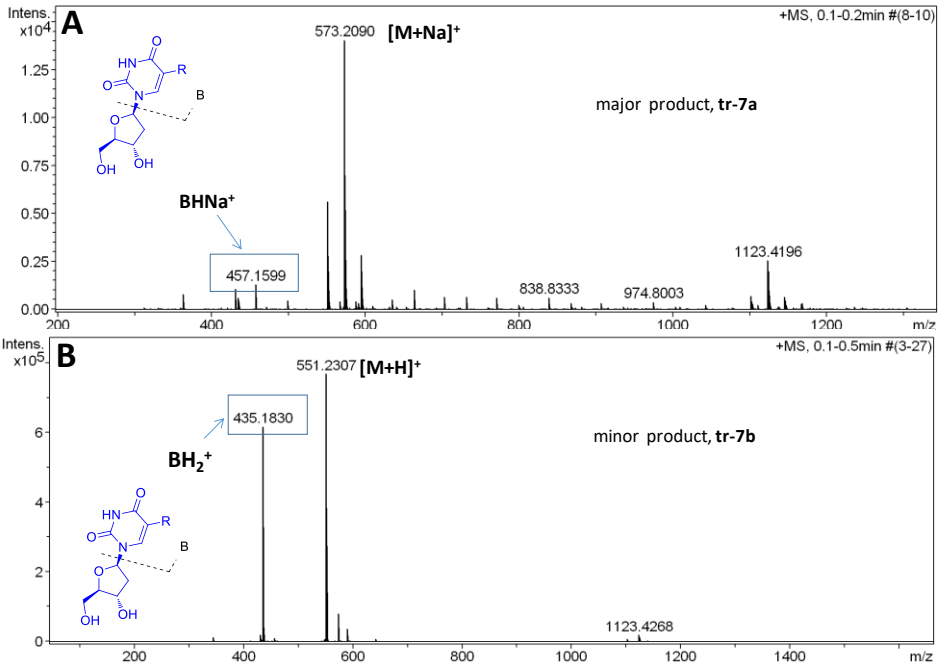


Figure 3B.2. ESI-Mass spectra of major product, ***tr*-7a** (A) minor product, ***tr*-7b** (B) and structural mass analysis of observed nucleobase fragment ions mass peaks.

In literature, during the synthesis of 5-alkynyl uridines via palladium/copper-catalyzed coupling reaction in presence of base, formation of bicyclic furanopyrimidines are also observed as minor

product.[15, 16] Importantly, 5-alkynyl uridines and their bicyclic furanopyrimidine are structural isomers and have same molecular weights. Since, the coupling of 5-bromo-N,N-dibenzyl-2-aminotropoimine intermediate **5** and 2'-deoxy-5-ethynyluridine **6** performed via palladium/copper-catalyzed coupling conditions, the formation of bicyclic furanopyrimidine product is highly predictable. Herein, this coupling reaction also resulted in two products having similar molecular masses. This suggests that along with expected product, furanopyrimidine product is also formed as minor product. To confirm the structure of these products, we carefully analyzed their $^{1}H/^{13}C$-NMR/^{1}H-^{1}H-COSY/D_2O exchange NMR experiments.

The ^{1}H-NMR spectra of structural isomers ***tr*-7a** and ***tr*-7b** was recorded in DMSO-d6 and their spectra are provided in Figure 3B.3. The ^{1}H-NMR spectra of mafor product (***tr*-7a)** featured broad peaks at δ10 ppm and δ5.1-5.3 ppm, from our D_2O exchange ^{1}H-NMR experiments these broad peaks was assigned as two NH and two OH protons respectively (D_2O exchange ^{1}H-NMR spectra provided in Appendix). All the others protons were assigned from analysis of ^{1}H-COSY spectra (Appendix). Tropolone ring attributed four protons as two doublets; in which tropolone ring protons-3/7 appeared at δ6.45 ppm and protons-4/6 appeared at δ6.9 ppm. These ^{1}H-NMR chemical shifts are consistent with its precursor **5.** Importantly, 6-H proton of thymidine ring appeared at δ8.25 ppm. Similarly, the ^{1}H-NMR spectra of ***tr*-7b** featured broad peaks at δ8.5 ppm and δ5.1-5.3 ppm, from our D_2O exchange experiments these broad peaks are assigned as benzyl-NH proton and two OH protons respectively. All the others protons were assigned from ^{1}H-COSY spectra (Appendix). Unlike ***tr*-7a,** the tropolone ring protons appeared downfield, tropolone ring protons-3/7 appeared at δ6.7 ppm and protons-4/6 appeared at δ7.31 ppm which is overlapped with phenyl protons. The 6-H proton of thymidine ring is also more deshielded and appeared at δ8.72 ppm. Importantly, a new peak at δ6.9 ppm is appeared only in ***tr*-7a** which is not exchangeable

with D_2O and not involved in any resonance in ^{1}H-COSY NMR (spectra provided in Appendix). The ^{1}H-NMR spectral analysis of ***tr*-7a** and ***tr*-7b,** suggest there are no chemical modification in sugar ring and tropolone residue. Importantly, the absence of broad NH peak for thymidine residue at δ10 ppm and appearance of isolated peak at δ6.9 ppm and it is not exchangeable with D_2O in minor product strongly suggests the bicyclic furanopyrimidine product as ***tr*-7b**.

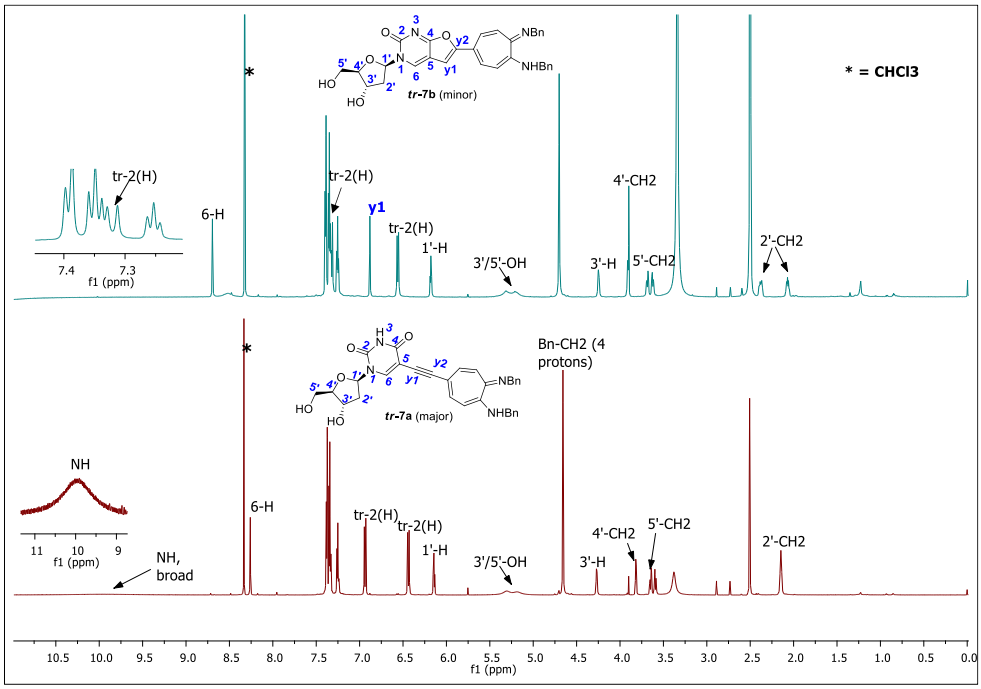


Figure 3B.3. ^{1}H-NMR spectra of ***tr*-7a** and ***tr*-7b** nucleosides and their peak assignments.

From mass and ^{1}H-NMR analysis, strongly indicate sugar ring and tropolone residues are conserved in both major and minor products (***tr*-7a** and ***tr*-7b**). These indicate that both major and minor products (***tr*-7a** and ***tr*-7b**) exhibit structural changes only at thymine and alkyne residues. The ^{13}C-NMR spectra of these structural isomers, (major and minor products) exhibit prominent changes which helped more in structure determination major and minor products as ***tr*-7a** and ***tr*-7b** (Figure 3B.4). The carbonyl carbons C-2 and C-4 in thymine residue of ***tr*-7a** appear at δ150 ppm and δ162 ppm respectively. The alkyne carbons (assigned as y1 and y2) appeared at δ96 ppm

and δ80 ppm**, w**hich are consistent with expected product. However, the carbonyl and alkyne carbon peaks in ***tr*-7b** significantly shifted to higher chemical shift. The carbonyl carbon C-4 and C-2 shifted to δ171 ppm and δ154 ppm respectively. Importantly, the alkyne carbon, y2 shifted to δ156 ppm which suggests this carbon experienced significant desheilding. Similarly, all the other carbon peaks of ***tr*-7a** and ***tr*-7b** are assigned. A prominent shift in y2 carbon strongly suggests the absence of alkyne functionality and presence of bicyclic furanopyrimidine unit. This analysis is further strongly supported by ^{1}H-/^{13}C-NMR spectral analysis of closely related compounds.[17-19]

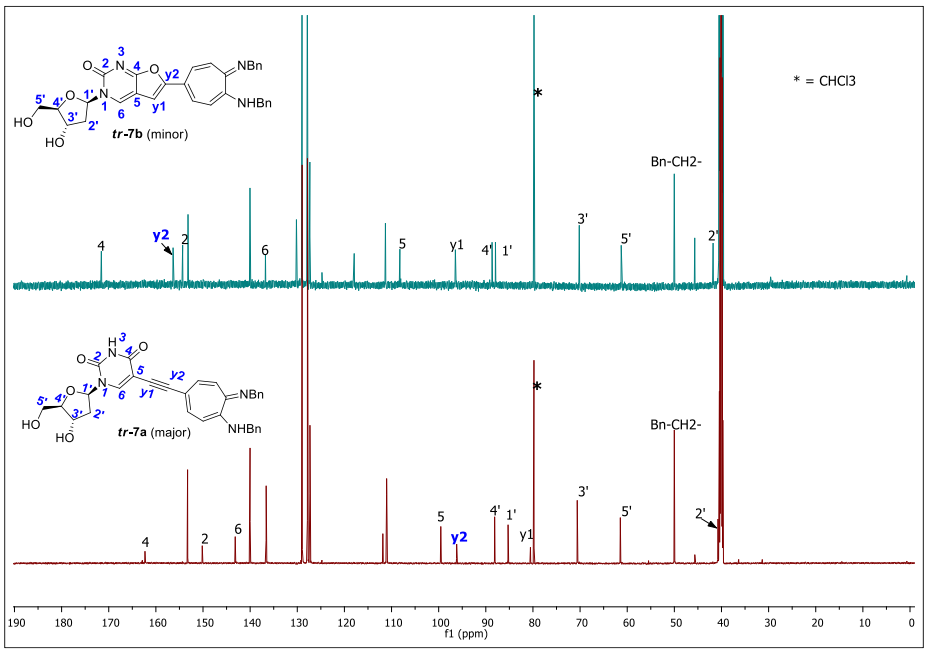
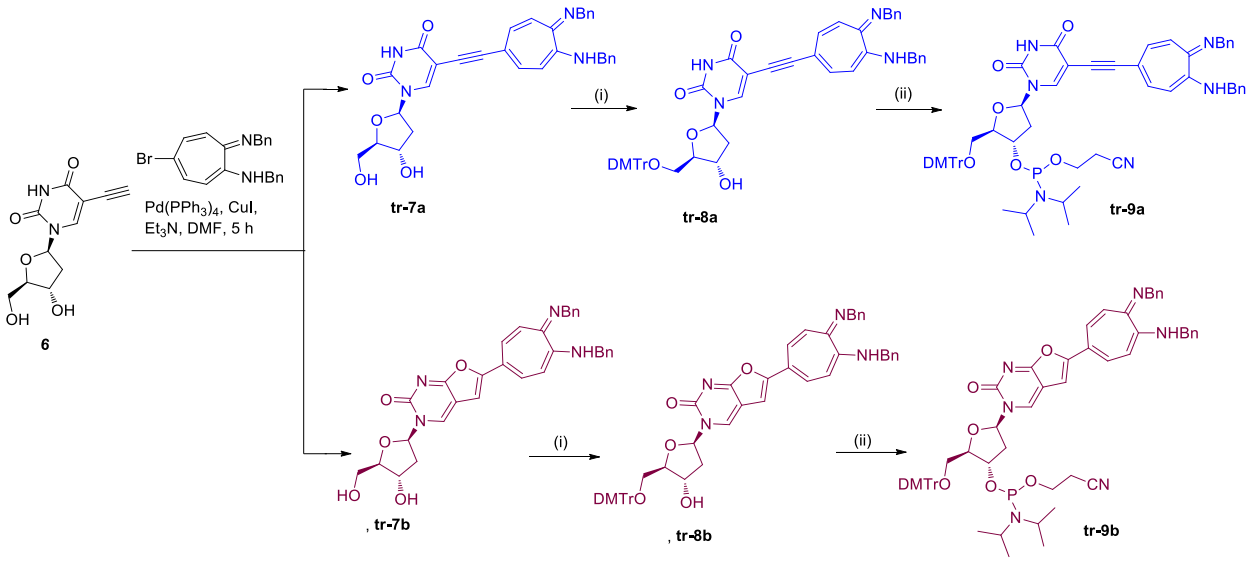


Figure 3B.4. ^{13}C-NMR spectra of ***tr*-7a** and ***tr*-7b** nucleosides and their nucleobase peak assignments.

Synthesis of phosphoramidites. After successful synthesis and structure assignment of major and minor products as, ***tr*-7a** and ***tr*-7b**, we planned to synthesize their phosphoramidites and incorporate into DNA oligonucleotides. For this, the 5'-OH of ***tr*-7a** and ***tr*-7b** is protected with DMTr group (**8a/8b**). After successful DMTr protection, 3'-*OH* of both ***tr*-8a** and ***tr*-8b** are

converted into 3'-*O*-phosphoramidite by treating with phosphoramidating reagent and obtained ***tr*-9a** in 91% yield and ***tr*-9b** in 81% yield (Scheme 3B.3).

Scheme 3B.3. Synthesis of alkynyl (***tr*-9a**) and furinyl (***tr*-9b**) phosphoramidites.



UV-Visible and Fluorescence studies. Tropolone exhibit strong absorbtion in the range of λ340-450 nm and also a very weak fluorophore. From physical appearance, major product ***tr*-7a** appeared as yellow and minor product ***tr*-7b** appeared as orange solids. The UV-/fluorescence spectra of ***tr*-7a** and ***tr*-7b** are provided in Figure 3B.5. Similar to tropolone, the UV-spectra of ***tr*-7a** exhibit UV-absorbtion at λ400 nm and λ430 nm, while ***tr*-7b** appeared at higher wavelength, which exhibit absorbtion at λ425 nm and λ456 nm. This higher wavelength absorbtion (~30 nm) of ***tr*-7b** is possibly due to extended conjugation with furanopyrimidine residue.[20] These UV-absorption peaks in ***tr*-7a** and ***tr*-7b** derivatives are mainly due to the π-π* and n-π* transitions of tropo-aminoimine residue. The ***tr*-7a** and ***tr*-7b** derivatives also exhibit weak fluorescence upon excitation at λ_{ex} 395 and 420 nm respectively. Comparatively, ***tr*-7b** exhibit higher fluorescence

than ***tr*-7a**. The higher fluorescence of ***tr*-7b** is attributed from tropo-aminoimne and furanopyrimidine residues. In literature, furanopyrimidine and its analogue pyrrolepyrimidne analogues are studied as fluorescent probes.[21]

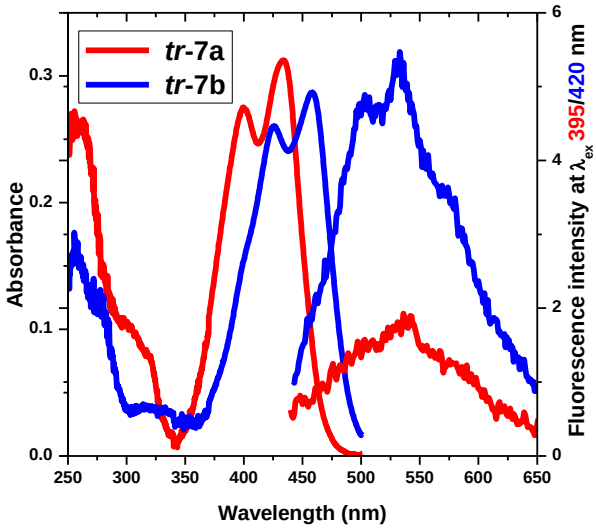


Figure 3B.5. UV-/fluorescence spectra of ***tr*-7a** and ***tr*-7b** in chloroform.

DNA oligonucleotide synthesis. After successful synthesis of phosphoramidites ***tr*-9a** and ***tr*-9b,** we incorporated these phosphoramidites into DNA oligonucleotide. The synthesized phosphoramidites ***tr*-9a** and ***tr*-9b** were incorporated in DNA oligonucleotide sequence (which we studied in section-A) under standard DNA synthesis conditions, except the coupling times of ***tr*-9a** and ***tr*-9b** increased to 10 min. The phosphoramidites, ***tr*-9a** and ***tr*-9b** are incorporated in DNA oligonucleotide at different positions having one incorporation, and two incorporations. After the synthesis, DNA oligonucleotides were cleaved from resin and incubated at 55 ^{o}C for 6 h in ammonium hydroxide for complete deprotection of cyanoethyl groups and nucleobase protecting groups. Further these were purified and desalted by passing through sephadex column and the

fractions containing DNA were characterized by ESI-MS and their mass spectra are provided in Appendix.

The DNA sequence, position of modifications, there observed ESI-mass and calculated mass for ***tr*-9a** incorporated DNA oligonucleotides, alkynyl DNA, (Alk-ONs) are provided in Table 3B.1. The mass of unmodified control DNA oligonucleotide C-ON1 appeared in charge state z = 3 with m/z = 3950, which is in agreement with calculated mass (m/z = 3950). Surprisingly, the mass of ***tr*-9a** modified DNA appeared less than their calculated DNA oligonucleotides mass. In alkynyl DNA oligonucleotides, Alk-ON2 and Alk-ON3 have single incorporated ***tr*-9a** residue and they differ in their position of modifications, have a mass difference of m/z 179 (M-179) from calculated mass (M). For example, the mass spectra of ALK-ON2 appeared at m/z 1358.975, in charge state z = 3, and this mass corresponds to the deconvoluted observed mass m/z = 4079.92, which is less than its calculated mass M = 4258.955 (Table 3B.1, Entry 2). The observed mass difference is m/z = 179.035. Similarly, Alk-ON4 has two ***tr*-9a** alkynyl incorporations and has mass difference of m/z = 358.351 from its calculated mass, again the difference for each unit is m/z 179 (179.17*2 units) (Table 3B.1, Entry 4).

Table 3B.1. DNA sequence, position of *tr-9a* incorporation and mass analysis data.

Entry	DNA	DNA sequence	Calculated mass (M)	Observed mass	Difference in mass
1	C-ON1	5'-CGC ATG TGT ACG C-3'	3950.626	3950.83	-
2	ALK-ON2	5'-CGC ATG XGT ACG C-3'	4258.955	4079.92	179.035
3	ALK-ON3	5'-CGC AXG TGT ACG C-3'	4258.955	4079.93	179.025
4	ALK-ON4	5'-CGC AXG XGT ACG C-3'	4567.331	4208.98	358.351

X = *tr*-9a incorporated nucleoside

Surprisingly, the mass of ***tr*-9b** modified DNA also appeared less than their calculated DNA oligonucleotide, mass analysis from mass spectra are provided in Table 3B.2 and their mass spectra are provided in Appendix. In ***tr*-9b** incorporated DNA, furyl DNA oligonucleotides, FUR-ON1/2

have single ***tr*-9b** modification and they differ in their position of incorporation (Table 3B.2, Entry 1-3). They have mass difference of m/z 180.1, (M-180) from their calculated mass (Table 3B.2, Column 6). Similarly, **FUR-ON3** has two ***tr*-9b** modifications and exhibit mass difference of m/z 360.44, (M-360), again the difference for each unit is also m/z 180 (180*2 units).

Table 3B.2. DNA sequence, position of ***tr*-9b** incorporation and mass analysis data.

Entry	**DNA**	**DNA sequence**	**Calculated mass (M)**	**Observed mass**	**Difference in mass**
1	FUR-ON1	5'-CGC ATG XGT ACG C-3'	4258.955	4078.81	180.145
2	FUR-ON2	5'-CGC AXG TGT ACG C-3'	4258.955	4078.82	180.135
3	FUR-ON3	5'-CGC AXG XGT ACG C-3'	4567.330	4206.89,	360.44

X = ***tr*-9b** incorporated nucleoside

Importantly, all the observed mass of ***tr*-9b** modified DNA (FUR-ON1/2/3/4) are approximately one unit less than ***tr*-9a** modified DNA (ALK-ON2/3/4) for each unit of modification. This is most possibly due to the conversion of furan residue into pyrrole residue and these functionalities differ only by 1 mass unit (Figure 3B.6). In literature reports, such kind of conversion is possible in presence of ammonia at elevated temperatures (~ 50-60 ^{o}C).[22, 23] A similar type of conversion is also reported in similar furanopyrimidine modified DNA oligonucleotides after ammonia treatment.[20, 24, 25]

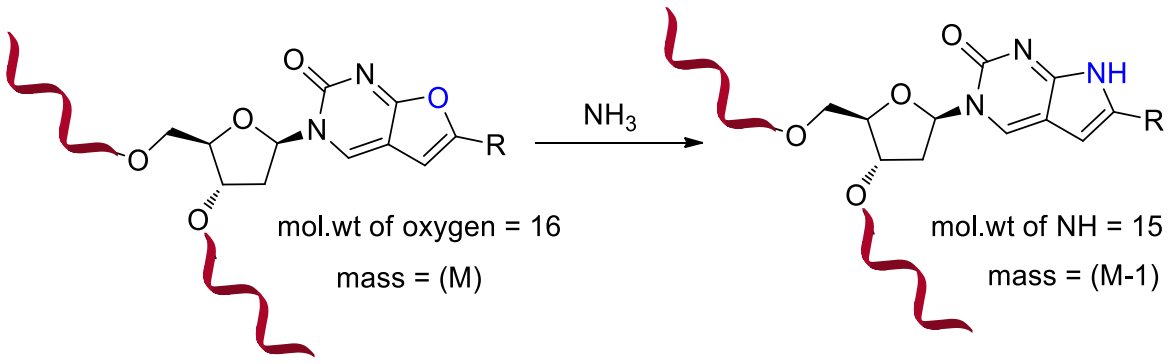


Figure 3B.6. Conversion of furan residue into pyrrole residue and changes in their molecular mass.

In both ***tr*-9a** and ***tr*-9b** incorporated DNA oligonucleotides, the mass of m/z 179 is less than calculated DNA mass for each unit of these modification. After careful structural analyses of these modifications, the mass of m/z 180 is attributed by two benzyl groups (m/z 90 for each benzyl group) present in these modifications. This indicates, the benzyl groups present in ***tr*-9a** and ***tr*-9b** nucleosides are deprotected during DNA synthesis conditions. After the DNA synthesis, oligonucleotides were cleaved from resin and incubated at 55 ^{o}C for 6 h in ammonium hydroxide for complete deprotection of cyanoethyl groups and nucleobase protecting groups. Under these conditions the benzylimine moiety is prone to hydrolyse. Most possibly, debenzylation resulted by the hydrolysis of benzylimine in ammonium hydroxide at 55 ^{o}C. Both ammonia and water can hydrolyze benzylimines and this result in formation of oxygen and imine functionality (Figure 3B.7). The mass of oxygen and imine functionality differ only by one mass unit. Generally in DNA/RNA mass analysis, it is difficult to assign DNA/RNA structures which differ by one mass unit, due to their ability to form diverse negatively charged species in mass spectroscopy. Hence it is difficult to assign the exact oxygen/imine functionality in ALK/FUR-DNA oligonucleotides. However, at nucleoside level, both ***tr*-8a** and ***tr*-8b** were treated with 3% TCA for 1-1.5 h to remove their respective DMTr protecting groups to obtain ***tr*-7a** and ***tr*-7b**. During these conditions benzyl groups are well tolerated and dint observed any debenzylated products.

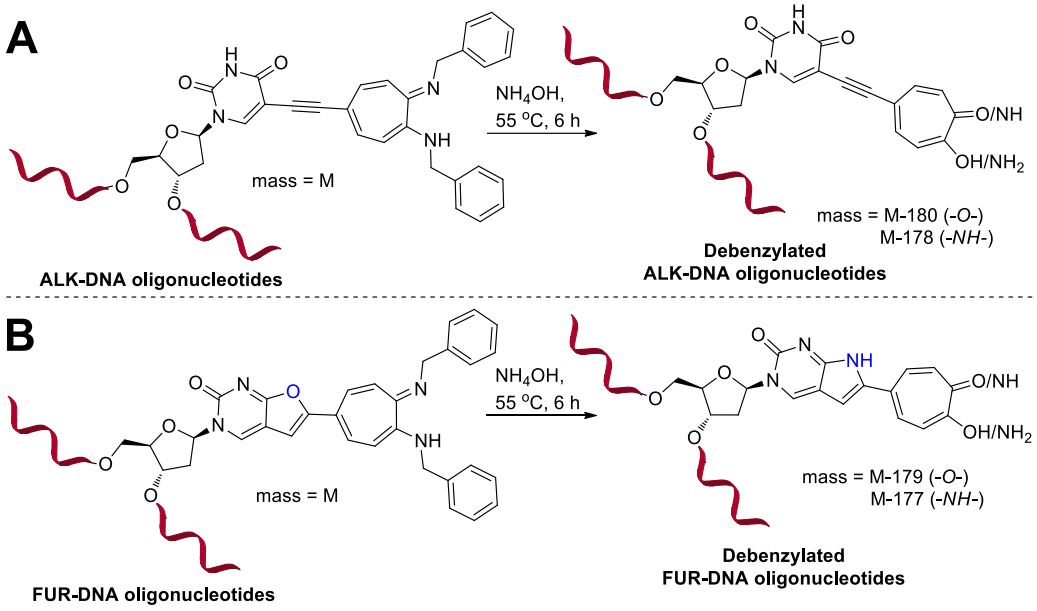


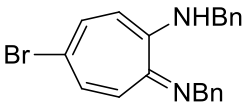
Figure 3B.7. Benzylimine hydrolysis in alkynyl (A) and furinyl (B) modified DNA oligonucleotides.

3B.4 Conclusion

This chapter describes the synthesis of metal chelating and environmental sensitive chromophore-tropo-aminoiminyl conjugated 2'-deoxyuridine nucleoside derivatives. The coupling reaction of 5-ethynyl 2'-deoxyuridine with corresponding 5-bromo tropolone diamine derivative produced two isomers and their structures are assigned as tropo-aminoiminyl-alkyne pyrimidine nucleoside and tropo-aminoiminyl bicyclic furanopyrimidine nucleoside. These isomers show characteristic UV spectra, and exhibit weak fluorescence. After successful synthesis of their phosphoramidites, these are incorporated into DNA oligonucleotides and characterized by ESI-MS. From their mass spectral analysis, the observed mass of tropoiminyl-alkyne DNA and tropoiminyl bicyclic furano DNA are less than the calculated mass. Though the tropoaminoiminyl-alkyne pyrimidine and tropoaminoiminylropoiminyl bicyclic furanopyrimidine nucleosides are stable in acidic condition,

but their deoxyoligonucleotides experienced an adverse effect of benzylimine hydrolysis during DNA synthesis and deprotection conditions. These tropo-aminoiminyl-alkyne pyrimidine /bicyclic furanopyrimidine nucleosides and their DNA could be potential candidates to screen the binding ability with different metals.

3B.5 Experimental section



4

Materials. All the required chemicals were obtained from commercial suppliers and used without any further purification. Pyridine and methylene chloride solvents were distilled over calcium hydride and stored under 4Å molecular sieves prior using for reactions. Tetrahydrofuran was freshly distilled over sodium and benzophenone for reactions. All the reactions were monitored by thin layer chromatography, visualized by UV and Ninhydrin. Column chromatography was performed with 230-400 mesh silica. MS and HRMS were obtained from Bruker micrOTOF-Q II Spectrometer. $^{1}H/^{13}C/^{31}P$ NMR spectra were recorded on Bruker AV-400 or 700 MHz at 298 K. ^{1}H and ^{13}C NMR chemical shifts were recorded in ppm downfield from tetramethyl silane or residual solvent peak. Splitting patterns are abbreviated as: s, singlet; d, doublet; dd, doublet of doublet; t, triplet; q, quartet; dq, doublet of quartet; m, multiplet.

Compounds **2/3/5/6** are synthesized by following literature reports.[26-28]

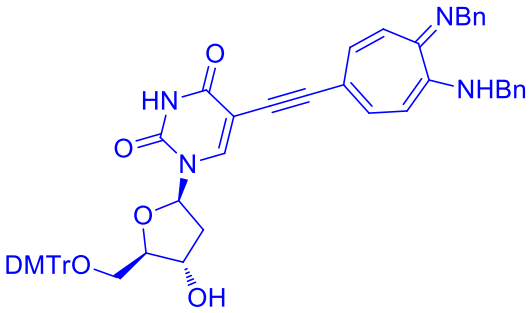
N-benzyl-7-(benzylimino)-4-bromocyclohepta-1,3,5-trienamine (**4**). The solution of compound **3** (0.84 g, 2.81 mmol) in freshly distilled CH_2Cl_2 (20 ml) is cooled in crushed ice/salt-bath. After brief cooling of solution, neat molecular bromine solution (0.18 ml, 3.3 mmol) is added drop wise and the reaction is monitored by TLC. After complete consumption of starting material (1.5 h), all

the solvents were removed under reduced pressure at 40 ^{o}C. The resulting brownish solid was purified with 2% MeOH/ CH_2Cl_2 to yield 0.66 g (65%) of compound **4** as yellow solid. R_f 0.6 (2% MeOH/CH_2Cl_2). ^{1}H NMR (700 MHz, $CDCl_3$) δ 7.35 – 7.31 (m, 8H), 7.26 (dt, *J* = 5.8, 3.1 Hz, 2H), 7.00 (d, *J* = 11.9 Hz, 2H), 6.15 (d, *J* = 11.8 Hz, 2H), 4.58 (s, 4H). ^{13}C NMR (176 MHz, $CDCl_3$) δ 152.48, 139.15, 135.71, 128.61, 127.30, 127.04, 111.51, 110.11, 50.37. HRMS (ESI-TOF) m/z: $[M + H]^+$ Calcd for $C_{21}H_{20}N_2Br$, 379.0804; Found 379.0841.

Sonogashira coupling reaction procedure:Compound **4** (1.7 g, 4.5 mmol) and Tetrakis (triphenylphosphine) palladium (0.26 g, 0.22 mmol)are dissolved in DMF (40 ml)and triethylamine (2 ml, 13.5 mmol) solution. To the resulting solution, CuI (0.38 g, 2 mmol) followed by compound **6** (1.24 g, 4.9 mmol) was added and stirred at room temperature. After 5 h, the reaction mixture is concentrated under reduced pressure. The resulting brown syrup was purified by column chromatography with 8% MeOH/ CH_2Cl_2. Two inseparable products, **tr-7a** and **tr-7b** were isolated, without further purification, preceded for next step.

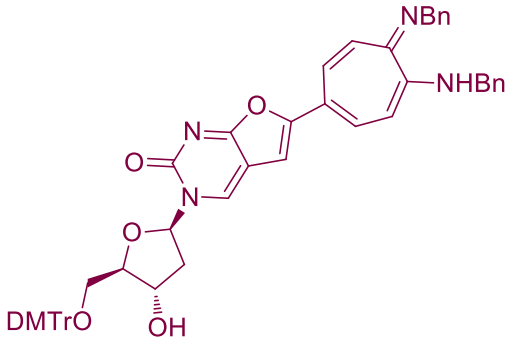
Two inseparable products, **tr-7a** and **tr-7b** from sonogashira coupling reaction (2.6 g, 4.8 mmol) was dissolved in dry CH_2Cl_2 (60 ml) and dry triethylamine (3.4 ml, 24.3 mmol) followed by the addition of DMTr-Cl (2.14 g, 6.32 mmol) was added and stirred at room temperature for 6 h. The red crude reaction mixture was concentrated to dryness and then subjected to column chromatography. Silica-gel prewashed with 0.5% triethylamine/70% EtOAc/n-Hexanes and eluted with 70% EtOAc/n-Hexanes –Neat EtOAc to yield 0.75 g (18%) and 0.43 g (11%) of compounds **tr-8a** and **tr-8b** as yellow and orange solids respectively. R_f 0.3 for ***tr*-8a** and 0.25 for ***tr*-8b** (75% EtOAc/n-Hexanes).

5-((4-(benzylamino)-5-(benzylimino)cyclohepta-1,3,6-trien-1-yl)ethynyl)-1-((2R,4S,5R)-5-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-4-hydroxytetrahydrofuran-2-yl)pyrimidine-2,4(1H,3H)-dione (***tr*-8a).** ^{1}H NMR (700 MHz, $CDCl_3$) δ 8.14 (s, 1H), 7.42 (d, *J* = 7.8 Hz, 2H), 7.37 – 7.29 (m, 11H), 7.28 – 7.24 (m, 2H), 7.22 (t, *J* = 7.4 Hz, 2H), 7.11 (t, *J* = 7.0 Hz, 1H), 6.81 – 6.75 (m, 4H), 6.44 (d, *J* = 11.6 Hz, 2H), 6.40 – 6.33 (m, 1H), 6.09 (d, *J* = 9.2 Hz, 2H), 4.66 – 4.36 (m, 5H), 4.13 (d, *J* = 2.3 Hz, 1H), 3.66 (d, *J* = 6.9 Hz, 6H), 3.48 (d, *J* = 9.8 Hz, 1H), 3.26 (d, *J* = 9.0 Hz, 1H), 2.59 – 2.43 (m, 1H), 2.39 – 2.24 (m, 1H). ^{13}C NMR (176 MHz, $CDCl_3$) δ 161.71, 158.60, 153.04, 149.56, 149.44, 144.49, 141.13, 139.19, 136.88, 136.28, 135.59, 135.55, 129.98, 129.94, 128.61, 128.06, 127.95, 127.34, 127.04 , 123.90, 113.38, 110.52, 101.26, 87.03, 86.79, 85.82, 72.30, 63.54, 55.23, 50.20. HRMS (ESI-TOF) m/z: $[M + H]^+$ Calcd for $C_{53}H_{49}N_4O_7$, 853.3596; Found 853.3573.



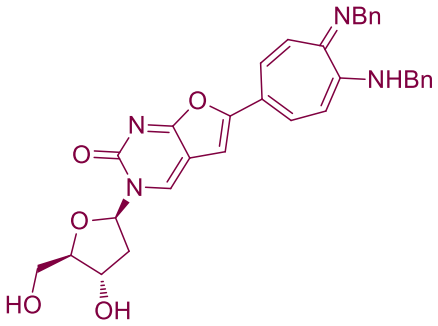
tr-8a

6-(4-(benzylamino)-5-(benzylimino)cyclohepta-1,3,6-trien-1-yl)-3-((2R,4S,5R)-5-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-4-hydroxytetrahydrofuran-2-yl)furo[2,3-d]pyrimidin-2(3H)-one (***tr*-8b**): ^{1}H NMR (700 MHz, $CDCl_3$) δ 8.14 (s, 1H), 7.42 (d, *J* = 7.8 Hz, 2H), 7.37 – 7.29 (m, 11H), 7.28 – 7.24 (m, 2H), 7.22 (t, *J* = 7.4 Hz, 2H), 7.11 (t, *J* = 7.0 Hz, 1H), 6.81 – 6.75 (m, 4H), 6.44 (d, *J* = 11.6 Hz, 2H), 6.40 – 6.33 (m, 1H), 6.09 (d, *J* = 9.2 Hz, 2H), 4.66 – 4.36 (m, 5H), 4.13 (d, *J* = 2.3 Hz, 1H), 3.66 (d, *J* = 6.9 Hz, 6H), 3.48 (d, *J* = 9.8 Hz, 1H), 3.26 (d, *J* = 9.0 Hz, 1H), 2.59 – 2.43 (m, 1H), 2.39 – 2.24 (m, 1H). ^{13}C NMR (176 MHz, $CDCl_3$) δ 171.59, 158.81, 158.76, 157.10, 155.01, 153.03, 144.13,



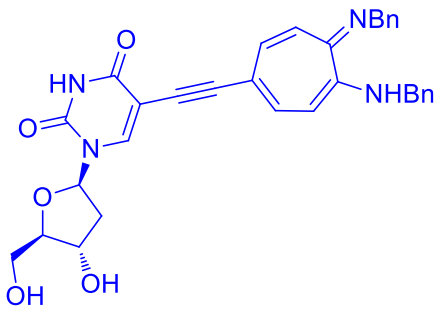
, **tr-8b**

139.06, 135.88, 135.49 , 135.31, 130.33, 130.19, 129.24, 128.75, 128.48, 128.26, 127.44, 127.24, 118.15, 113.51 113.27, 110.90, 108.95, 94.95, 87.89, 87.22, 86.14, 69.21, 61.8, 55.33, 50.41, 42.32, 29.81. HRMS (ESI-TOF) m/z: $[M + H]^+$ Calcd for $C_{53}H_{49}N_4O_7$, 853.3596; Found 853.3597.



, **tr-7b**

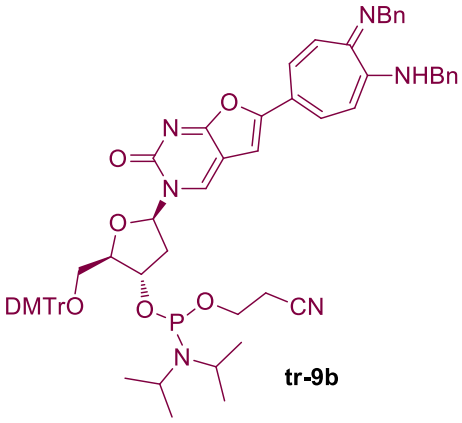
5-((4-(benzylamino)-5-(benzylimino)cyclohepta-1,3,6-trien-1-yl)ethynyl)-1-((2R,4S,5R)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)pyrimidine-2,4(1H,3H)-dione (***tr*-7a**). Compound ***tr*-8a** (0.05 g) was treated with 3% trichloroacetic acid (0.5 ml) and allowed to stir for 60 min. The solvents were removed under vacuum and the resulting pale orange residue was purified via silica-gel chromatography with 8% MeOH/CH_2Cl_2 to produce compound **9a** as a yellow solid. R_f 0.2 (8% MeOH/CH_2Cl_2). ^{1}H NMR (700 MHz, DMSO) δ 8.26 (s, 1H), 7.42 – 7.30 (m, 8H), 7.25 (t, *J* = 7.2 Hz, 2H), 6.94 (d, *J* = 11.7 Hz, 2H), 6.44 (d, *J* = 11.8 Hz, 2H), 6.14 (t, *J* = 6.6 Hz, 1H), 5.24 (br, 2H), 4.66 (s, 4H), 4.27 (d, *J* = 3.5 Hz, 1H), 3.82 (d, *J* = 3.2 Hz, 1H), 3.62 (ddd, *J* = 38.5, 11.8, 3.1 Hz, 2H), 2.14 (d, *J* = 5.9 Hz, 2H). ^{13}C NMR (176 MHz, DMSO) δ 162.31, 153.27, 150.16, 143.16, 140.02, 136.60, 128.97, 127.84, 127.31, 111.85, 111.04, 99.55, 96.19, 88.12, 85.27, 80.53 79.78, 70.60, 61.46, 50.01, 40.74, 40.57. HRMS (ESI-TOF) m/z: $[M + Na]^+$ Calcd for $C_{32}H_{30}N_4O_5Na$, 573.2108; Found 573.2085.



tr-7a

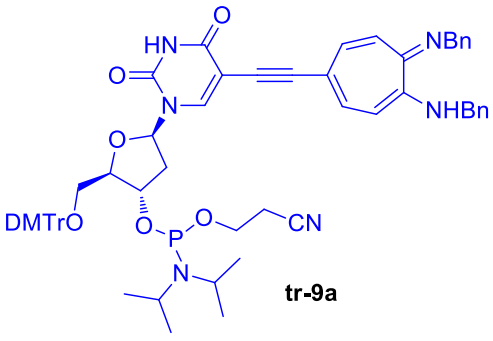
6-(4-(benzylamino)-5-(benzylimino)cyclohepta-1,3,6-trien-1-yl)-3-((2R,4S,5R)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)furo[2,3-d]pyrimidin-2(3H)-one (***tr*-7b**). Under similar conditions compound **tr-7b** was prepared from Compound ***tr*-8b** (0.05 g). R_f 0.15 (10% MeOH/CH_2Cl_2). ^{1}H NMR (700 MHz, DMSO) δ 8.70 (s, 1H), 8.50 (br, 1H), 7.46 – 7.16 (m, 10H),

6.88 (s, 1H), 6.56 (d, J = 11.8 Hz, 2H), 6.18 (t, J = 6.1 Hz, 1H), 5.26 (br, 2H), 4.70 (s, 4H), 4.25 (d, J = 3.4 Hz, 1H), 4.00 – 3.81 (m, 2H), 3.65 (dd, J = 44.1, 8.9 Hz, 2H), 2.44 – 2.30 (m, 1H), 2.16 – 1.95 (m, 1H). ^{13}C NMR (176 MHz, DMSO) δ 171.57, 156.35, 154.32, 153.18, 140.04, 136.75, 130.18, 128.99 , 127.86, 127.33 , 124.76, 117.92, 111.30, 108.23, 96.46, 88.67, 87.95, 70.17, 61.27, 50.01, 45.67, 41.78, 40.58 . HRMS (ESI-TOF) m/z: [M + H] $^{+}$ Calcd for $C_{32}H_{31}N_4O_5$, 551.2289; Found 551.2307.



tr-9b

(2R,3S,5R)-5-(5-((4-(benzylamino)-5-(benzylimino)cyclohepta-1,3,6-trien-1-yl)ethynyl)-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)-2-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)tetrahydrofuran-3-yl (2-cyanoethyl) diisopropylphosphoramidite (**9a**). 2-Cyanoethyldiisopropylchlorophosphoramidite (0.13 ml, 0.53 mmol) was added at room temperature to a stirred solution of compound **8a** (0.3 g, 0.35 mmol) in freshly distilled dry THF (2 ml) and *N,N*-Diisopropyl ethylamine (0.3 ml, 1.7 mmol). After stirred for 2.5 h, the crude reaction mixture was directly subjected to column chromatography (silica-gel was prewashed with 0.5% triethylamine/80% EtOAc/ n-Hexanes and eluted with 80% EtOAc/ n-Hexanes) to yield 0.34 g (91%) of compound **9a** as yellow foam. R_f 0.6 (70% EtOAc/ n-Hexane). ^{31}P NMR (162 MHz, $CDCl_3$) δ 149.01 (s), 148.53 (s). HRMS (ESI-TOF) m/z: [M + H]$^{+}$ Calcd for $C_{62}H_{66}N_6O_8P$, 1053.4674; Found 1053.4603.



tr-9a

(2R,3S,5R)-5-(6-(4-(benzylamino)-5-(benzylimino)cyclohepta-1,3,6-trien-1-yl)-2-oxofuro[2,3-d]pyrimidin-3(2H)-yl)-2-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)tetrahydrofuran-3-yl (2-cyanoethyl) diisopropylphosphoramidite (**9b**). Under similar conditions compound **8b** (0.2 g,

0.24 mmol) produced 0.21 g (81%) of compound **9b** as orange foam. R_f 0.5 (80% EtOAc/ n-Hexane). ^{31}P NMR (162 MHz, $CDCl_3$) δ 149.84 (s), 149.04 (s). HRMS (ESI-TOF) m/z: $[M + H]^+$ Calcd for $C_{62}H_{66}N_6O_8P$, 1053.4674; Found 1053.4641.

3B.6 Reference

1. Rinaldi, C.; Wood, M. J., Antisense oligonucleotides: the next frontier for treatment of neurological disorders. *Nat. Rev. Neurol.* **2018,** *14*, 9-21.
2. Opalinska, J. B.; Gewirtz, A. M., Nucleic-acid therapeutics: basic principles and recent applications. *Nat. Rev. Drug Discov.* **2002,** *1*, 503-514.
3. Shin, D.; Sinkeldam, R. W.; Tor, Y., Emissive RNA alphabet. *J. Am. Chem. Soc.* **2011,** *133*, 14912-14915.
4. Dai, N.; Kool, E. T., Fluorescent DNA-based enzyme sensors. *Chem. Soc. Rev.* **2011,** *40*, 5756-5770.
5. Xu, W.; Chan, K. M.; Kool, E. T., Fluorescent nucleobases as tools for studying DNA and RNA. *Nat. Chem.* **2017,** *9*, 1043.
6. Chung, S.; Himmel, D. M.; Jiang, J.-K.; Wojtak, K.; Bauman, J. D.; Rausch, J. W.; Wilson, J. A.; Beutler, J. A.; Thomas, C. J.; Arnold, E., Synthesis, activity, and structural analysis of novel α-hydroxytropolone inhibitors of human immunodeficiency virus reverse transcriptase-associated ribonuclease H. *J. Med. Chem.* **2011,** *54*, 4462-4473.
7. Morita, Y.; Matsumura, E.; Okabe, T.; Shibata, M.; Sugiura, M.; Ohe, T.; Tsujibo, H.; Ishida, N.; Inamori, Y., Biological activity of tropolone. *Biol. Pharm. Bull.* **2003,** *26*, 1487-1490.

8. Jacobsen, F. E.; Lewis, J. A.; Heroux, K. J.; Cohen, S. M., Characterization and evaluation of pyrone and tropolone chelators for use in metalloprotein inhibitors. *Inorg. Chim. Acta.* **2007,** *360*, 264-272.

9. Jacobsen, J. A.; Fullagar, J. L.; Miller, M. T.; Cohen, S. M., Identifying chelators for metalloprotein inhibitors using a fragment-based approach. *J. Med. Chem.* **2010,** *54*, 591-602.

10. Fullagar, J. L.; Garner, A. L.; Struss, A. K.; Day, J. A.; Martin, D. P.; Yu, J.; Cai, X.; Janda, K. D.; Cohen, S. M., Antagonism of a zinc metalloprotease using a unique metal-chelating scaffold: tropolones as inhibitors of P. aeruginosa elastase. *Chem. Commun.* **2013,** *49*, 3197-3199.

11. Bryant, B. E.; Fernelius, W. C.; Douglas, B. E., Formation Constants of some Metal-Tropolone Complexes. *Nature* **1952,** *170*, 247.

12. Muetterties, E.; Roesky, H.; Wright, C., Chelate Chemistry. V. Metal Chelates Based on Tropolone and Its Derivatives. *J. Am. Chem. Soc.* **1966,** *88*, 4856-4861.

13. Brasen, W.; Holmquist, H.; Benson, R., N, N'-Disubstituted-1-amino-7-imino-1, 3, 5-cycloheptatrienes, a Non-classical Aromatic System. *J. Am. Chem. Soc.* **1961,** *83*, 3125-3135.

14. Muetterties, E.; Wright, C., Chelate chemistry. I. Tropolone and aminotroponimine derivatives of the main-group elements. *J. Am. Chem. Soc.* **1964,** *86*, 5132-5137.

15. Luoni, G.; McGuigan, C.; Andrei, G.; Snoeck, R.; De Clercq, E.; Balzarini, J., Bicyclic nucleoside inhibitors of Varicella–Zoster virus: The effect of branching in the p-alkylphenyl side chain. *Bioorg. Med. Chem. Lett.* **2005,** *15*, 3791-3796.

16. McGuigan, C.; Bidet, O.; Derudas, M.; Andrei, G.; Snoeck, R.; Balzarini, J., Alkenyl substituted bicyclic nucleoside analogues retain nanomolar potency against varicella zoster virus. *Bioorg. Med. Chem.* **2009,** *17*, 3025-3027.

17. Sniady, A.; Durham, A.; Morreale, M. S.; Marcinek, A.; Szafert, S.; Lis, T.; Brzezinska, K. R.; Iwasaki, T.; Ohshima, T.; Mashima, K., Zinc-catalyzed cycloisomerizations. Synthesis of substituted furans and furopyrimidine nucleosides. *J. Org. Chem.* **2008,** *73*, 5881-5889.

18. Meneni, S.; Ott, I.; Sergeant, C. D.; Sniady, A.; Gust, R.; Dembinski, R., 5-Alkynyl-2′-deoxyuridines: chromatography-free synthesis and cytotoxicity evaluation against human breast cancer cells. *Bioorg. Med. Chem.* **2007,** *15*, 3082-3088.

19. Mei, H.; Röhl, I.; Seela, F., Ag+-mediated DNA base pairing: extraordinarily stable pyrrolo-dC–pyrrolo-dC pairs binding two silver ions. *J. Org. Chem.* **2013,** *78*, 9457-9463.

20. Woo, J.; Meyer Jr, R. B.; Gamper, H. B., G/C-modified oligodeoxynucleotides with selective complementarity: synthesis and hybridization properties. *Nucleic Acids Res.* **1996,** *24*, 2470-2475.

21. Li, Y.; Soni, P. B.; Liu, L.; Zhang, X.; Liotta, D. C.; Lutz, S., Synthesis of fluorescent nucleoside analogs as probes for 2′-deoxyribonucleoside kinases. *Bioorg. Med. Chem. Lett.* **2010,** *20*, 841-843.

22. Seela, F.; Sirivolu, V. R., Pyrrolo-dC oligonucleotides bearing alkynyl side chains with terminal triple bonds: synthesis, base pairing and fluorescent dye conjugates prepared by the azide–alkyne “click” reaction. *Org. Biomol. Chem.* **2008,** *6*, 1674-1687.

23. Berry, D. A.; Jung, K.-Y.; Wise, D. S.; Sercel, A. D.; Pearson, W. H.; Mackie, H.; Randolph, J. B.; Somers, R. L., Pyrrolo-dC and pyrrolo-C: fluorescent analogs of cytidine

and 2′-deoxycytidine for the study of oligonucleotides. *Tetrahedron Lett.* **2004,** *45*, 2457-2461.

24. Liu, C.; Martin, C. T., Fluorescence characterization of the transcription bubble in elongation complexes of T7 RNA polymerase. *J. Mol. Biol.* **2001,** *308*, 465-475.

25. Liu, C.; Martin, C. T., Promoter Clearance by T7 RNA Polymerase initial bubble collapse and transcript dissociation monitored by base analog fluorescence. *J. Biol. Chem.* **2002,** *277*, 2725-2731.

26. Balachandra, C.; Sharma, N. K., Novel fluorophores: Syntheses and photophysical studies of boron-aminotroponimines. *Dyes Pigm.* **2017,** *137*, 532-538.

27. Sanz, P.; Mó, O.; Yáñez, M.; Elguero, J., Bonding in Tropolone, 2-Aminotropone, and Aminotroponimine: No Evidence of Resonance-Assisted Hydrogen-Bond Effects. *Chem. Eur. J.* **2008,** *14*, 4225-4232.

28. Seo, S.; Onizuka, K.; Nishioka, C.; Takahashi, E.; Tsuneda, S.; Abe, H.; Ito, Y., Phosphorylated 5-ethynyl-2′-deoxyuridine for advanced DNA labeling. *Org. Biomol. Chem.* **2015,** *13*, 4589-4595.

3B.7 Appendix

1. ^{1}H-/ ^{13}C-/ESI-MS/HRMS spectra of compound **4**

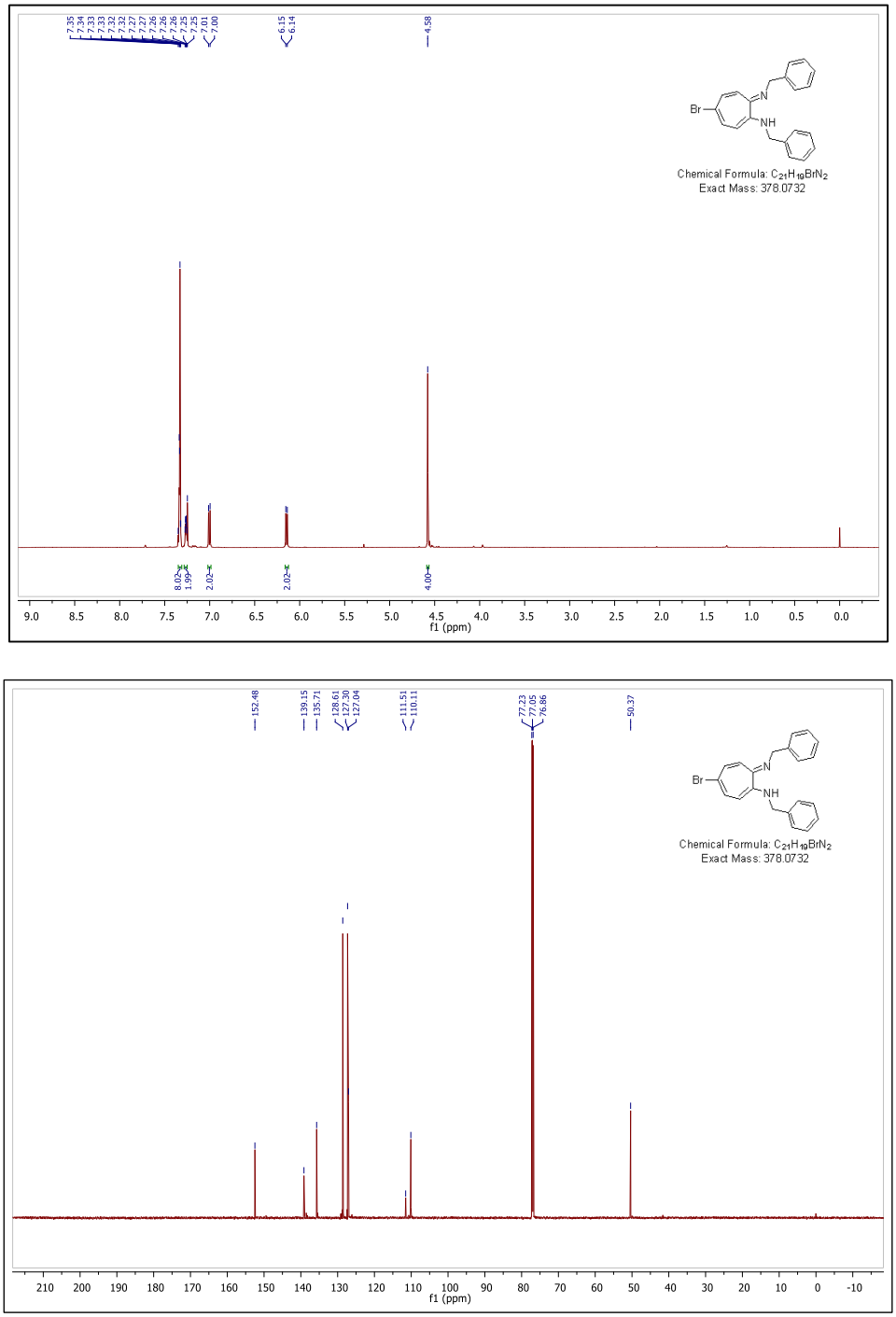


Figure S1. ^{1}H and ^{13}C NMR spectra of compound **4** in $CDCl_3$.

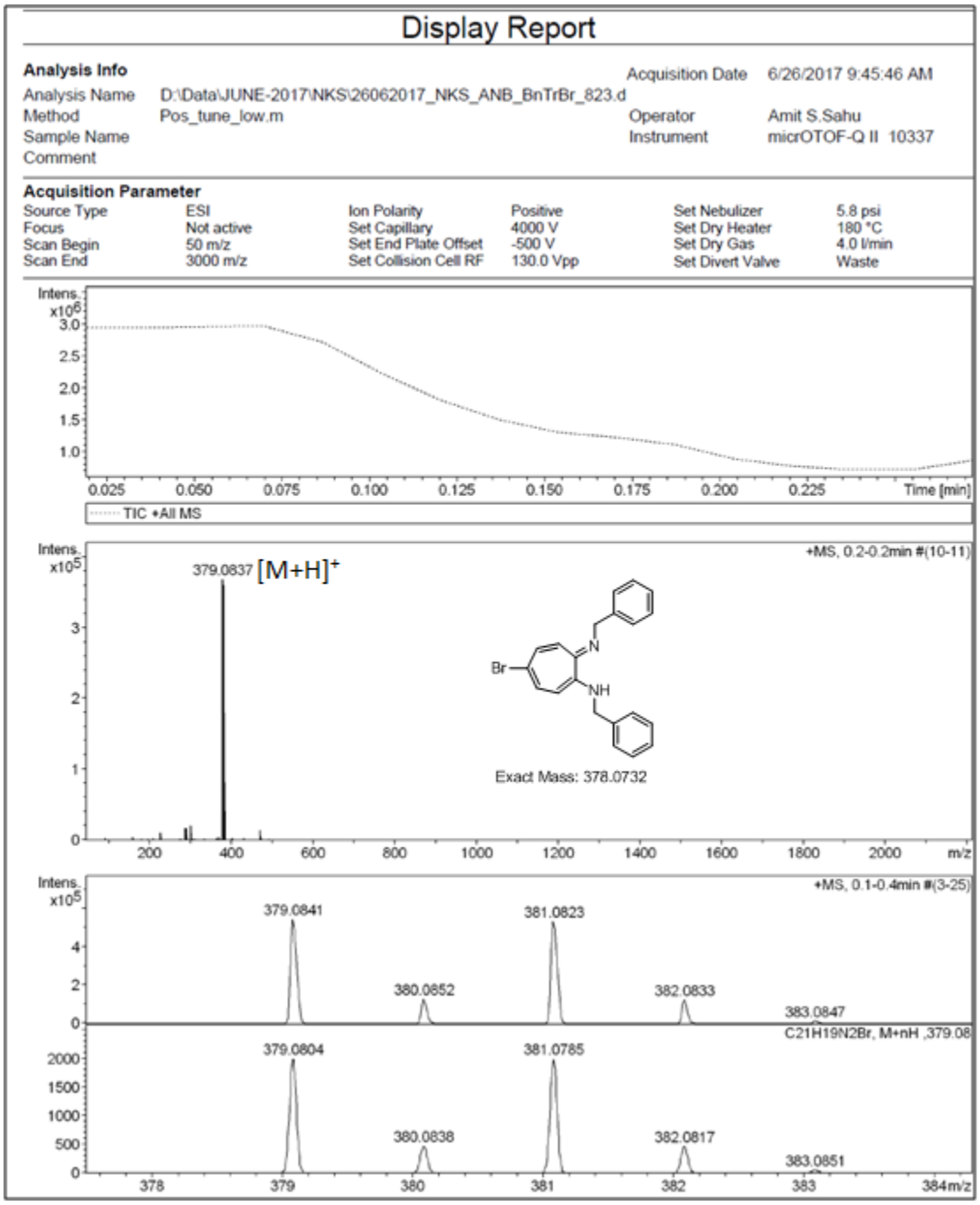


Figure S2. ESI-MS of compound **4**.

2.^{1}H-/ ^{13}C-/^{1}H-COSY/ESI-MS/HRMS spectra of compound ***tr*-7a**.

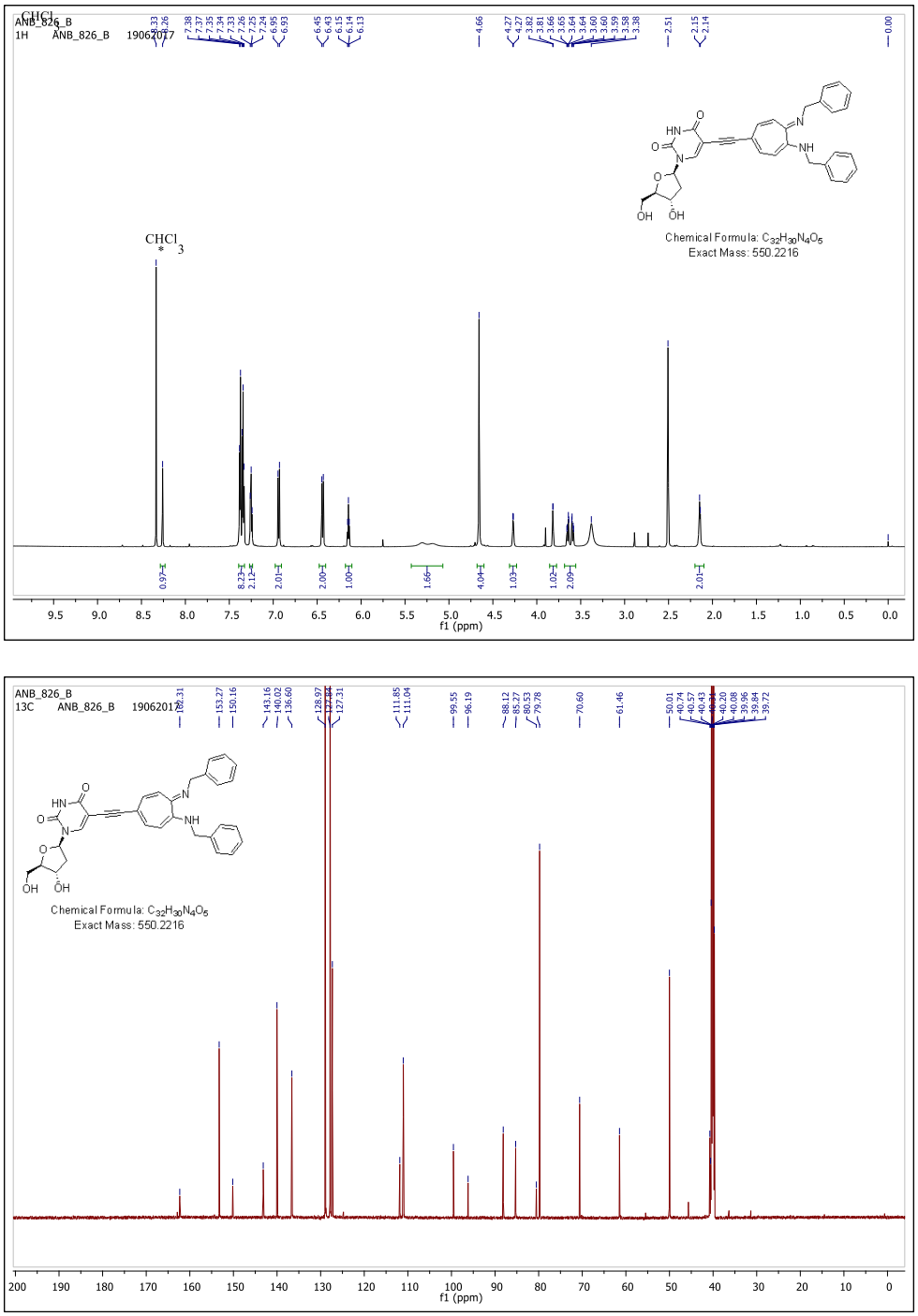


Figure S3. ^{1}H-/ ^{13}C spectra of compound ***tr*-7a** in DMSO-*d*$_6$.

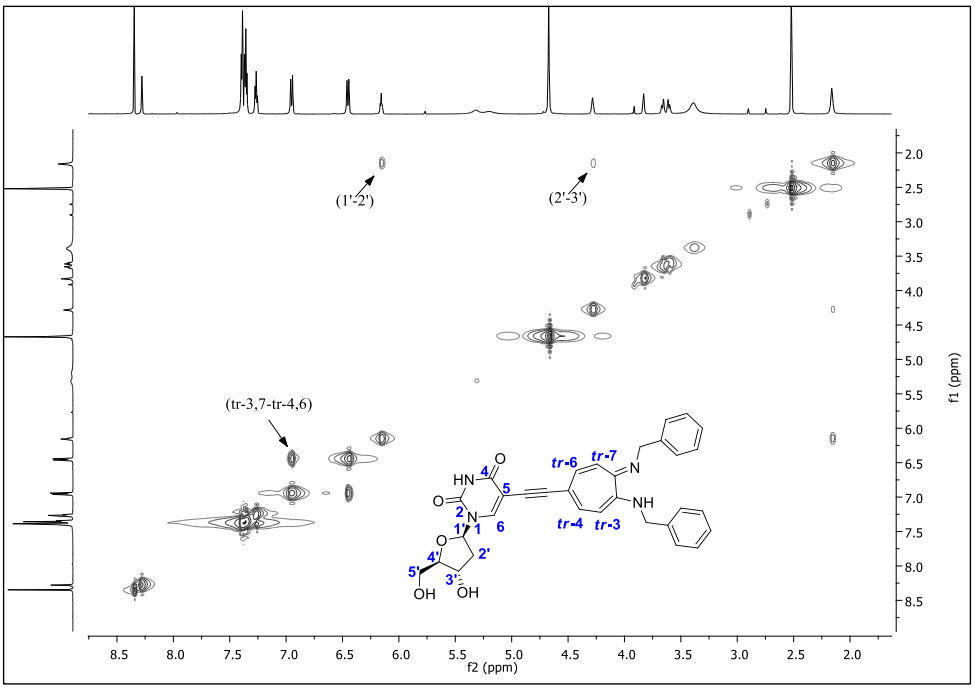


Figure S4. ^{1}H-COSY spectra of compound ***tr*-7a** in DMSO-d_6

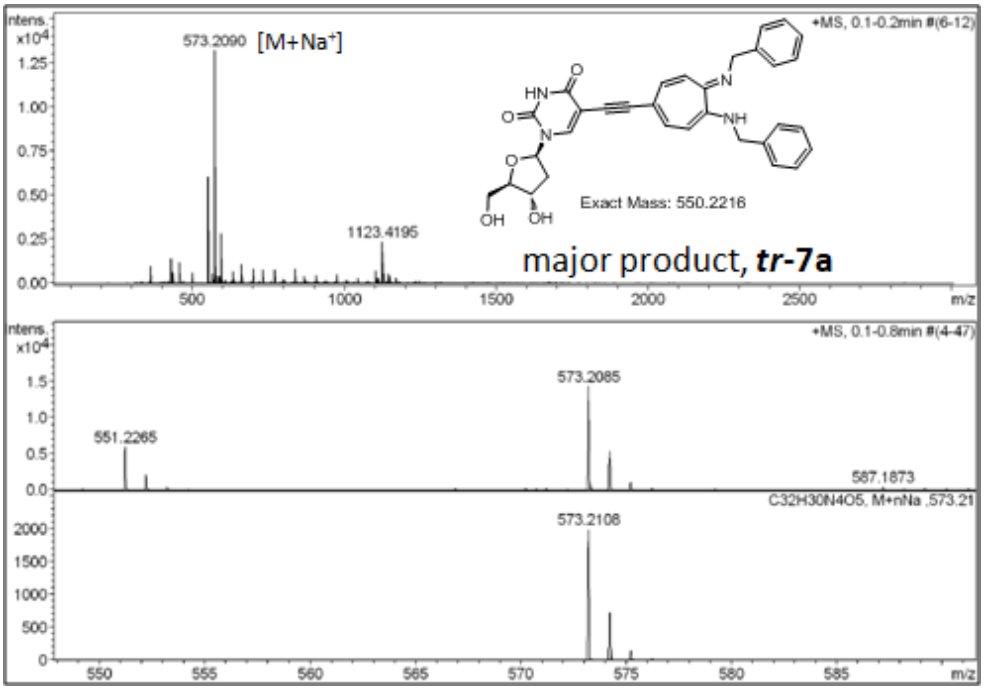


Figure S5. ESI-MS of compound ***tr*-7a.**

3.^{1}H-/ ^{13}C-/^{1}H-COSY/ESI-MS/HRMS spectra of compound ***tr*-7b**.

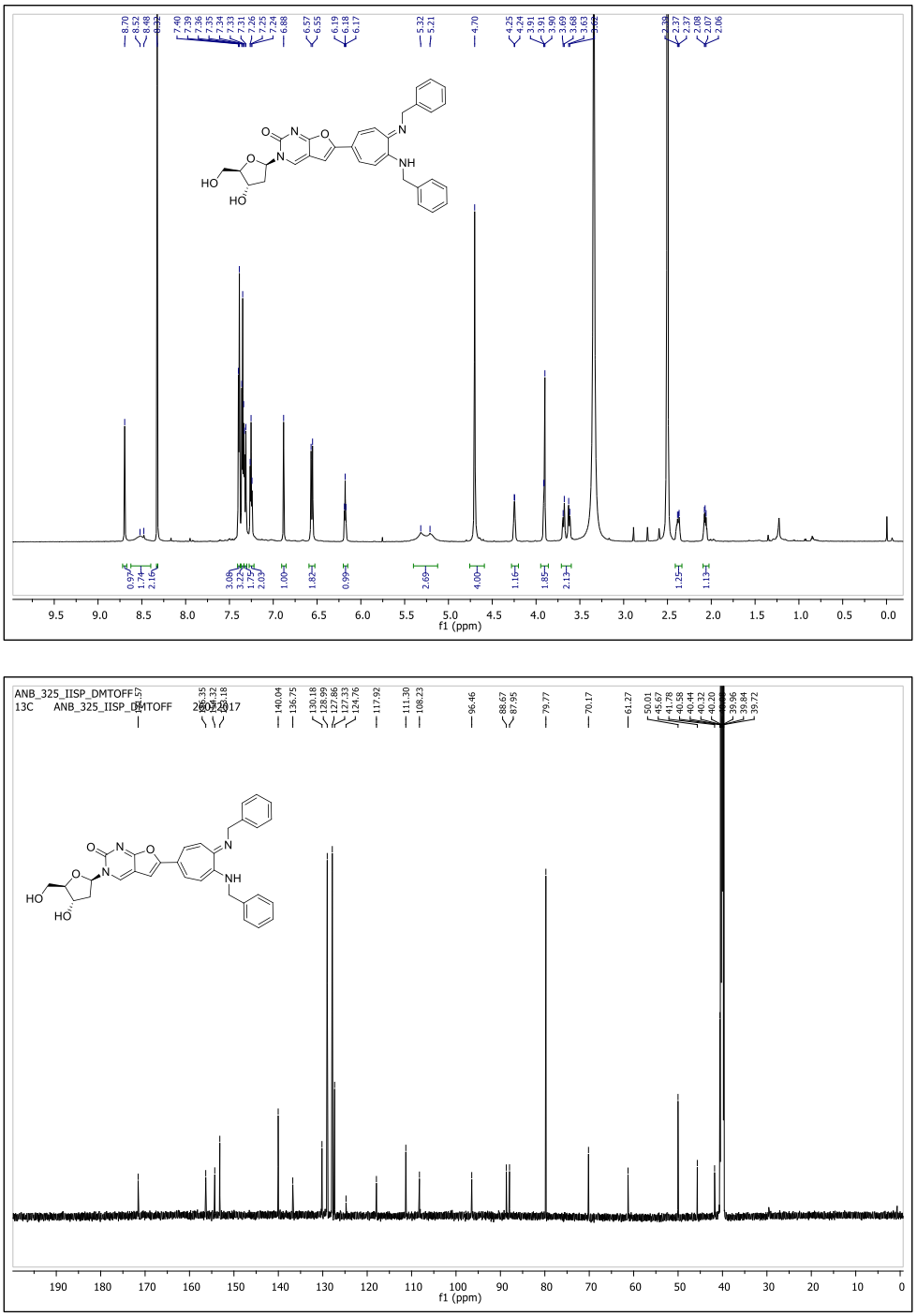


Figure S6. ^{1}H-/ ^{13}C spectra of compound ***tr*-7b** in DMSO-d_6.

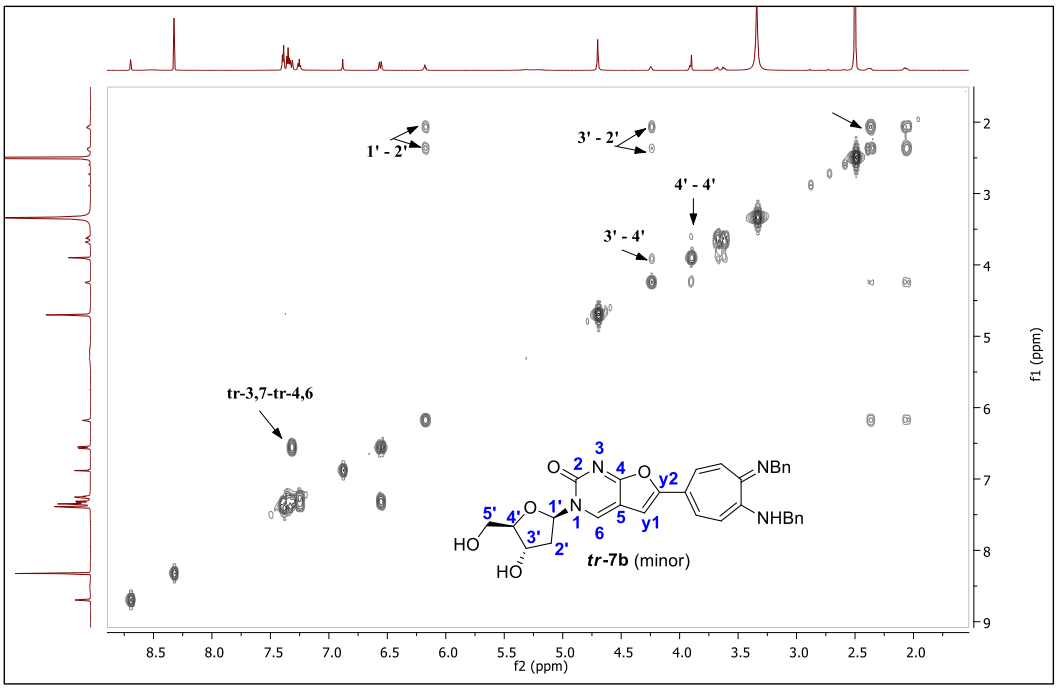


Figure S7. ^{1}H-COSY spectra of compound ***tr*-7b** in DMSO-d_6.

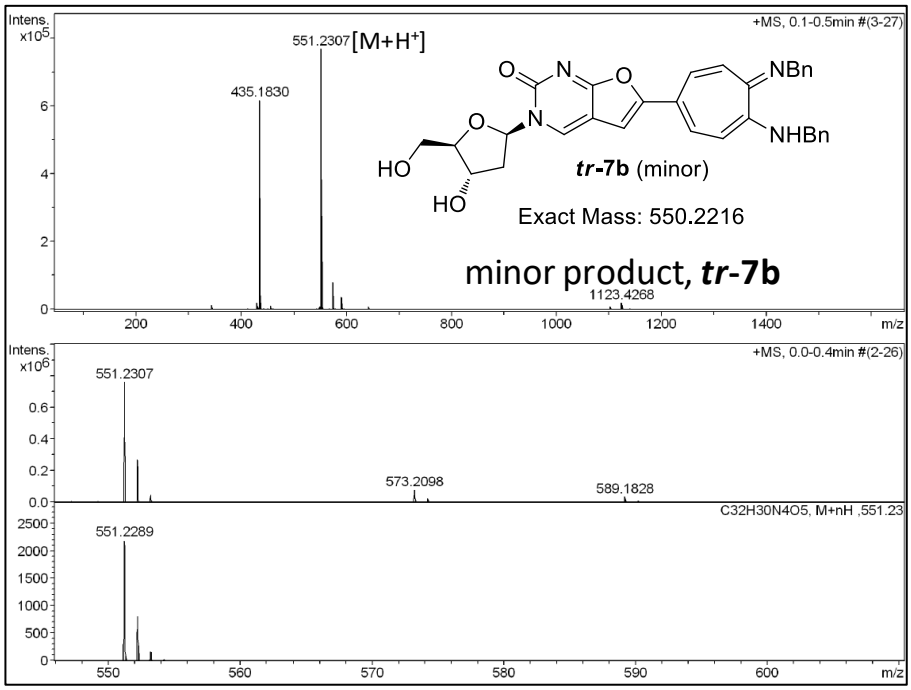


Figure S8. ESI-MS of compound ***tr*-7b**.

4.Stacked ^{1}H /^{13}C-NMR spectra of compounds ***tr*-7a** and ***tr*-7b.**

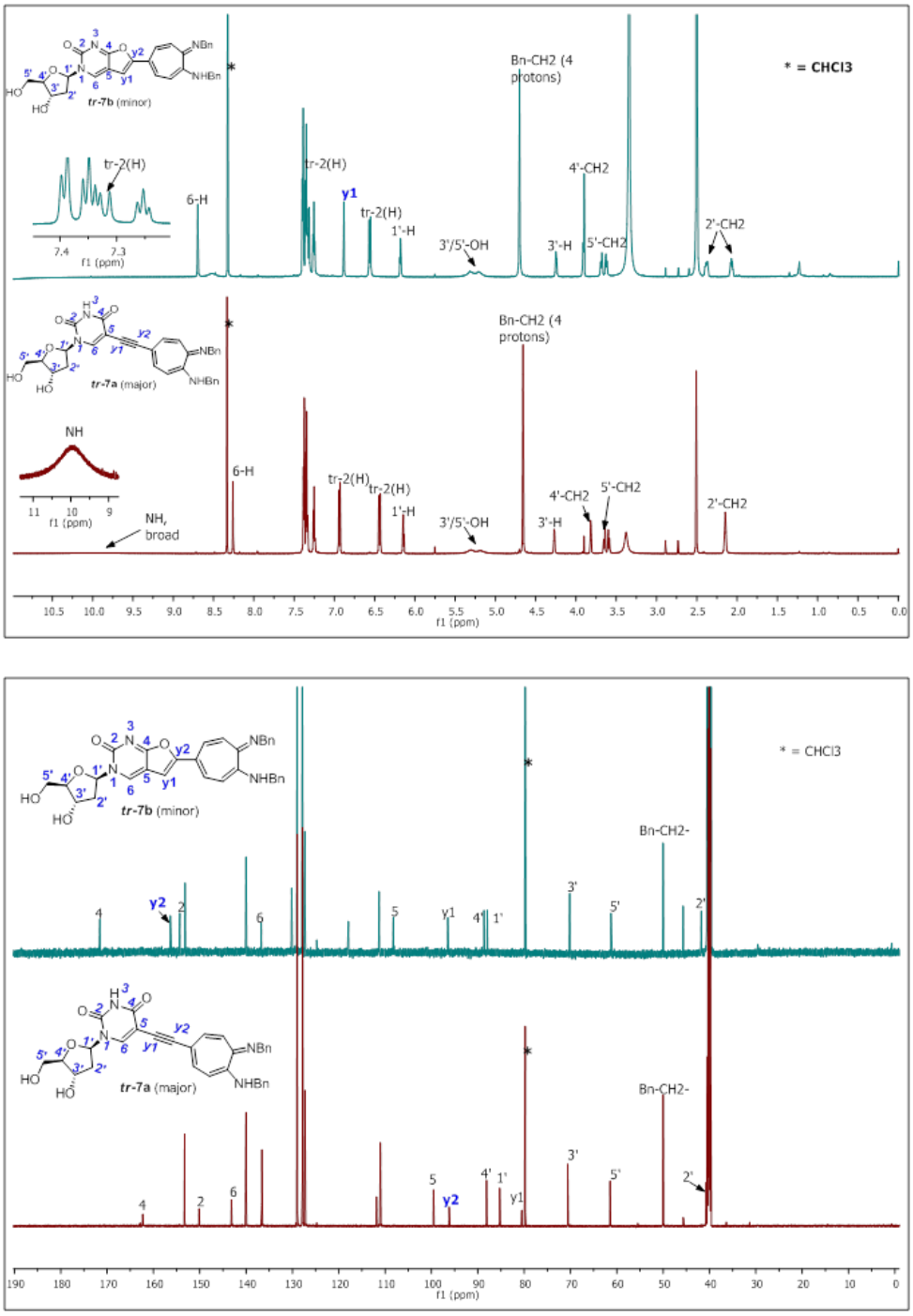


Figure S9. Stacked ^{1}H /^{13}C-NMR spectra of compound ***tr*-7a** and ***tr*-7b** for comparative understanding.

5.D_2O exchange ^1H NMR studies of compounds ***tr*-7a** and ***tr*-7b.**

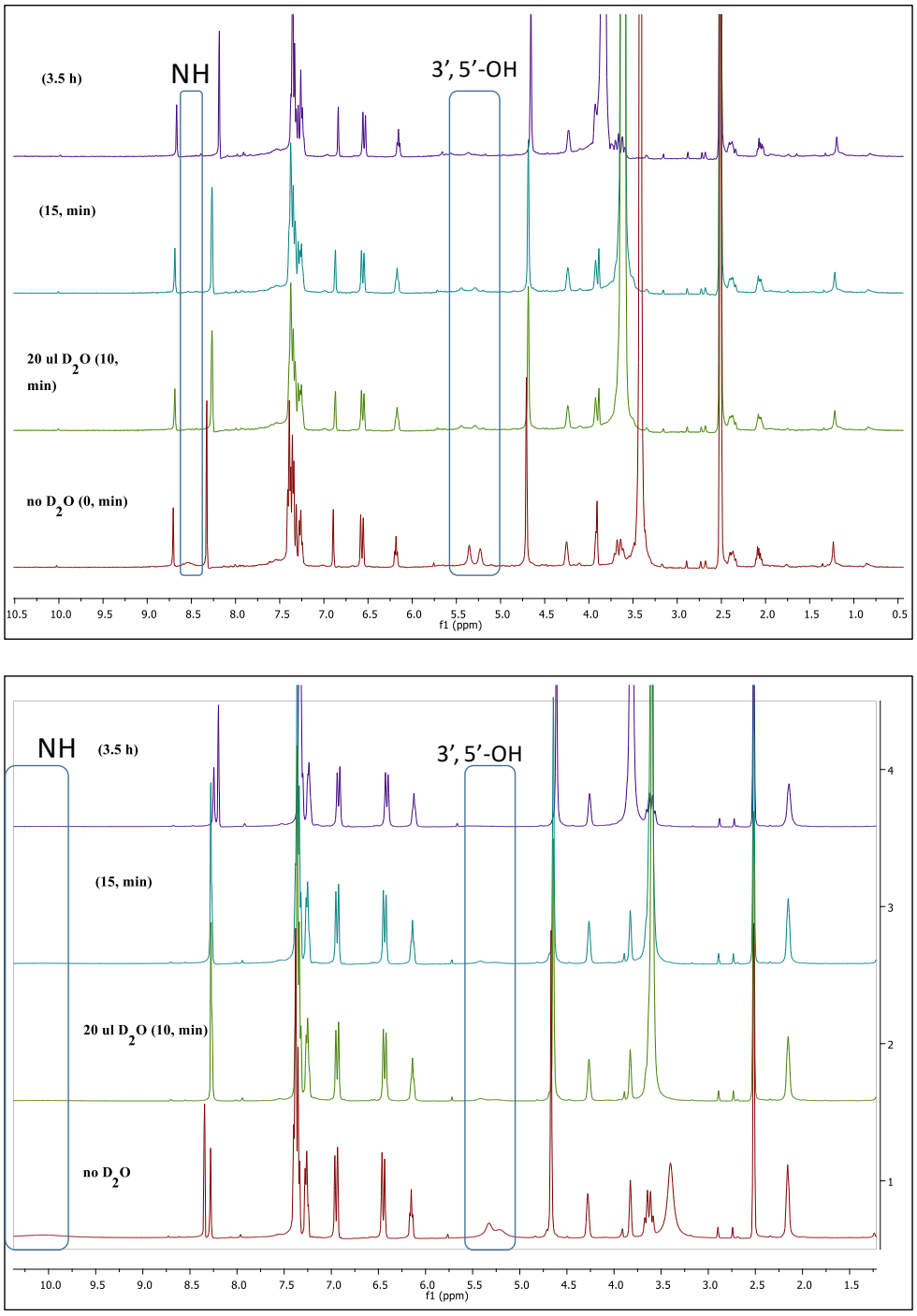


Figure S10. D_2O exchange ^1H NMR spectra of compounds ***tr*-7a** and ***tr*-7b**, exchangeable 3'-/5'-OH and NH before and after D_2O shake are framed.

6.^{1}H-/ ^{13}C-/ESI-MS/HRMS spectra of compound ***tr*-8a**.

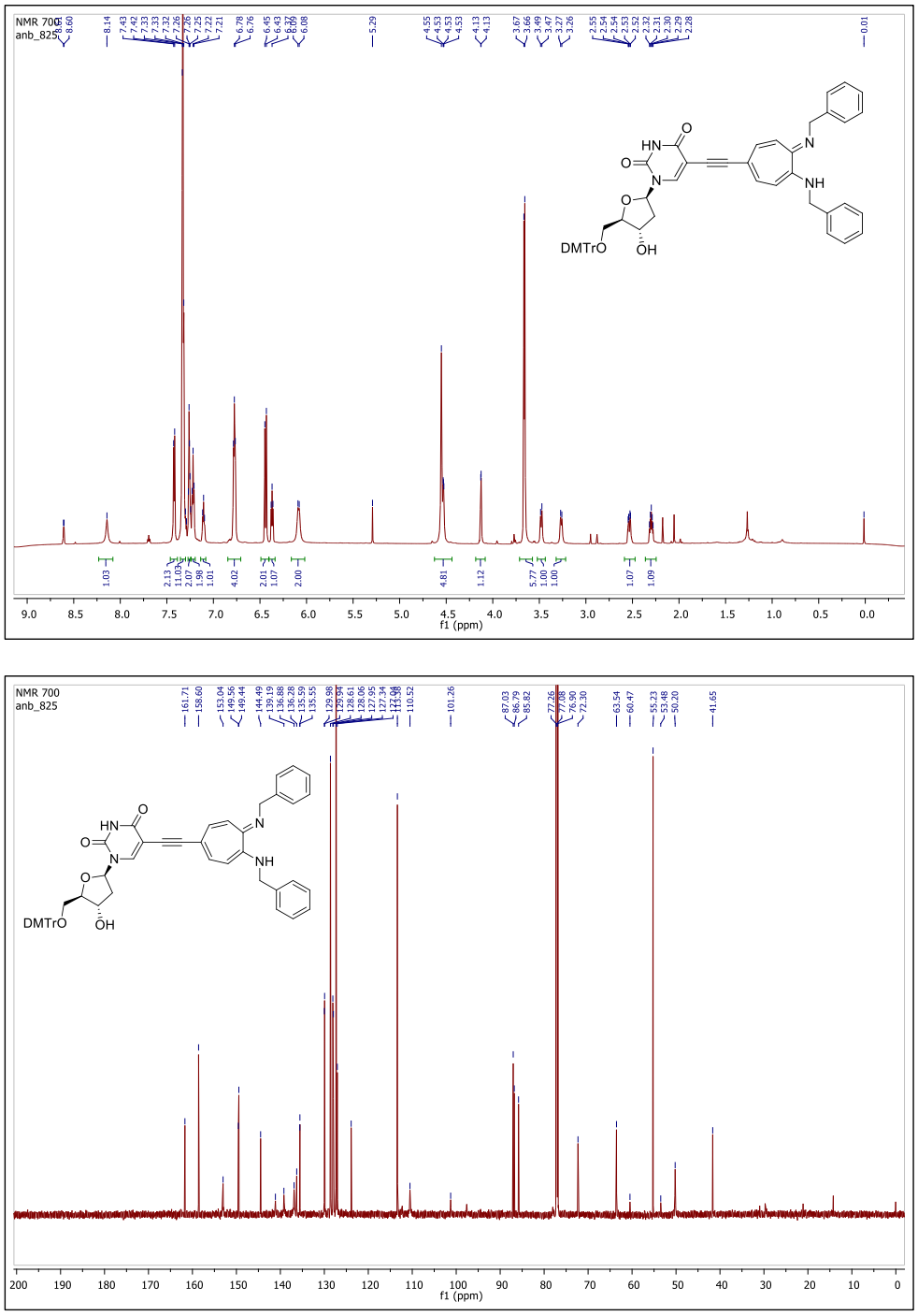


Figure S11. ^{1}H-/ ^{13}C spectra of compound ***tr*-8a** in $CDCl_3$.

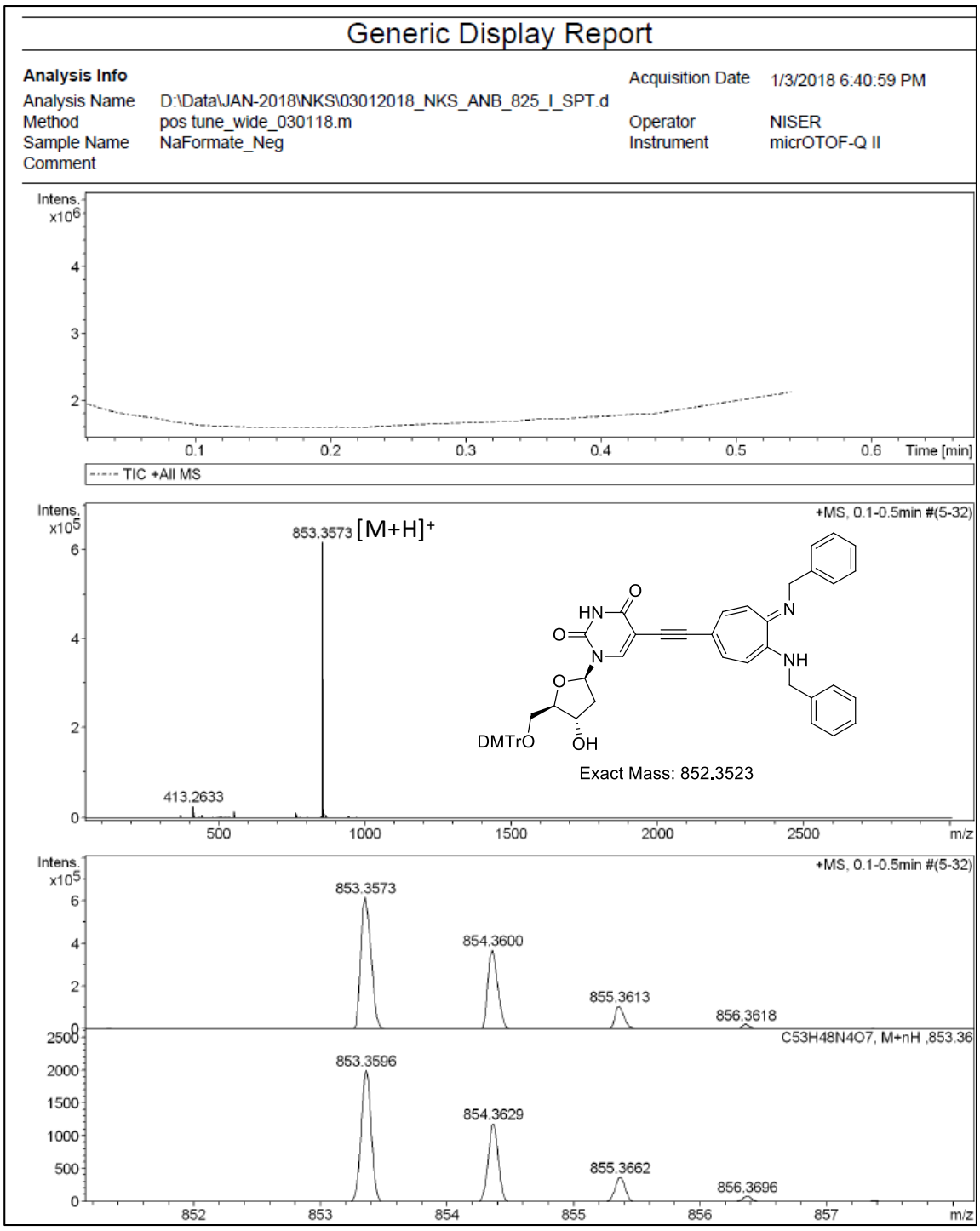


Figure S12. ESI-MS of compound **8a**.

7.^{1}H-/ ^{13}C-/ESI-MS/HRMS spectra of compound ***tr*-8b**.

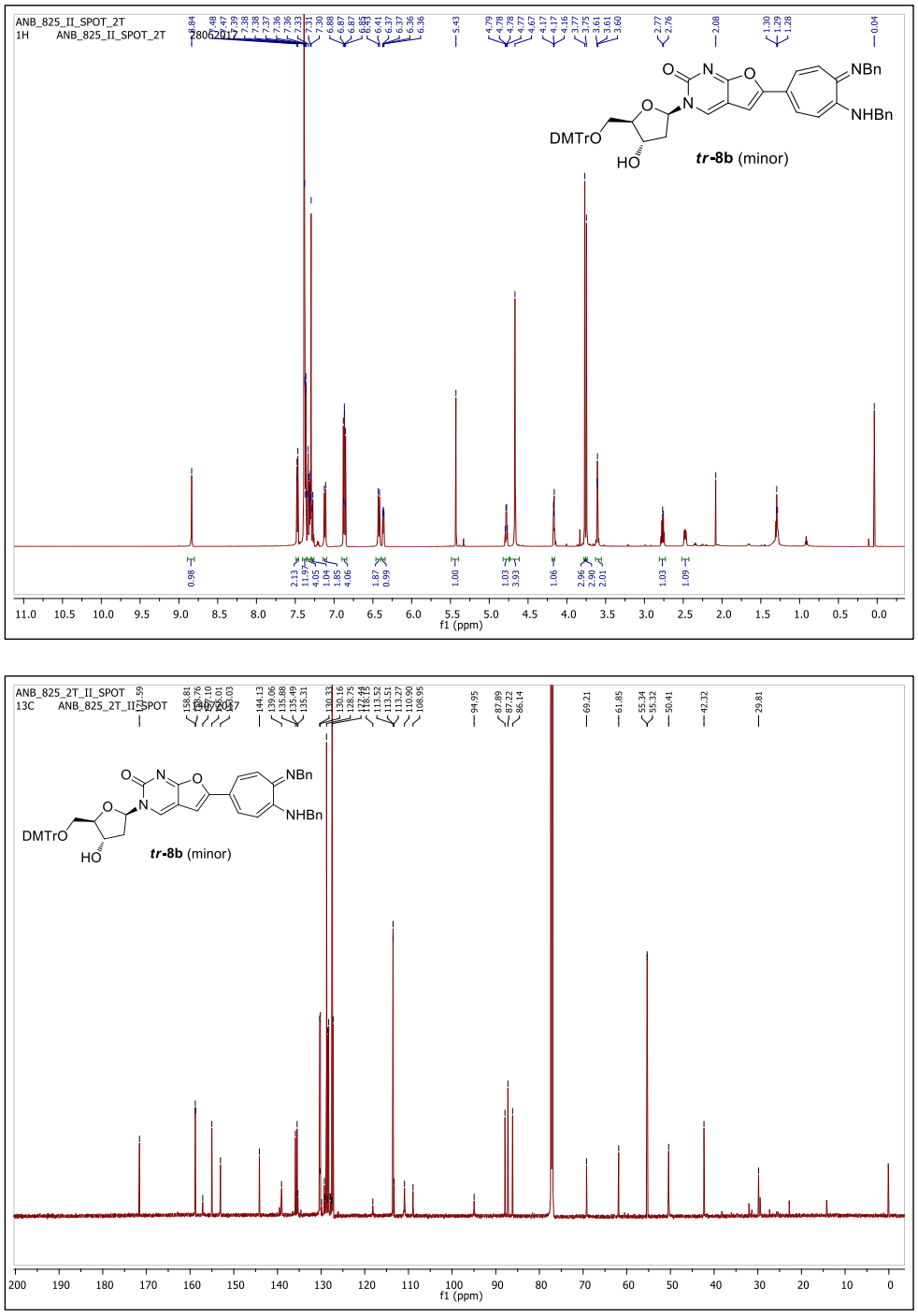


Figure S13. ^{1}H-/ ^{13}C spectra of compound ***tr*-8b** in $CDCl_3$.

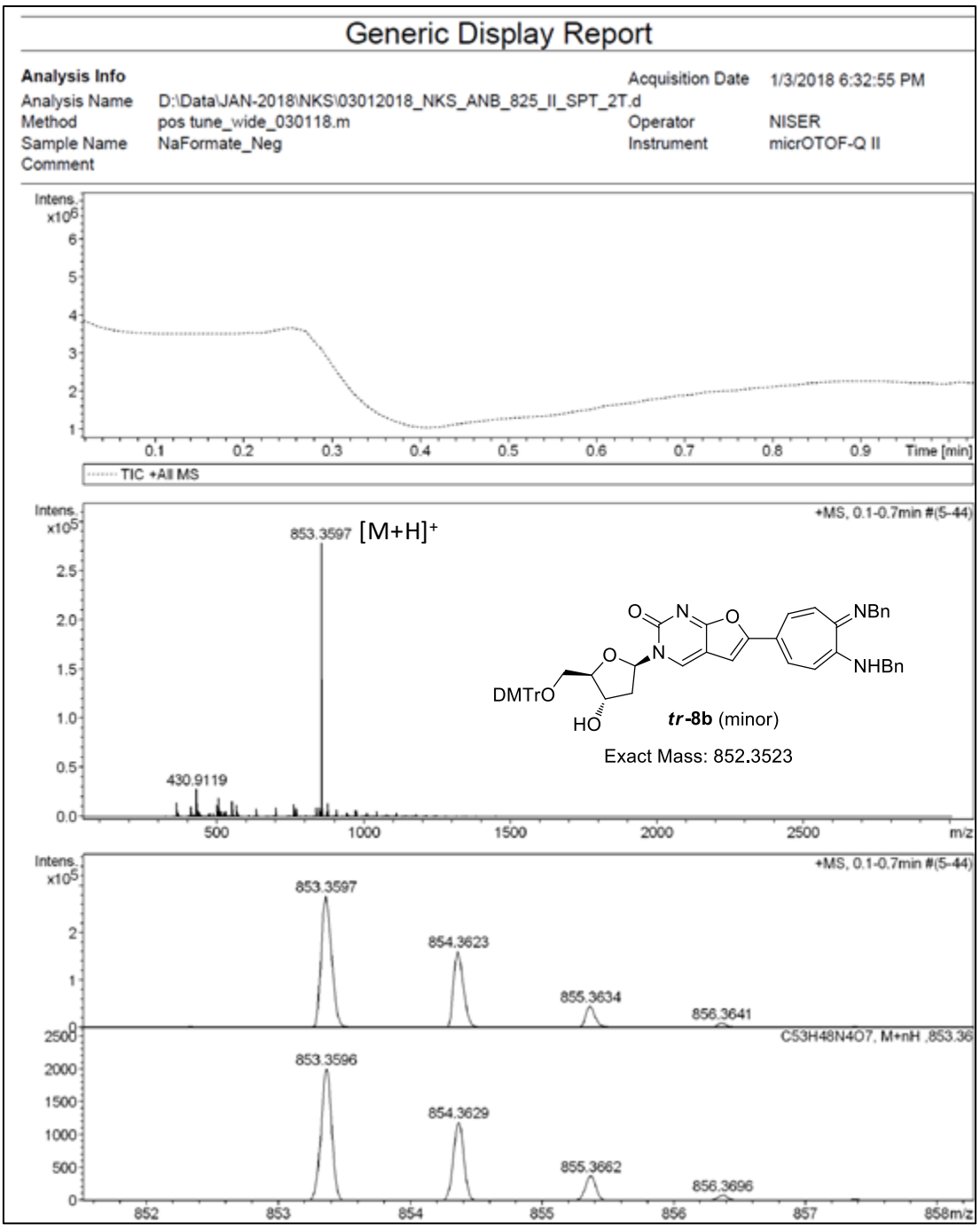


Figure S14. ESI-MS of compound **8b**.

8.^{1}H-/^{31}P-/ESI-MS/HRMS spectra of compound ***tr*-9a** phosphoramidite.

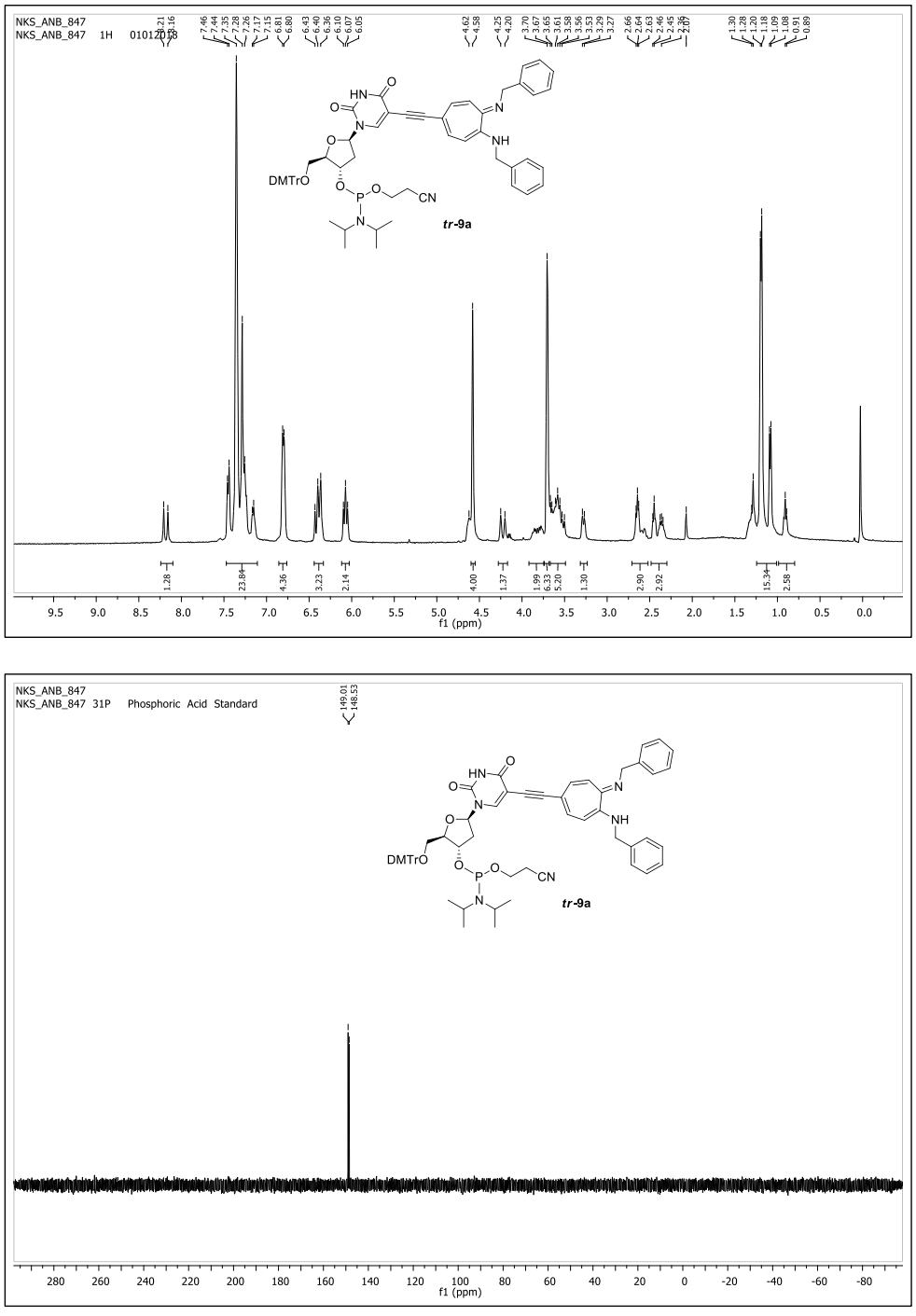


Figure S15. ^{1}H and ^{31}P NMR spectra of compound ***tr*-9a** phosphoramidite in $CDCl_3$.

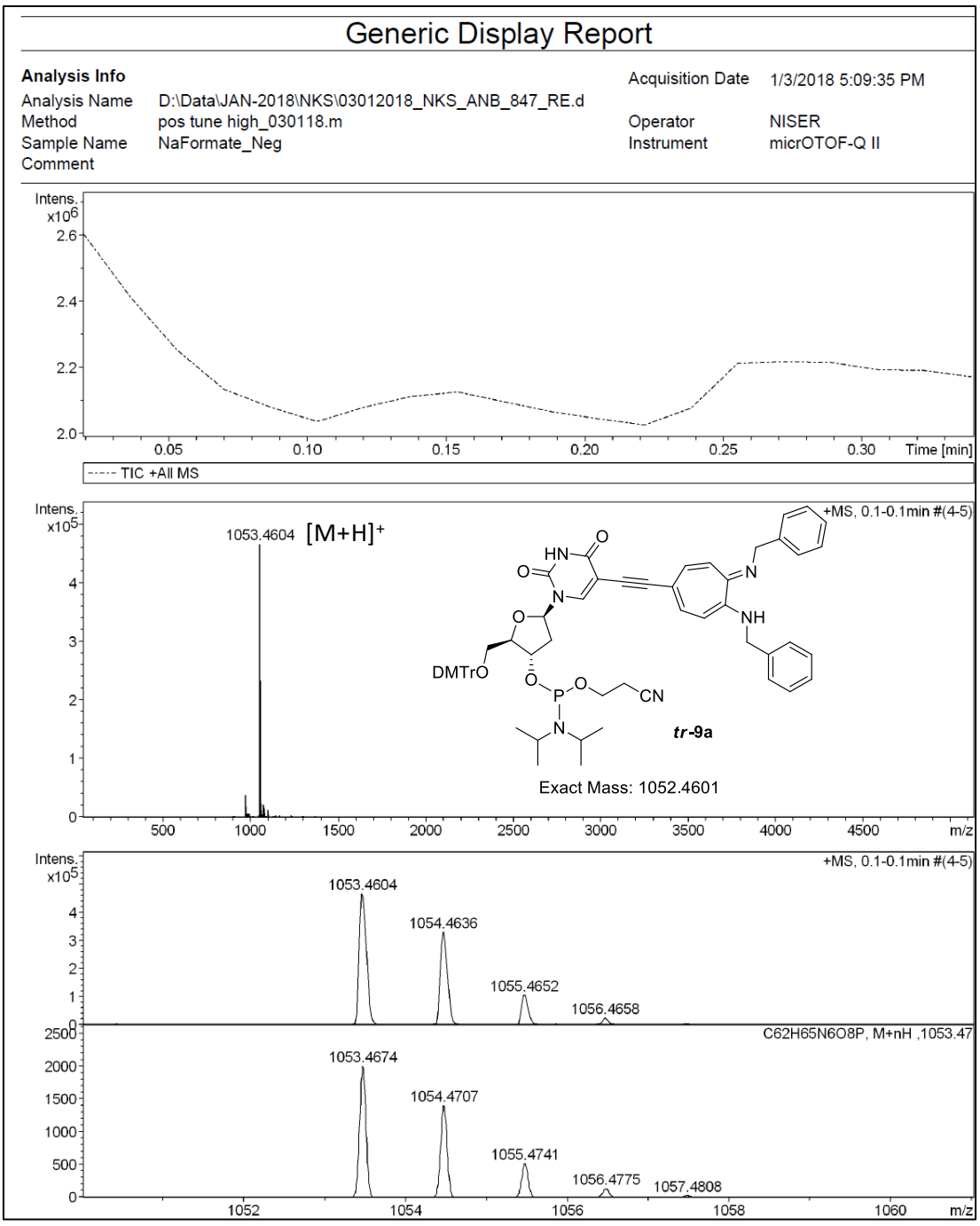


Figure S16. ESI-HRMS spectra of compound ***tr*-9a** phosphoramidites

9.^{1}H-/^{31}P-/ESI-MS/HRMS spectra of compound ***tr*-9b** phosphoramidite.

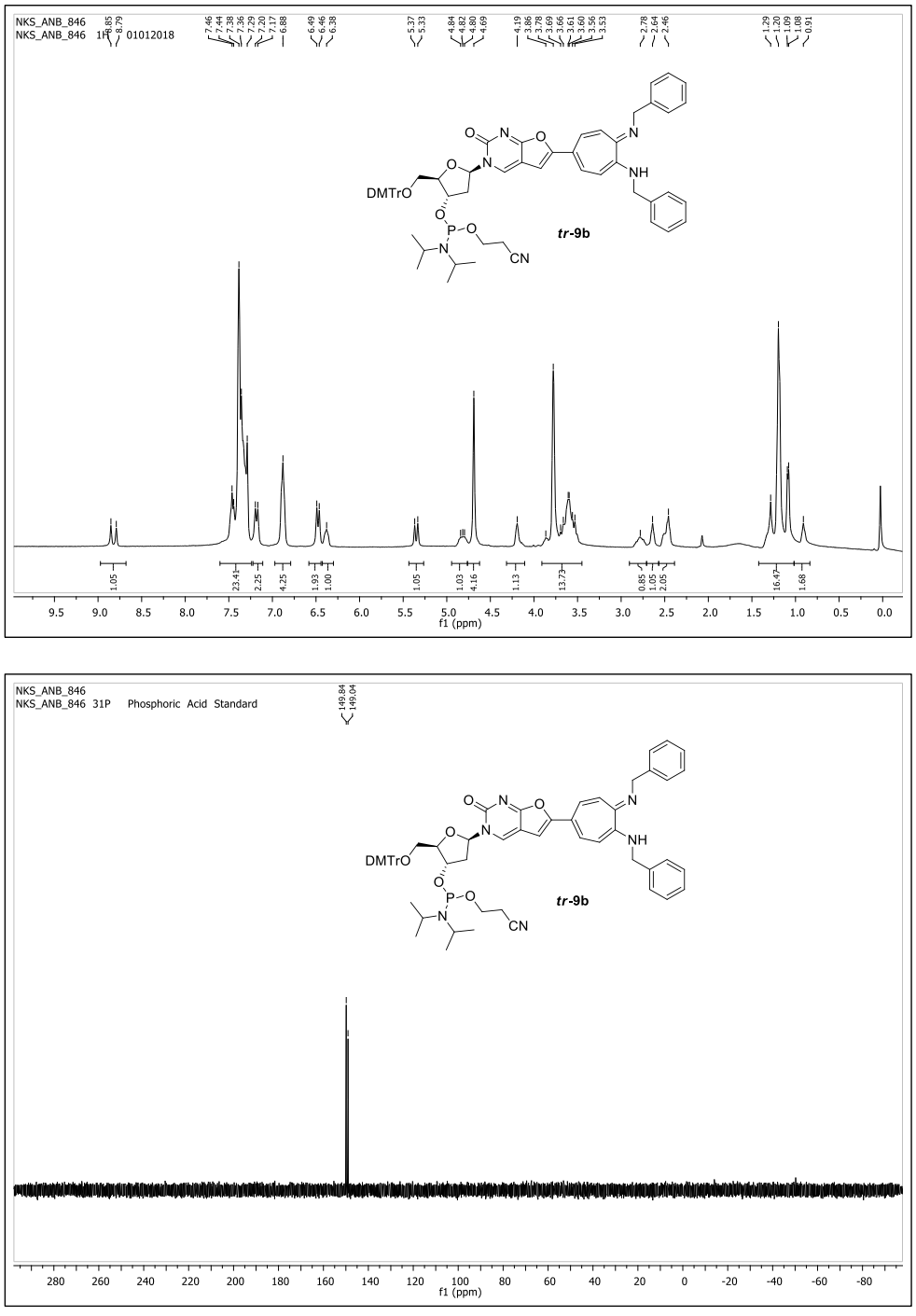


Figure S17. ^{1}H and ^{31}P NMR spectra of compound ***tr*-9b** phosphoramidite in $CDCl_3$.

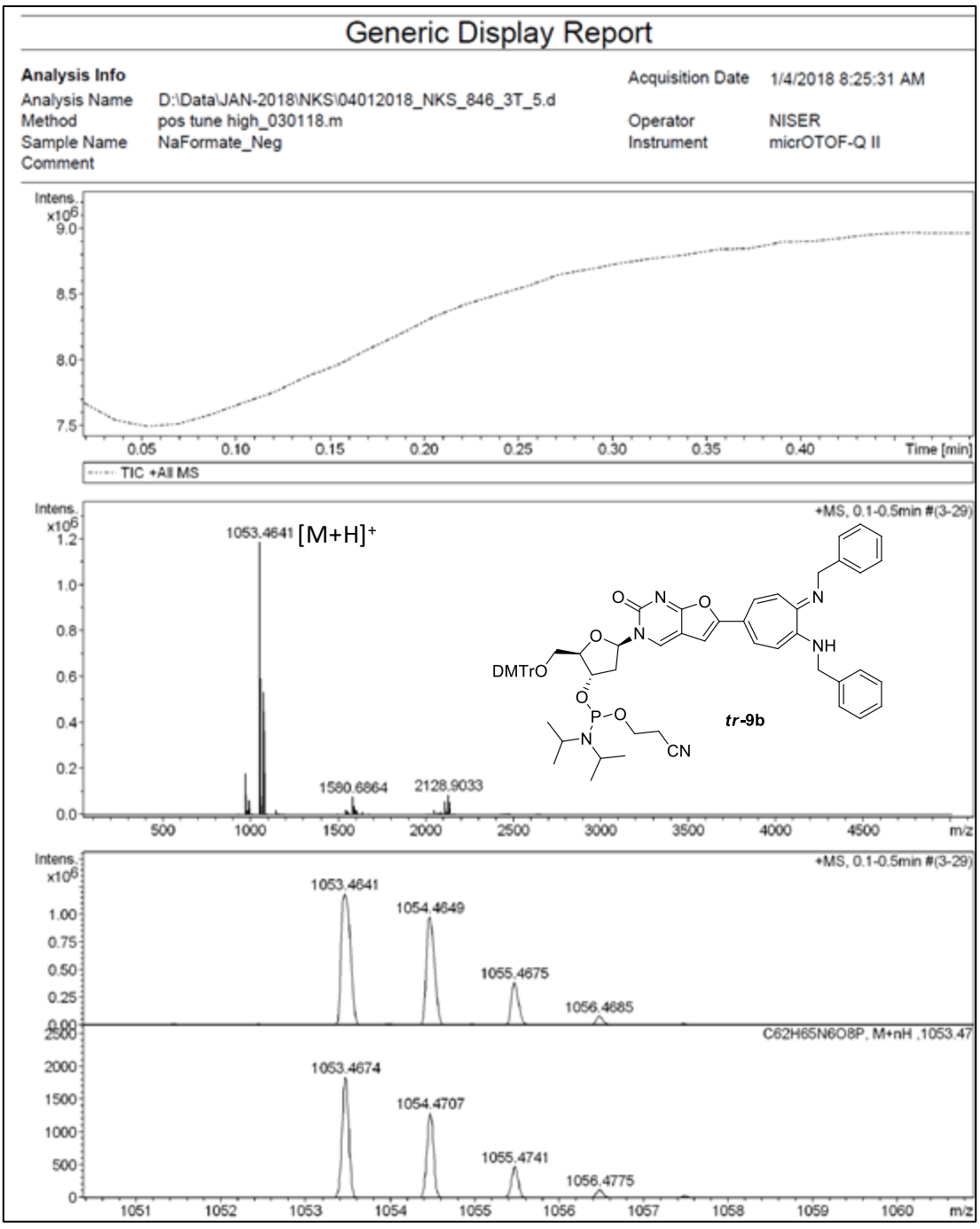


Figure S18. ESI-HRMS spectra of compound ***tr*-9b** phosphoramidite

10.ESI-MS of ALK/FUR DNA oligonucleotides.

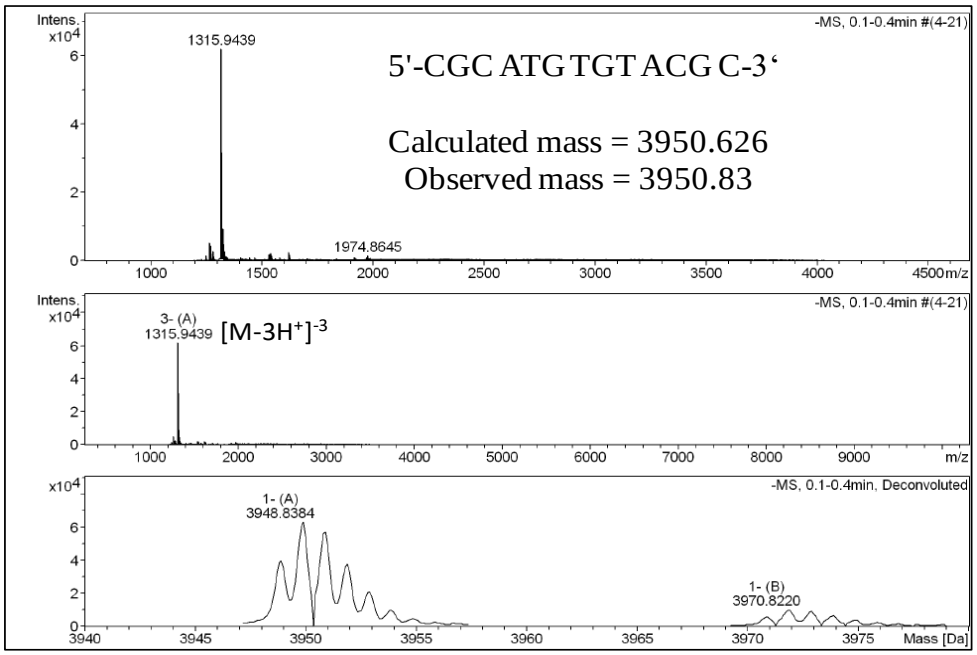


Figure S19. ESI-MS spectra of C-ON1 DNA oligonucleotide.

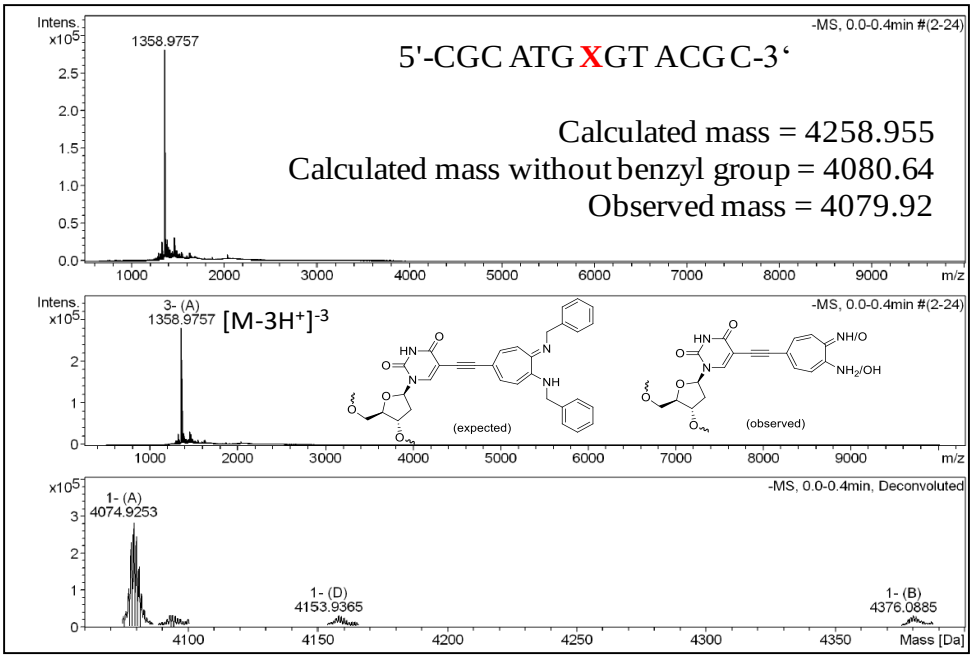


Figure 20. ESI-MS spectra of Alk-ON2 DNA oligonucleotide.

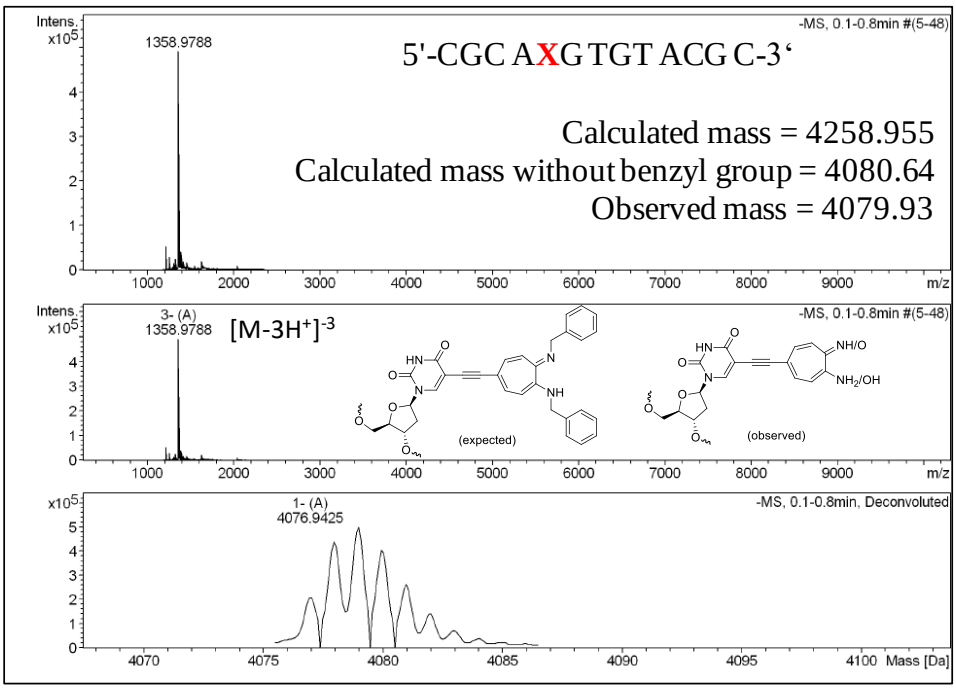


Figure 21. ESI-Mass spectra of ALK-ON3 DNA oligonucleotide.

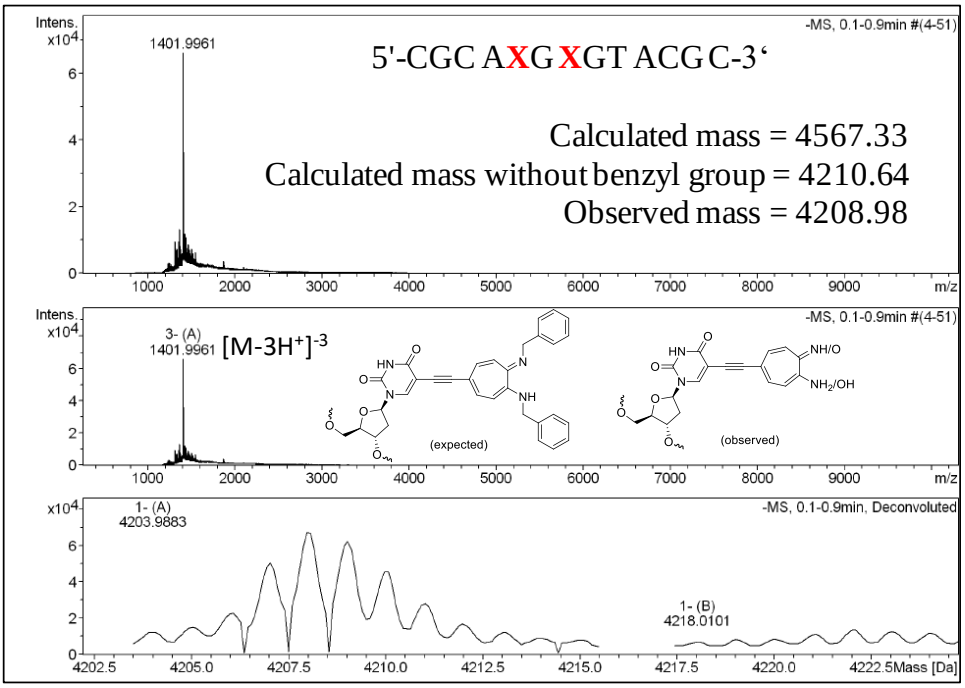


Figure 22. ESI-Mass spectra of **ALK-ON4** DNA oligonucleotide.

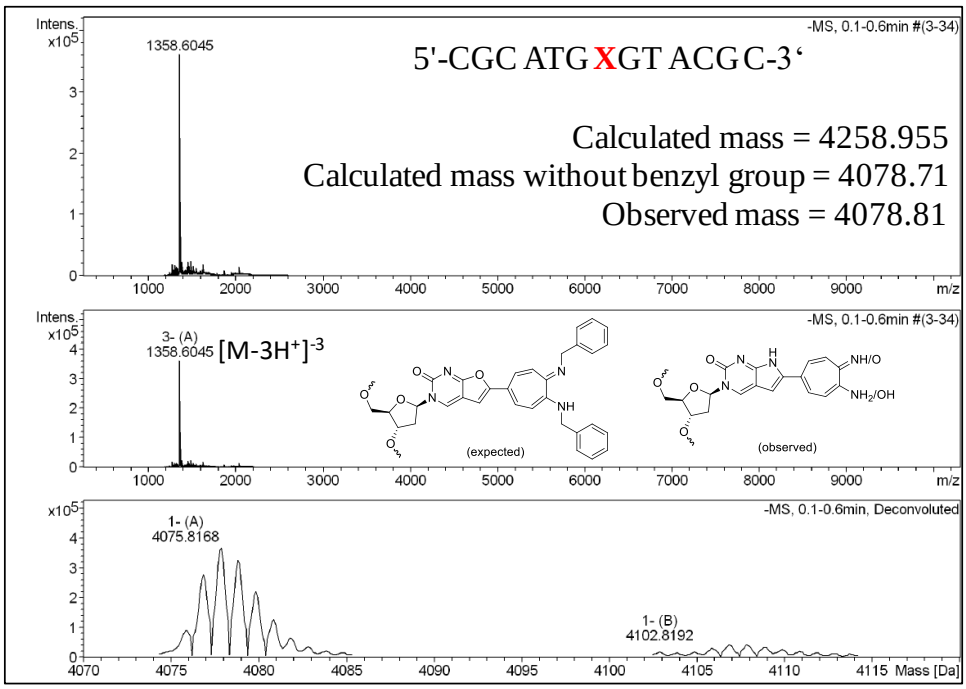


Figure 23. ESI-Mass spectra of **FUR-ON1** DNA oligonucleotide.

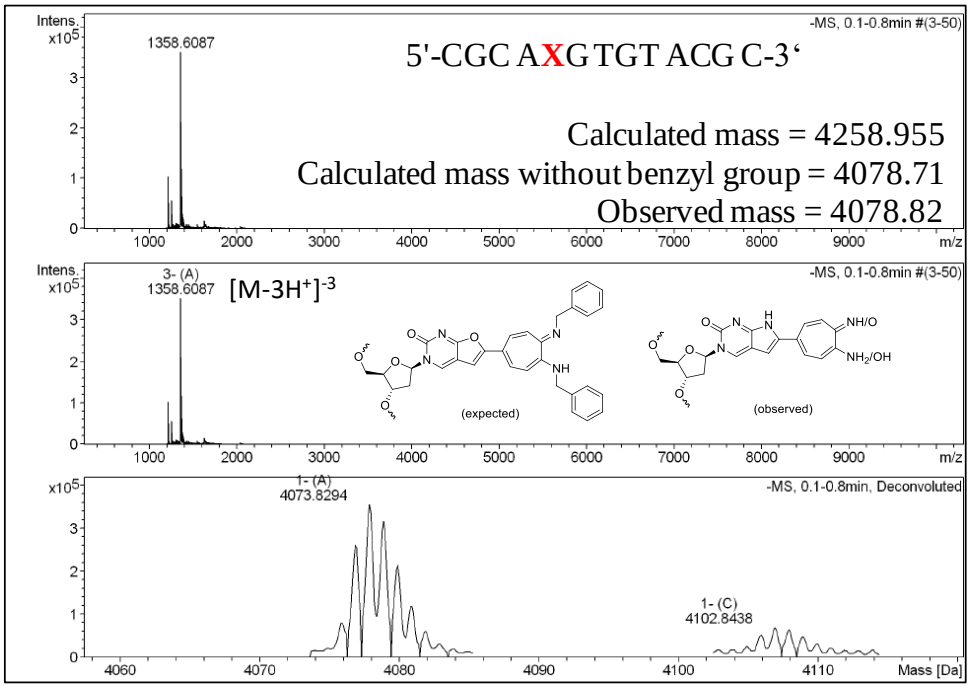


Figure 24. ESI-Mass spectra of **FUR-ON2** DNA oligonucleotide.

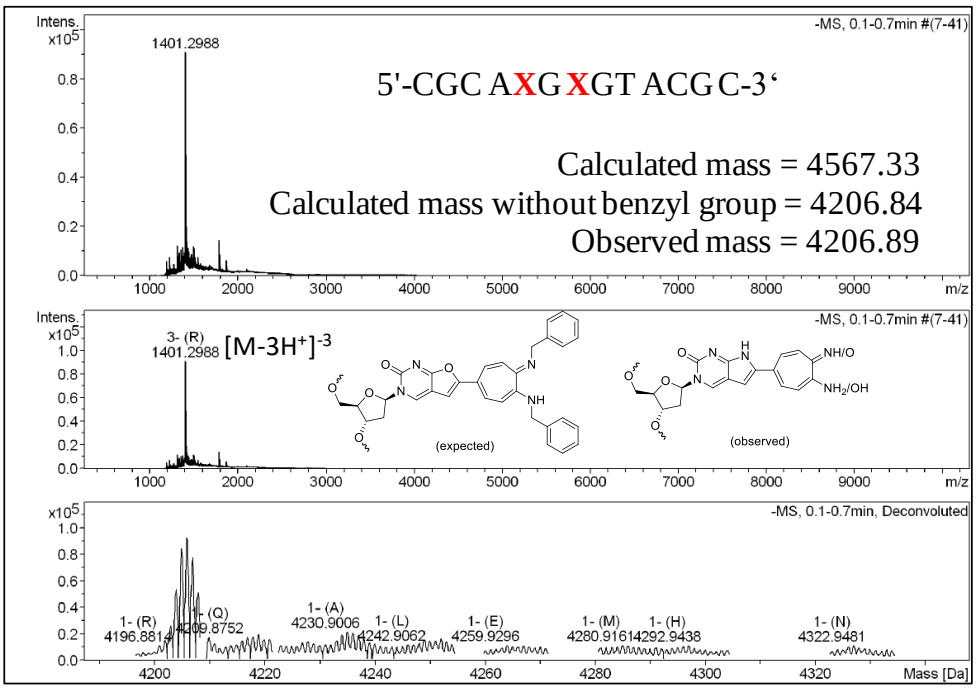


Figure 25. ESI-Mass spectra of FUR-ON3 DNA oligonucleotide.

CHAPTER 4

Part-A

Synthesis and Conformational Analysis of *Aminopyrazolonyl* Amino Acids (APA)/Peptides

Chapter 4A. Synthesis and Conformational Analysis of *Aminopyrazolonyl* Amino Acids (APA)/Peptides

4A.1 Introduction

Pyrazole, pyrazolone and aminopyrazolone are constituents of several natural products, synthetic drug molecules, agrochemical reagents, and metal chelating agents (Figure 4A.1).[1-3] For example, pyrazolone derivatives are well-known antipyretic and analgesic drugs for more than one century.[4, 5] Many other pyrazolone derivatives are reportedly known as neuroprotective agents, HIV integrase inhibitors, phosphodiesterase inhibitors, and antibacterial agents.[6] Further the substituted pyrazolone derivative, 4-aminophenazone (Aminopyrene), and its metabolite (Ampyrone) have shown excellent analgesic, anti-inflammatory and antipyretic activities, though some risk factors as agranulocytosis are also associated [7-10] (Figure 4A.1A). Substituted pyrazole and pyrazolone derivatives have shown strong DNA binding affinities and considered as potential anti-cancer agents.[11-13] For an example, the amino substituted pyrazolone molecule, ampyrone, has been modified as Schiff base derivatives and employed for metal complexation with metal ions (Co^{2+}/Ni^{2+}/Fe^{3+}) which are considered as antimicrobial agents.[14]

Further, ampyrone is explored as a biochemical reagent for determination of phenol concentration including inside the cell.[15-18] Moreover, pyrazole and pyrazolone derivatives have abilities to form hydrogen bonding and have been explored in screening of their biological activities.[19, 20] Thus pyrazole and pyrazolone could be considered as potential scaffold for syntheses of unnatural aromatic amino acids for tuning the peptide structures. Schrader and co-workers have reported that aminopyrazole containing amino acids are the artificial templates for stabilization of β-sheet

conformation [21-23] (Figure 4A.1A). Recently, Sutherland and co-workers have reported the new class of conjugated unnatural α-amino acid containing 5-arylpyrazole residue as probe of serine proteases[24] (Figure 4A.1A).

4A.2 Objective of our work

Herein we report the syntheses of rationally designed aminopyrazolonyl amino acid (APA) derivatives (Figure 4A.1B) and their roles in structural organization of native dipeptides.

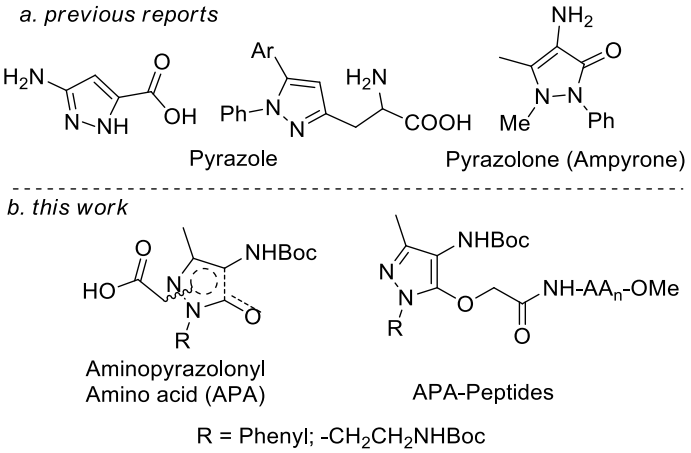
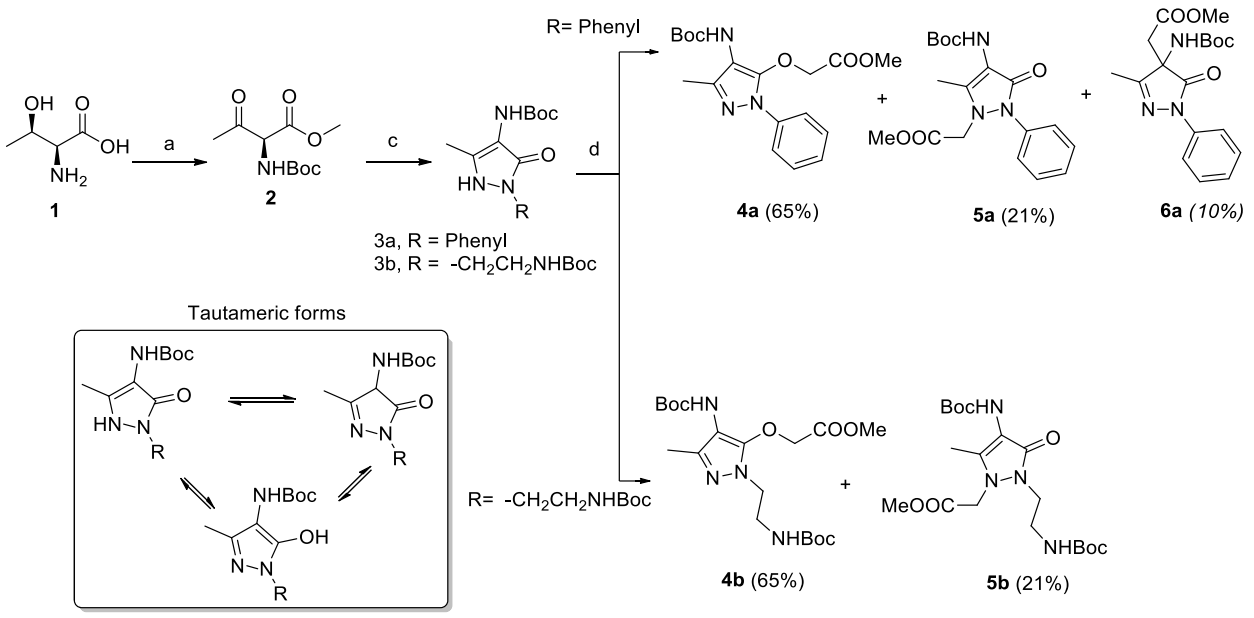


Figure 4A.1. (a) Chemical structures of pyrazole/pyrazolone/aminopyrazolone; (b) rationally designed amino acids/peptides.

4A.3 Results and Discussion

The synthesis of APA peptides was commenced from L-threonine amino acid as shown in Scheme 4A.1. The amine and acid groups of threonine was protected as NBoc and methyl esters respectively. The hydroxyl group of threonine derivative **1** is oxidized to β-keto ester **2** with Dess–Martin oxidizing reagent. The aminopyrazolone derivative (**3a**/**3b**) was prepared by refluxing β-keto ester (**2**) with corresponding phenylhydrazine/(NHBoc)ethyl-2-hydrazine in presence of acetic acid. The (NHBoc)ethyl-2-hydrazine was prepared by following literature report from 2-

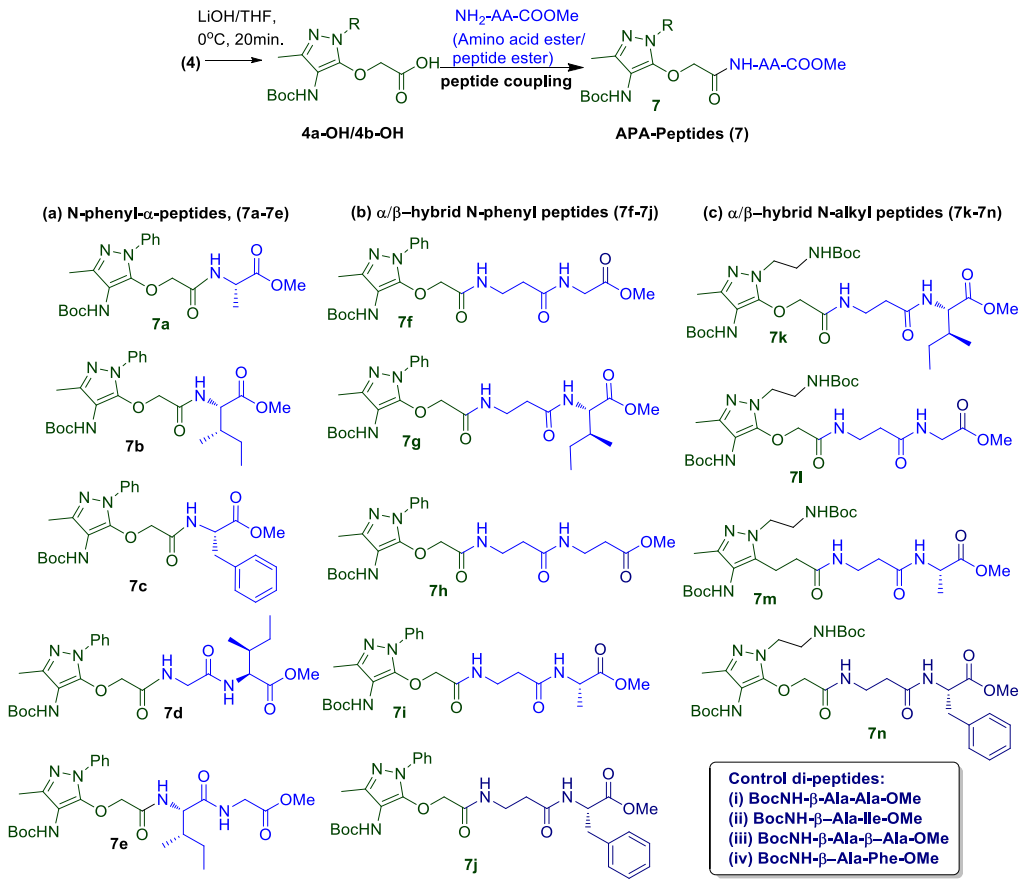
bromoethylamine in two steps. The alkylation of aminopyrazolone derivatives (**3a**) with methyl bromoacetate under basic conditions resulted in isolation of corresponding O/N/C alkylated products, **4a**/**5a**/**6** in 81% yields with regioselectivity of 69:21:10 (**4a**/**5a**/**6**). However, the alkylation of 3b resulted in isolation of only O/N alkylated products **4b**/**5b**, we couldn't observe C-alkylated product with **3b**. The structure assignment of O/N alkylated products of **4b**/**5b** was determined from NMR data by comparing with **4a**/**5a**/**6**.



Scheme 4A.1. Synthesis of 4-aminopyrazolone amino acids. Reagents and reaction conditions: (a) (i) SOCl2, MeOH, 0 °C then reflux, 12 h, (ii) CH_3CN:H_2O (85:15), Et_3N, $(Boc)_2O$, 0 °C-rt, 13 h, 85 % (iii) Dess–Martin periodinane, DCM, 2.5 h, rt, 85 % (b) Phenyl hydrazine, CH_3COOH, MeOH, reflux, 16 h, 86 %; (c) methylbromo acetate ($BrCH_2COOMe$), K_2CO_3, CH_3CN, 3 h, rt.

Herein, we explored only *O*-alkylated amino acid derivatives and investigated the role of pyrazolonyl amino acid residue in conformational changes of their peptides. We choose neutral hydrophobic, simple and chiral amino acids with increasing bulkiness such as glycine (*Gly*), alanine (*Ala*), isoleucine (*Ile*) and phenylalanine (*Phe*). Moreover, *Phe* may involve in

conformational changes of peptides by π-π stacking with APA moiety. Additionally, we planned to couple APA with dipeptides, where the sterically hindered *Ile* group is switched between APA unit and flexible *Gly* unit. Accordingly, the *O*-alkylated amino acid derivatives (**4a**/**4b**) were incorporated at *N*-terminal of natural amino acids and dipeptides. Hence, **4a**/**4b** was hydrolyzed with LiOH (1 M) to obtain desired *N*-Boc-Aminopyrozolonyl acid (**4a/4b-OH**) and coupled with neutral natural amino acids/peptides. The synthesized APA peptide library as shown in Scheme 4A.2,



Scheme 4A.2. Synthesis of 4-aminopyrazolonyl (APA) peptides (**7a**–**7n**).

X-ray studies. After several attempts under various crystallizing conditions we obtained crystal of **3**/**4a**/**5a**/**6a** derivatives, APA dipeptide (**7a**) and tripeptide (**7e**). These single crystal structures

were analyzed by X-ray studies which confirmed the structure of respective APA derivatives and peptides. The ORTEP diagrams of these compounds are provided in Figure 4A.2 and their crystal parameters and CCDC numbers are provided in Appendix Table S2.

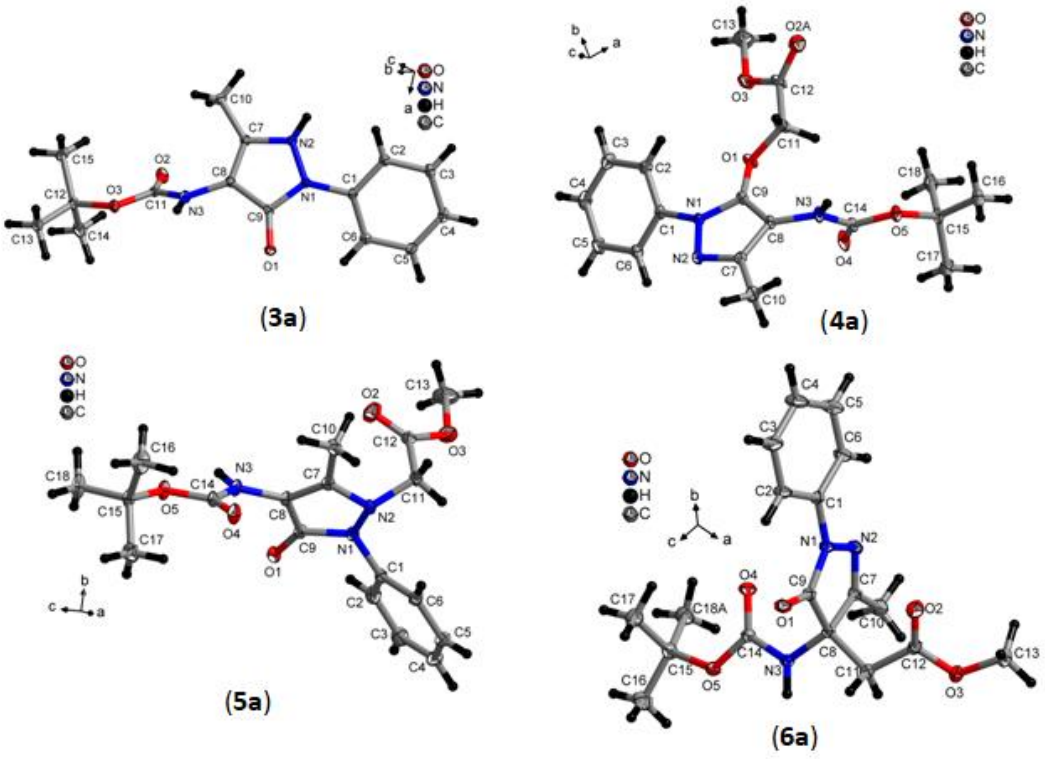


Figure 4A.2. Crystal structures of compound **3a**, and regioisomers **4a**, **5a**, **6a**.

To understand the acquired APA peptide conformation in crystal lattice, we studied their solid state crystal packing arrangement. The crystal packing arrangement of amino pyrazolone **3** exhibit intermolecular hydrogen bonding. Each molecule of amino pyrazolone **3** is involved in intermolecular H-bonding with three other such units to form a non-linear type of sheets. In this intermolecular H-bonding network, three H-bonding sites are involved in each molecule to form four hydrogen bonds. Two hydrogen bonds between pyrC=O···H-NBoc and other two between pyrC=O···H-Npyr (Figure 4A.3). Importantly, pyrazolonyl ring carbonyl, pyrC=O is involved in

bifurcated hydrogen bonding with BocN-H and pyrN-H, which generated unique type of supramolecular self-assembly with marginal difference in their hydrogen bond length. At the outset, these potential hydrogen bonding sites in pyrazolone residue encouraged us to design their peptides.

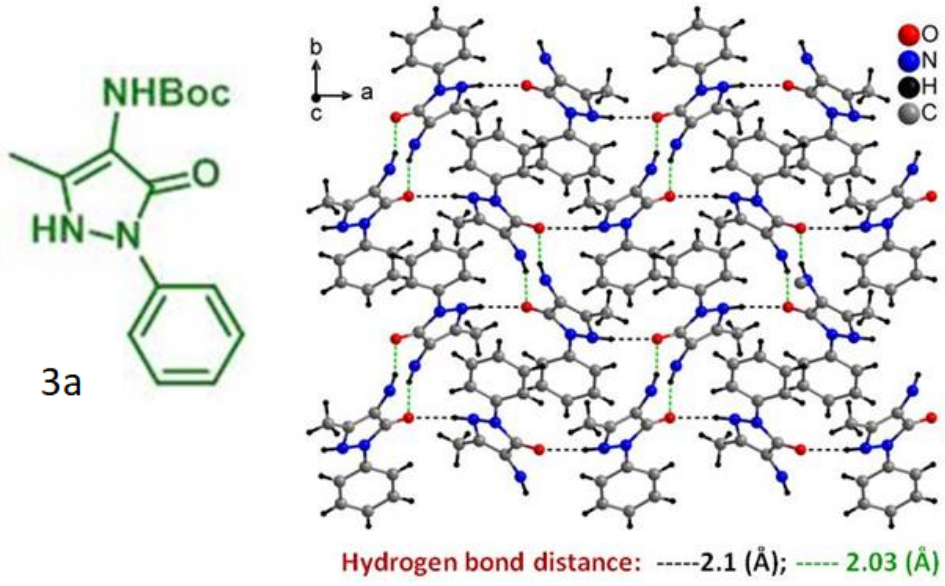


Figure 4A.3. Crystal packing arrangement of 4-aminopyrazolone (**3a**) (Boc group has omitted for representation).

The crystal packing arrangement of all three APA regio-isomers **4a**/**5a**/**6** acquired distinctive types of packing arrangements. The crystal packing arrangement of *O*-alkylated ester derivative (**4a**) exhibits intermolecular H-bonding between pyrN···H-NBoc with a bond length of 2.1 Å and forms helical structure along b-axis with the pitch value of ca. 10.8 Å (Figure 4A.4)

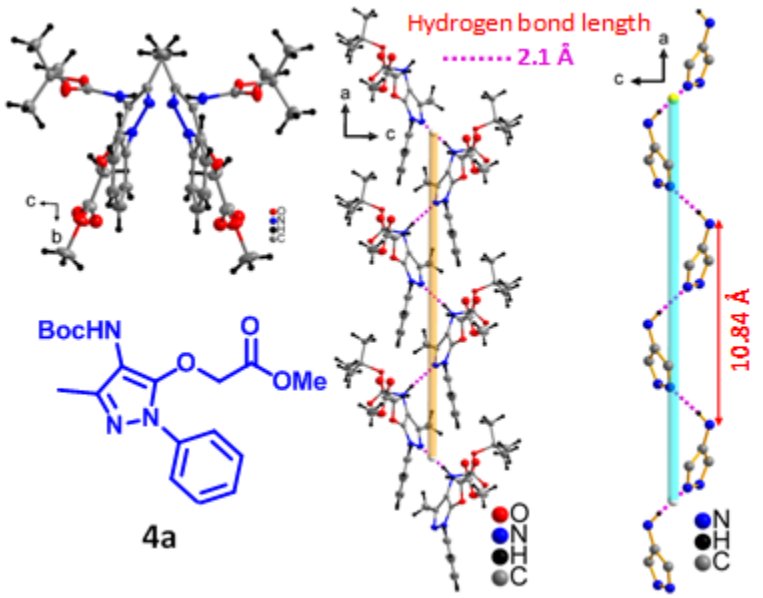


Figure 4A.4. Crystal packing pattern of APA ester derivative (**4a**) with intermolecular hydrogen bonding (pink dotted line) in different axis (few atoms are omitted for representation in middle and right).

The crystal packing arrangement of *N*-alkylated APA derivative **5a** featured intermolecular H-bonds between pyrC=O···H-NBoc with distance H-bond length of 2.0 Å. This H-bonding plays important role to provide extra stability to **5a** and form repeated dimmers in anti-parallel manner, which featured in a supramolecular linear β-sheet type of structures (Figure 4A.5A). In the C-alkylated APA derivative **6a** noticed two types of H-bonding, one is classical and other is non-classical H-bonding. The classical H-bonding is observed between pyrC=O···H-NBoc and forms repeated anti-parallel dimmers (Figure 4A.5B). These anti-parallel dimmers were further stabilized through non-classical H-bonding extended between π-C of phenyl group and BocCH_3 (bond length ranging from 3.1-3.9 Å) and resulted into a unique H-bonded three-dimensional network structure (Appendix). Overall, these crystal lattice studies suggest that pyrazolone ring N-atom (pyrN), pyrazolone ring carbonyl (pyrC=O) and BocN-*H* of APA esters (**4**/**5**/**6**) are involved in

intermolecular hydrogen bonding, and play important role for the formation of helical and anti-parallel-β-sheet types of supramolecular self-assembled structures. In recent literature, native dipeptide derivatives (BocNH-Val-Val-OMe and BocNH-Ile-Ala-OMe) are also known to form parallel β-sheet type of structure.[25]

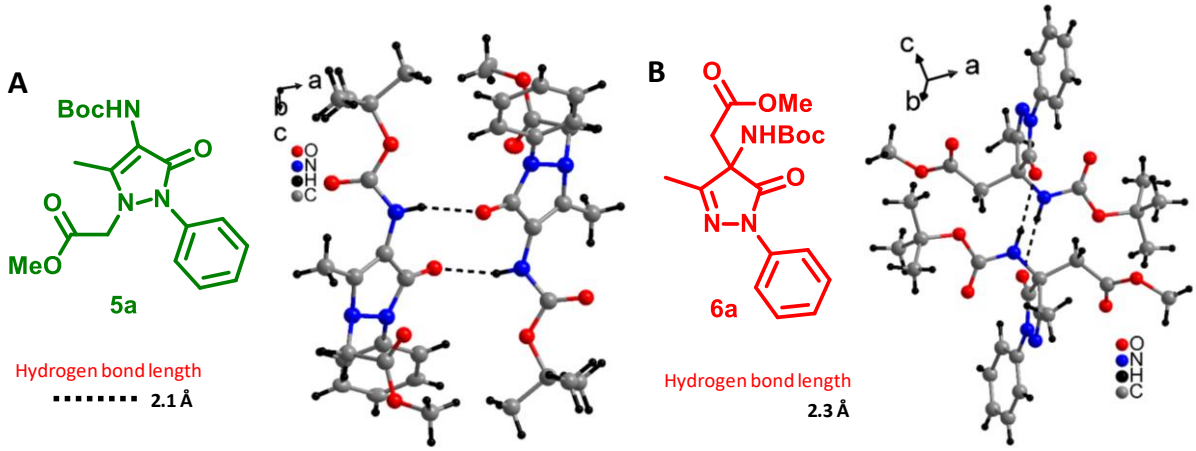


Figure 4A.5. Intermolecular hydrogen bonding in APA ester derivatives, **5a** (A) and **6a** (B).

In case of peptides, the crystal lattice of dipeptide, *BocNH-APA-Ala-OMe* (**7a**) show both classical and T-shaped multiple phenyl CH-π interactions. The classical intermolecular H-bonding is observed among amide N-H with adjacent amide carbonyl (N–H···O=C), and BocN-H with adjacent Boc carbonyl (N–H···O=C). These H-bonding were further stabilized by perpendicularly placed phenyl groups through non-classical T-shaped multiple CH-π interactions with bond lengths ranging from 2.9–3.7 Å (∠140° to ∠180°). This resulted in formation of unique semi-cylindrical channel along *b*-axis with radius of 2.0 Å (Figure 4A.6). In literature, the similar type of self-assembly are reported and considered as parallel β-sheet type of structure.[25]

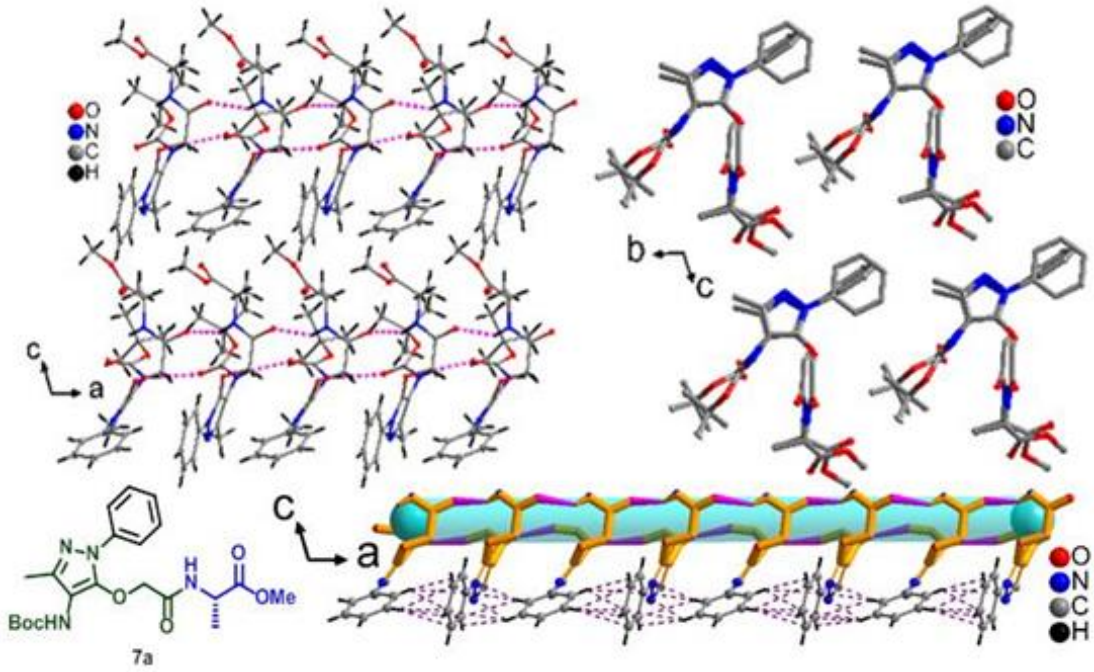


Figure 4A.6. Crystal packing arrangement in APA peptide in different axis (**7a)** with intermolecular hydrogen bonding with phenyl CH-π interactions. Purple dotted line (CH-π interactions) and classical H-bonding (pink) (few atoms are omitted for representation).

The crystal packing arrangement of APA-tripeptide, **7e** featured two types of intermolecular H-bonding, one is Boc carbonyl (–C=O) with *amide* N-H of *Ile* (BocC=O···H-N*Ile*) having bond length of 2.7 Å; and other one among pyrazolone carbonyl with *Gly* amide N-H (*Pyr*-C=O···H–N-*Gly*) with bond length of 2.1 Å (Figure 4A.7). These H-bonding extended along *a*-axis and formed a new type of helical structure with pitch ca. 8.9 Å. Thus the single crystal X-ray analyses reveal that APA-peptides adopt unique conformations with the involvement of pyrazolonyl ring in formation of β-sheet and helices type of supra-molecular structures.

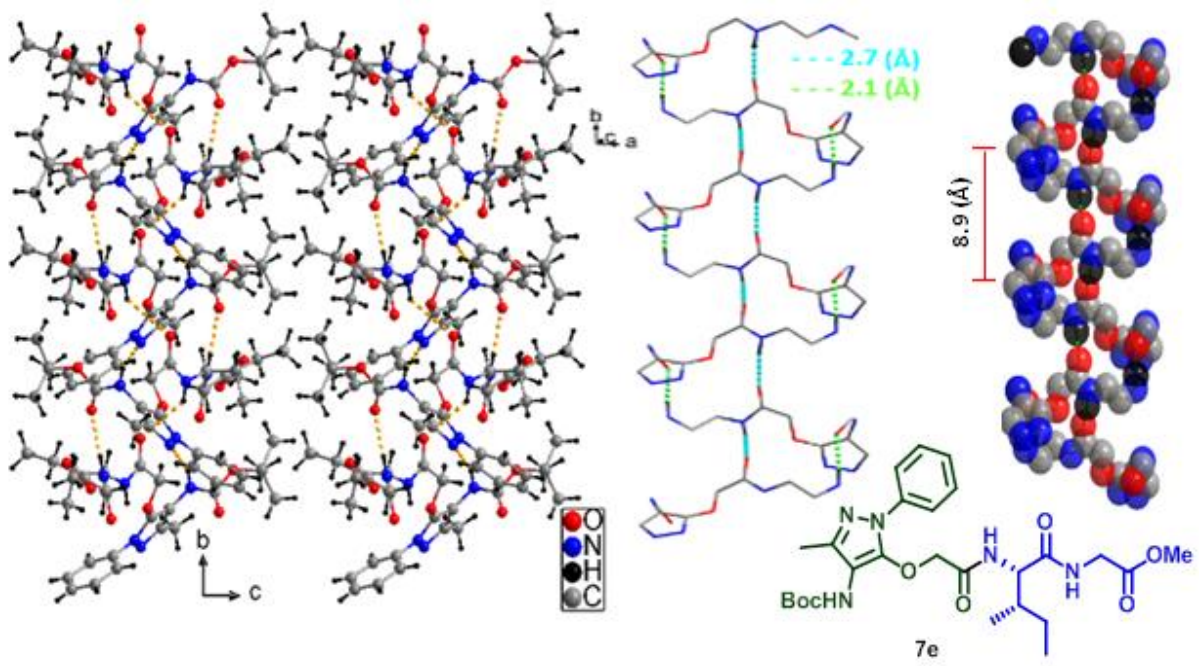


Figure 4A.7. Crystal packing pattern of APA peptide in (**7e**) with intermolecular H-bonding (few atoms are omitted for representation).

It is important to note that the crystal structures of *O*-alkylated ester (**4**) and their peptides (**7a** & **7e**) attain an unusual *syn* conformation between –***O***–*C*H$_2$–*C*O–*N*H– and also exhibit a weak intramolecular H-bonding between *H*-atom (–CO–N*H*–) and *O*-atom (–***O***–*C*H$_2$–). We noticed the H-bond length of 2.25 Å (with ∠108°) and 2.30 Å (with ∠106°) in peptide **7a** and **7e** respectively, Appendix, (Figure S44 and S47–S48).

NMR Studies: Generally, peptides form secondary structures through amide bonds, mainly by inter/intramolecular H-bonding. Herein, APA peptides (**7a**-**7e**) also have amide bonds which can involve in H-bonding. To determine the H-bonding in APA peptides, we performed ^{1}H-NMR experiments which includes: ^{1}H-COSY, ^{1}H-NOESY and DMSO-d_6 ^{1}H-NMR titrations in, $CDCl_3$. Since APA peptides have two to three amide N-H, henceforth, we define chemical shifts for

BocNH as NH(1), and other amide NH's as NH(2)/NH(3) in order of appended peptide chain to APA residue.

The ^{1}H-COSY NMR spectra of APA peptides (**7a**–**7e**) was recorded in $CDCl_3$ and assigned their respective cross peaks. Importantly the N-H protons and their chemical shifts are analyzed which are provided in Table 4A.1. In all APA peptides **7a**-**7e**, the chemical shifts of amide N-H (NH(2)), which is close to APA moiety, appear downfield as compared to NH(1)/NH(3) in their respective peptides. It appears, the NH(2) of APA peptide is more deshielded than their BocNH and other amide N-H.

Table 4A.1. Chemical shifts values of amide N-*H* of APA-peptides (**7a**–**7e**) in $CDCl_3$

Entry	APA peptide	Chemical shift, (δ, in ppm)		
		NH(1)	NH(2)	NH(3)
1	*APA-Ala-OMe* (**7a**)	6.072	6.87	–
2	*APA-Ile-OMe* (**7b**)	5.936	6.851	–
3	*APA-Phe-OMe* (**7c**)	6.092	6.855	–
4	*APA-Gly-Ile-OMe* (**7d**)	6.302	7.118	6.507
5	*APA-Ile-Gly-OMe* (**7e**)	6.568	6.957	6.275

The ^{1}H-NOESY NMR spectra of peptides (**7a**) and (**7c**–**7e**) are also recorded in $CDCl_3$ and assigned all the resolved cross peaks. A portion of NOE spectra in amide region for APA peptide 7a/7e is interpreted and provided in Figure 4A.8. The NOESY spectra of **7a** has seven prominent cross-peaks resulted from amide NH proton, (NH(2)), BocNH proton, (NH(1)) and methylene protons (–*O*–*C****H***$_2$–). The amide NH proton, NH(2), show three correlation peaks (cross peaks assigned as 2, 6 & 7), corresponding to methylene protons (–*O*–*C****H***$_2$–), alanyl residue (–C*H*$_3$) and *N*-phenyl (*ortho*-protons) residue. Interestingly, the methylene protons (–*O*–*C****H***$_2$–) also show

NOE (cross peak 1) with *N*-phenyl (*ortho*-protons) group. While the BocNH, NH(1) shows three weak correlation peaks (cross peaks 3, 4 & 5) with methylene protons (–*O*–*C****H***$_2$–), *tert*-butyl group of BocNH and methyl (*pyr*-CH_3) group on pyrazole ring. Similar NOE interactions were noticed in all other APA peptides (**7c**–**7e**).

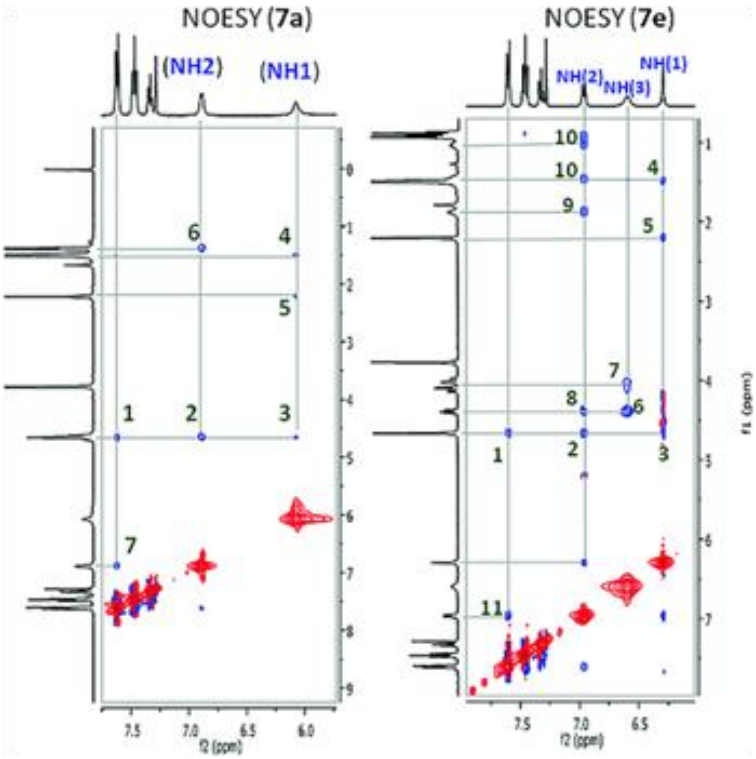


Figure 4A.8. ^{1}H-NOESY spectra of APA peptides **7a** & **7e** in $CDCl_3$. (Showing only amide region).

All these NOE interactions in APA-peptides **7a, 7c**–**7e** was carefully analyzed and summarized in Figure 4A.9. Importantly, in Figure 4A.9, we notice NOE's between methylene protons (–*O*–*C****H***$_2$–) of APA residue with *N*-phenyl group (*ortho*-protons), amide NH [NH(2)] and BocNH [NH(1)]. This support that one of the methylene proton's have NOE with *N*-phenyl group and the other one with BocNH [NH(1)], Additionally amide NH [NH(2)] also shows NOE with *N*-phenyl group

(*ortho*-protons). The observed NOEs also indicate that the methylene protons (–*O*–*C****H***$_2$–) are in close spatial proximity with amide NH.

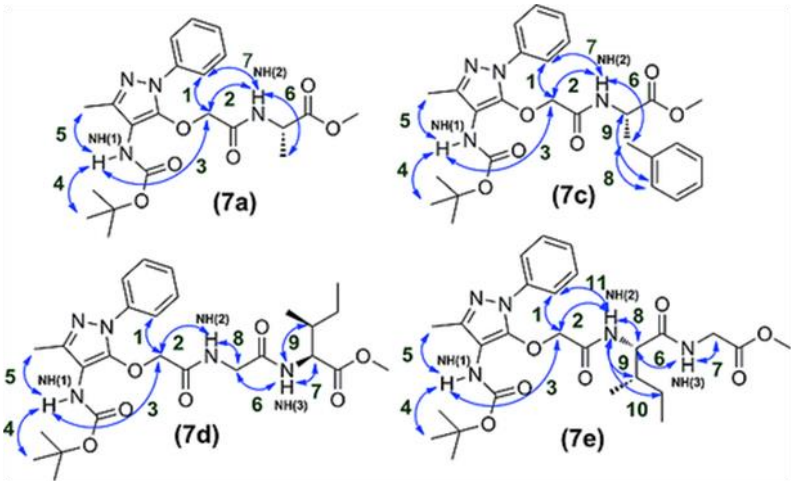


Figure 4A.9. NOE interactions observed in APA peptides in $CDCl_3$ (**7a, 7c**–**7e**).

In general, amide bonds in peptides can form H-bonding; however, they vary in their strength of H-bonding. Some H-bonds are held very strongly, while some of them held weakly and even some of them may not be involved in H-bonding. The relative strength of H-bonding can be studied by NMR. In particular, the NMR studies of DMSO-d6 titration in aprotic deuterated solvent such as $CDCl_3$, Toluene-d8, and CD_2Cl_2 will provide quick insights about the nature of H-bonding. In these solvents the ability to form hydrogen bonding is more pronounced. Since DMSO-d6 has strong H-bond acceptor (S=O), the titration with DMSO-d6 will perturb weekly held H-bonds. As a resultant, the NH peak of weakly held H-bond will shift to downfield as titrated with DMSO-d6, while strongly held H-bonds have relatively no shifts or little shifts.[26] To understand the intricate inter-/intra molecular H-bonding in peptides **7a**-**7e** via amide/BocNH, we carried out DMSO-d6 1H-NMR titration.

The DMSO-d6 titration profile of N-H protons for peptide **7e** is provided in Figure 4A.10. As the titration proceeded with increasing DMSO-d6 volume (upto 20 µl), the amide N-H proton appended close to APA residue, NH(2) has marginal downfield shift, where as BocN-H proton, NH(1) and amide N-H proton, NH(3) relatively experience a significant downfield shift. This indicate that the NH(3) and NH(1) are involved in weak intramolecular H-bonding, while NH(1) experience relatively stronger H-bonding. Similar results were obtained for other peptides featuring the amide N-H proton appended close to APA residue, NH(2) experience relatively stronger H-bonding (Appendix). The weak intramolecular H-bonding is further supported from its supramolecular self assembly structure in its single crystal X-Ray studies, where the amide N-H proton, NH(2), BocN-H proton, NH(1) and amide N-H proton, NH(3) are involved in intermolecular H-bonding, additionally amide N-H proton NH(2) also experience weak intramolecular H-bonding with adjacent methylene protons (–*O*–*C****H***$_2$–).

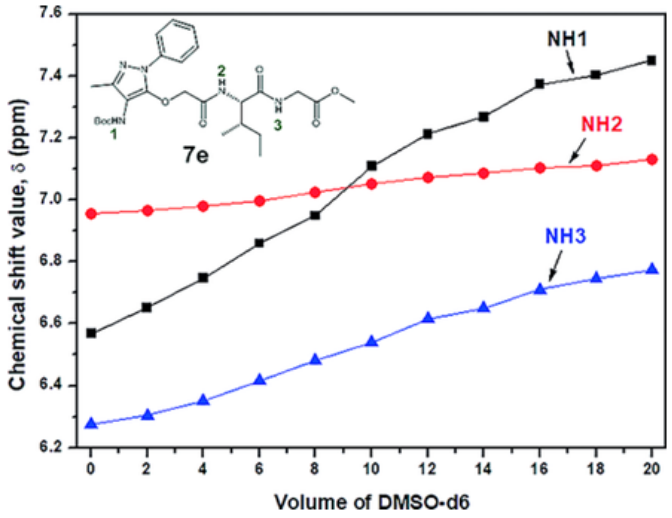


Figure 4A.10. Changes in δ (in ppm) with DMSO-d6 (in µL) titration of peptide **7e** in $CDCl_3$ (≈ 40 mm).

Further, from this titration experiment it is observed that the amino acid residue coupled to APA unit also affected the strength of H-bonding through its amide bond. This is observed from the difference in chemical shift values (dδ) after addition of 20 µl DMSO-d6. The difference in chemical shifts (dδ) of NH(2) for sterically less crowded glycine or alanine residue is relatively more than the bulky isoleucine or phenyl alanine residue (Figure 4A.11). With increasing bulkiness of amino acid appended next to APA residue in peptides **7a**-**7e**, the amide N-H proton (NH(2)) experience relatively stronger H-bonding. This supports, the steric hindrance at that amide position in APA peptides has important role in the formation of rigid structure via intramolecular hydrogen bonding as compared to other N-H.

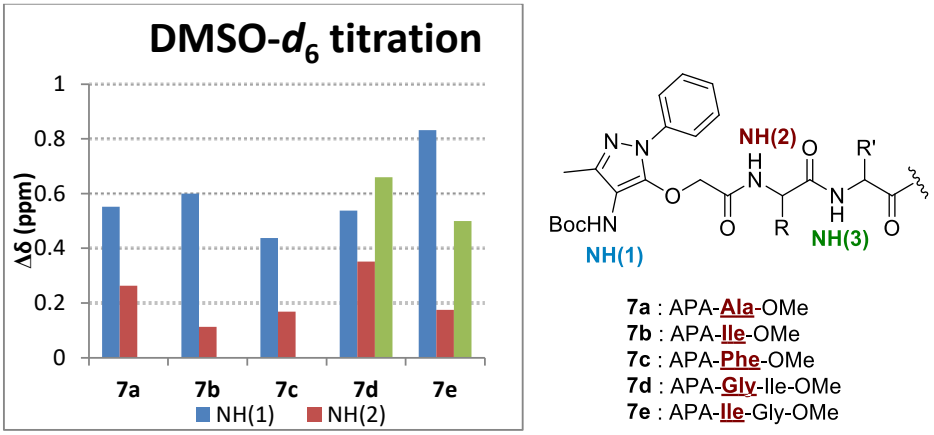


Figure 4A.11. Difference in chemical shift values (dδ) after addition of 20 µl DMSO-d6 for APA-peptides **7a**-**7e**. Respective amides NH are shown in left representative APA structure.

Our NMR studies suggest that NH(2) (amide NH) is most deshielded in solution ($CDCl_3$) as compared to NH(1)/NH(3). Further, DMSO titration NMR experiments confirm that NH(2) is involved in weak intramolecular H-bonding while NH(1)/NH(3) are involved in intermolecular hydrogen bonding in aprotic polar solvents. Importantly, NOESY results show that NOE interactions of APA phenyl (*ortho* C-*H*) residue with –O–CH_2- and amide N*H*(2) which are

consistent in all APA-peptides. Herein, we propose the preferable conformation of APA-peptides in solution as shown in Figure 4A.12A which have acquired *syn* conformation between O–CH_2– and amide N*H*(2) and oriented appropriately for intramolecular hydrogen bonding as N–H···O–CH_2 as Figure 4A.12B. Further, crystal structure of two APA-peptides (**7a**/**7e**) have also acquired the similar type of *syn* conformation by O-CH_2-and amide N*H*(2) in solid state (Figure 4A.12C) From NMR and X-ray correlation studies, we propose structure of APA peptides as Figure 4A.12D which show inter-/intramolecular hydrogen bonding in solution state.

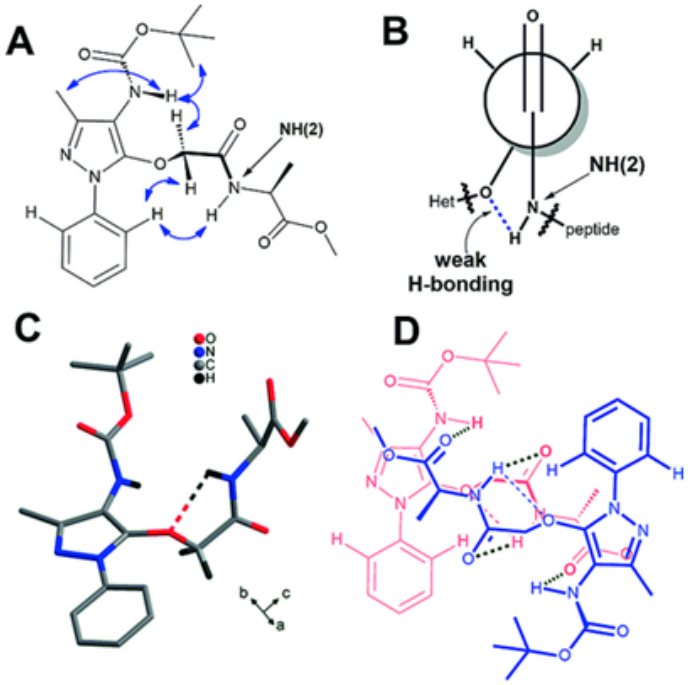


Figure 4A.12. Proposed structures and conformations in solution state and observed single-crystal X-ray structure of **7a**. A. possible structure of **7a** from NOESY NMR; B. *syn* conformation with weak intramolecular H-bonding; C. Single-crystal X-ray structure of **7a** with intramolecular H-bonding; D. Possible intermolecular H-bonding in solution for **7a**.

Organogelation:

Gelation is a well defined structural organization resulting from intermolecular interactions such as: intermolecular H-bonding, pi-interactions, van der waals interactions, solvophobic effects. Reversible gellators form self-assembly by non-covalent interactions in three dimensional networks which are easily destroyed and reorganized. The process of gelation from organic solvents is known as organogelation. Oligopeptides and small peptides are widely applied for the formation of versatile supramolecular organogels through these non-covalent interactions (known as physical gels). Sono-gels are class of gels, which are formed under sonication. The process of gelation activated by ultrasound could at first sight be contradictory, because sonication is also responsible for the breaking of oligomeric and supramolecular entities in the liquid phase. However, sonication can break the larger aggregates, otherwise disordered aggregates and can induce the formation of well-defined larger uniform aggregates, which lead to the formation of gels.[27] By seeing a small peptide sequence we cannot predict the formation of organogels. It is the unique feature of that peptide under the employed experimental condition. For example, *N*-Fmoc-Leucine- *C*-benzoyl- Leucine dipeptide forms organogel by sonication, but within this dipeptide lacking the benzoyl or with a benzoyl replacing the Fmoc group do not show any solvent gelation. This is specific structural feature of this dipeptide to gelate.[28] The formation of peptide organogels is not fully understood and it is still at infancy. From our X-ray/NMR/CD studies of APA peptides reveal the presence of intermolecular interactions, which is one of the important criteria for gelation. Hence, these results encouraged us to investigate APA peptides for organogelation.

The formation of organogels for small peptides is reported in recent literature. For an example, *di-/tri*-peptides form organogel in co-solvent systems (hexane:ethylacetate) by sonication.[29, 30] Herein, we also attempted the organogelation of APA-linked *di-/tri*-peptides (**7a-7j**) in similar solvent systems by sonication. All APA peptides **7a-7j** 0.7-0.9% (wt/vol) were dissolved in solvent

systems EtOAc:Hexane (1:3) and sonicated for two minutes above the room temperature (~50 ^{o}C); then allowed to cool at room temperature. Pleasantly we noticed that the homogenous solutions of peptides (**7g-7j)** were transformed into colourless organogel within 10 minutes after sonication. We also noticed the variation in physical appearance such as transparent or opaque in the gels. The gel formation is not observed in other solvents such as hexane, ethyl acetate, benzene, chloroform, acetonitrile, and methanol. Mostly these peptides are sparingly soluble/or appeared as precipitates in these solvents. Formation of weak gel was observed for **7h** in toluene-d_8 after 2-3 days left after sonication. We observed initial precipitation of APA peptide **7j** at room temperature; these precipitates were dissolved upon heating. After sonication for two minutes and cooling, we observed the formation of gel. In literature, precipitates can also help in formation of larger aggregates which can transform into organogels.[31, 32] The APA-peptide organogels are stable up to 50-55 ^{o}C, but melted at higher temperature and eventually resulted in clear solutions. The formation of organogels was validated by widely accepted inverted test tube method as shown in Figure 4A.13. The surface morphology of peptide organogels (**7g-7j**) are studied by FE-SEM imaging technique. Their SEM images are depicted in Figure 4A.13.The SEM images of organogel from peptide **7g** reveal the nature of oragnogel surface as fibre type of self-assembly structure (Figure 4A.13A). The surface morphology of transparent organogel from peptide **7h** exhibits a polymeric structure (Figure 4A.13B). The organogel from peptide **7i**, exhibits flower type of self-assembly structure (Figure 4A.13C). Finally SEM-images of organogel from peptide **7j** also reveal the formation of fibre type of self-assembled structure (Figure 4A.13D). We could not find the formation of organogel from other APA-linked peptides (**7a-7f**).

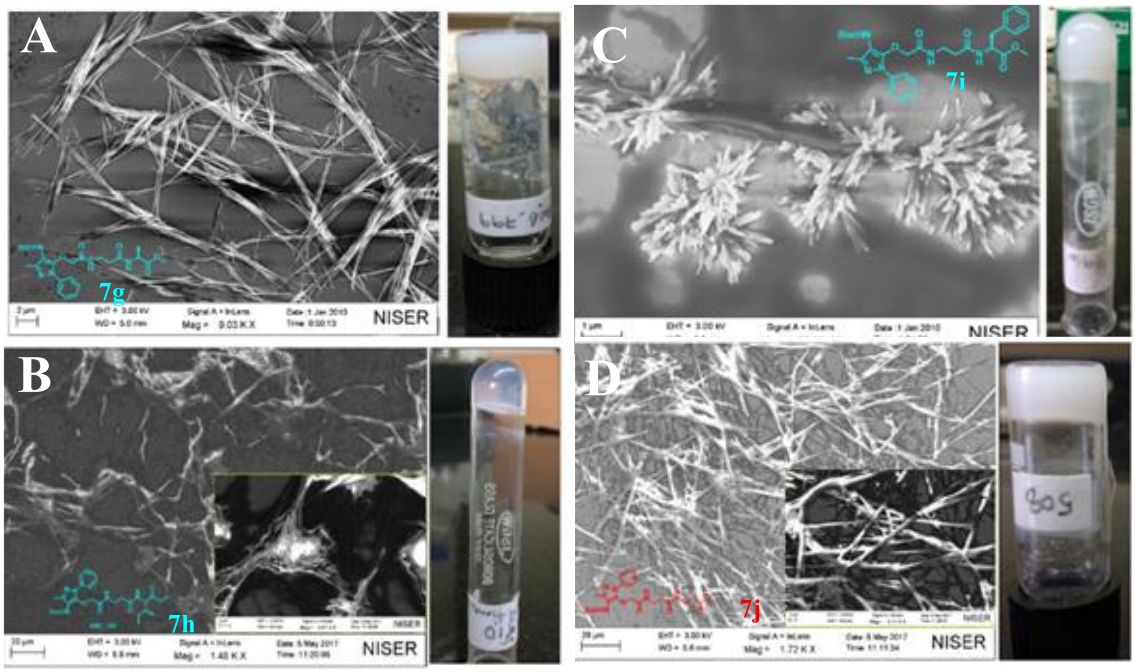


Figure 4A.13. Aminopyrazolonylated α/β-hybrid peptide gels. SEM-images and organogels in inverted test tube: (A) peptide **7g**; (B) Peptide **7h**; (C) Peptide **7i**; and (D) peptide **7j**.

The stabilization of homogenous aggregates to form gel network needs intermolecular interactions and π-π interactions. NMR and X-Ray studies of APA α-peptides (**7a-7e**) also suggest the presence of intermolecular interactions and π-π interactions. Gelation is not observed when β-Ala is replaced with α-amino acids such as gly, ilu, and ala. Interestingly, the physical organogels are formed only from hybrid peptides (**7g-7h**) but not from α-peptides (**7a-7e**) and hybrid peptides lacking N-phenyl groups (**7k-7n**). Indicating, the presence of β-Ala at that position is crucial for the formation of gels. Possibly, the presence of β-Ala (extra methylene) group will increase the chain length which affects the intramolecular H-bonding interactions and additionally, this also increases the flexibility in sol-state which can reorganize easily during the formation of the rigid gel. The optimum flexibility and relatively dominant intermolecular interactions in these hybrid peptides (**7g-7j**), induce the formation of homogenous ordered networking aggregates, which

appear as physical gels. Similar networking aggregates are also possible in other APA peptides (**7a-7e** and **7k-7n)** but they may not appear as relatively stronger physical gels as in **7g-7j** under these examined conditions.

CD-Studies. The peptide secondary structures were exclusively studied by circular dichroism (CD). However, the peptide conformations are sensitive to the nature of solvent environment and play significant role in formation of secondary structures.[33] The solvent dependent CD spectra of peptides have prominent role in finding the defined structure and conformation of peptides.[34, 35] [15] To examine the role of APA residue in the conformational changes in their peptides, we recorded the CD spectra of APA peptides (**7a**–**7j**) in different solvent systems AcCN, MeOH, $CHCl_3$ and TFE, and their CD spectra of **7a**–**7j** are provided in Figure 4A.14, while other spectra are provided in Appendix (Figure S50–S53). The CD spectra of APA di-/tri-peptides (**7a**–**7e**) are similar in AcCN and MeOH, which exhibit Cotton effect with distinctive CD signals as the maxima at wavelength (λ) ≈ 200 nm/220 nm and the minima at $\lambda \approx 260$ nm, though the maxima at $\lambda = 220$ nm is more prominent in MeOH. In solvent $CHCl_3$, however, the CD signal of peptides (**7a**–**7e**) exhibit poorly resolved maxima and minima. In solvent TFE, the CD spectra of peptides (**7a**–**7e**) exhibit only maxima at $\lambda = 200$ nm & $\lambda = 220$ nm. Under similar conditions, the CD spectra of β/α-hybrid APA peptides **7f-7j** exhibit similar CD signature as APA α-peptides **7a-7e**. These CD signals of peptides are possibly due to electronic transitions of amide carbonyl group (π–π*/n–π*) and pyrazolonyl/phenyl aromatic rings (π–π*). We recorded the UV-spectra of APA peptide **7c** in CH_3CN and extracted wavelength maxima (λ_{max}) as 220 nm and 250 nm. In CD spectra of these peptides, overlapping of aromatic chromophoric (pyrazolonyl/phenyl) absorption of APA peptide (220 nm to 280 nm) with finger print region of peptide secondary structure (190 nm to 240 nm) is observed. The unresolved CD spectra in all these solvents made difficult to interpret the secondary

structure. However, it is anticipated that, by changing the solvents, we can study the structural variations and may facilitates the possibilities to understand the structural behavior. Herein, the CD spectra of these APA peptides are found to be solvent dependent. The CD structure is more prominent in AcCN, MeOH solvents compared to TFE. There are no strong CD signals in TFE. In literature, TFE is well known to induce the formation of stable helical conformations in peptides which are generally undefined in other solvents.[36] Hence, there are no such strong intramolecular H-bonds to form stable helical conformations in all APA peptides. The CD structures of APA peptides in other solvents are presumed due to intermolecular H-bonding; this is further supported from our NMR and X-ray studies. Importantly, the CD pattern for APA α-peptides and β/α-hybrid peptides are very similar. The presence of β-Ala in hybrid peptides (**7f**-**7j**) has no significant affect in CD pattern.

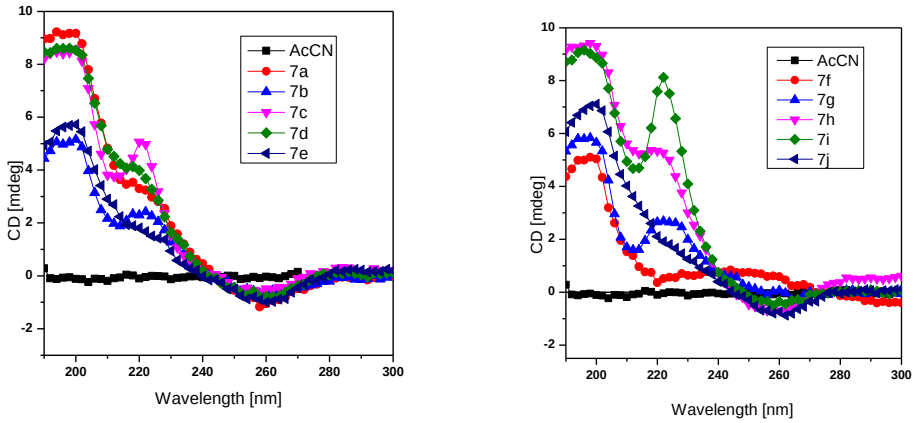


Figure 4A.14. CD spectra of APA peptides **7a**-**7e** (A) and **7f**-**7j** (B) in AcCN solvent.

Since β/α-hybrid peptides are forming gels in ethyl acetate and hexane mixture (1:3) hence, we recorded CD spectra in this solvent mixture for at least two peptide gels, **7g** and **7h** (Figure 4A.15). The CD spectra of organogels are comparatively different from other solvents such as AcCN,

$CHCl_3$, MeOH and TFE. Importantly, in oragnogels the minima at $\lambda \approx 260$ nm, is shifted to $\lambda \approx$ 300 nm in ethyl acetate and hexane mixture (1:3). This red shift (≈ 40 nm) indicates the presence of predominant intermolecular interactions in ethyl acetate and hexane mixture (1:3), which are not observed in other solvents systems. However, the CD spectra of organogels before and after sonication are similar in pattern and not affected by sonications.

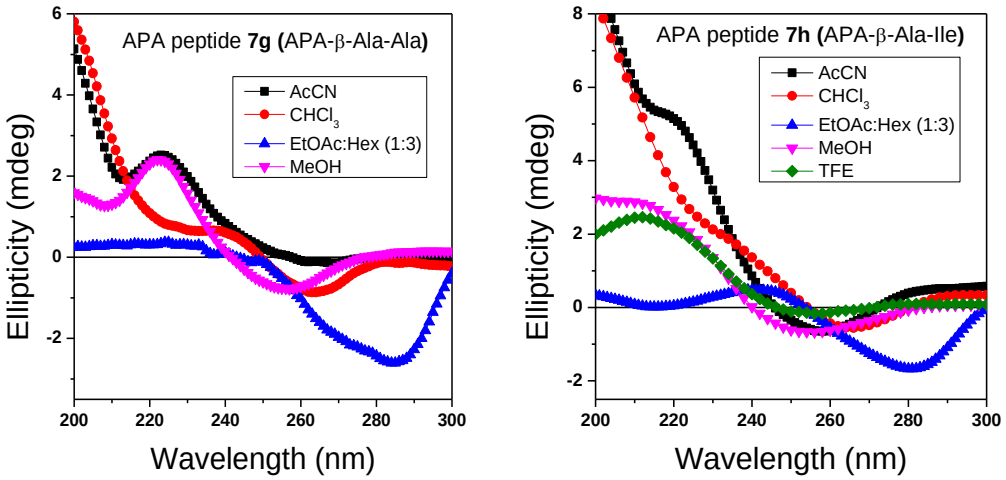


Figure 4A.15. CD spectra of organogels **7g** and **7h** in different solvents. The CD spectra in ethyl acetate and hexane mixture (blue) is significantly shifted.

NMR-Studies. In literature, the formation of peptide organogels are also studied by ^{1}H-NMR in deuterated solvent (toluene-d_8) which exhibit significant downfield chemical shift of amide N-H.[29] Herein, we also performed the similar NMR studies with one of organogel forming peptide **7h** in toluene-d_8. The NMR spectra of N-H protons of peptide **7h** are depicted in Figure 4A.16, which clearly indicates the notable chemical shift of its N-H protons after sonication. This indicates the presence of H-bonding and they reorganize to relatively stronger H-bonding environment after sonication. Similar NMR experiment with others organogel forming peptides (**7g**/**7i**/**7j**) could not succeed to record NMR in toluene-d_8 either they are instantly soluble and/or gelate instantly.

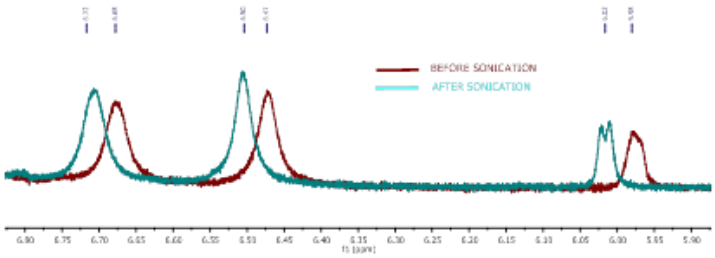


Figure 4A.16. Expanded NMR of peptide **7h** amide N-H region of organogel in toluene-d_8 before (red) and after sonication (turquoise).

FT-IR studies of organogels.

FT-IR is widely used for determining the secondary structure of proteins and peptides. Peptides acquire wide structural diversity through H-bonding via amide bonds. Due to variations in strength of H-bonding, the amide bond vibrations perturb from normal amide bond vibrations. The study of amide carbonyl and amide NH vibrations provides valuable information about the acquired structures such as α-helix and β-sheets/turns. From proteins and peptides there are nine characteristics bands from amide vibrations named as amide A, B, I, II ... VII. From these, amide-A band about 3220-3320 cm^{-1} is associated with NH stretching vibrations and amide I band about 1620-1710 cm^{-1} is associated with C=O stretching vibrations. Primarily, the amide-A and amide-I bands are most intense bands and depict the peptide conformations. The common correlation between protein/peptide secondary structure and Amide-I bands are provided in Table 4A.2. Importantly, B-turns have two bands, arising from acceptor carbonyls at 1645-1635 cm^{-1} and constrained carbonyls at 1690-1660 cm^{-1}. The structure assignment of α-helix and β-turn will become difficult if these two bands are not resolved.

Table 4A.2. Commonly observed Amide-I frequency band in peptide and protein

Structure	Amide-I frequency band (cm^{-1})
α-helix	1648-1660
β-sheet	1625-1640
β-turn (have two types of bands)	1645-1635; 1690-1660

Similar to these intrinsic secondary structures, small peptides such as di/tri/tetra peptides can also form supramolecular self-assembled α-helix/β-sheets/β-turns like structures through intermolecular interactions. Since, peptide organogels are also self-assembled supramolecular structures, their amide carbonyl and amide NH vibrations provide valuable information about the acquired structures in their gel forms. Since these structures are non-covalent in nature, these are solvent dependent. To acquire information on the secondary structure of the self-assembled structure as organogel, we performed Fourier Transform Infrared Spectroscopy (FT-IR) analysis with neat xerogels (dried organogels) of peptides (**7f-7j**) and HFIP dissolved xerogels. We choose HFIP here because, it is known to stabilize or induce helical like structures including β-turns by destabilizing β-sheets. But this will not completely discriminate H-bonding.[37, 38]

The neat xerogel from peptides (**7f-7j**) showed highly resolved four carbonyl peaks in their respective spectra: two peaks at ~1645-1657 cm^{-1} for amide carbonyls (Amide-I), one peak at ~1684-1693 cm^{-1} for carbamate carbonyl, and one peak at ~1736-1763 cm^{-1} for ester carbonyl (Figure 4A.17A-E). The N-H (Amide-1) stretching vibrations appear at ~3276-3312 cm^{-1}, which indicates their involvement in H-bonding (Figure 4A.17F-J). However, no such resolved IR peaks were noticed with HFIP dissolved xerogels. All the carbonyls of amide, carbamate, and ester appeared as one broad peak at ~1642-1674 cm^{-1}. While N-H (Amide-1) appears slightly red-shift

at ~3302-3427 cm^{-1}, this also indicates their involvement in H-bonding. In literature FT-IR signals of amide carbonyl in peptides near 1648-1660 cm^{-1} belong to α-helix or β-turn and another band near 1690 cm^{-1} belongs to the second band of β-turn conformation.[37, 39] Thus our FT-IR spectral analyses support the formation of α-helix and β-turn type of structures in organogels (**7f-7j**).

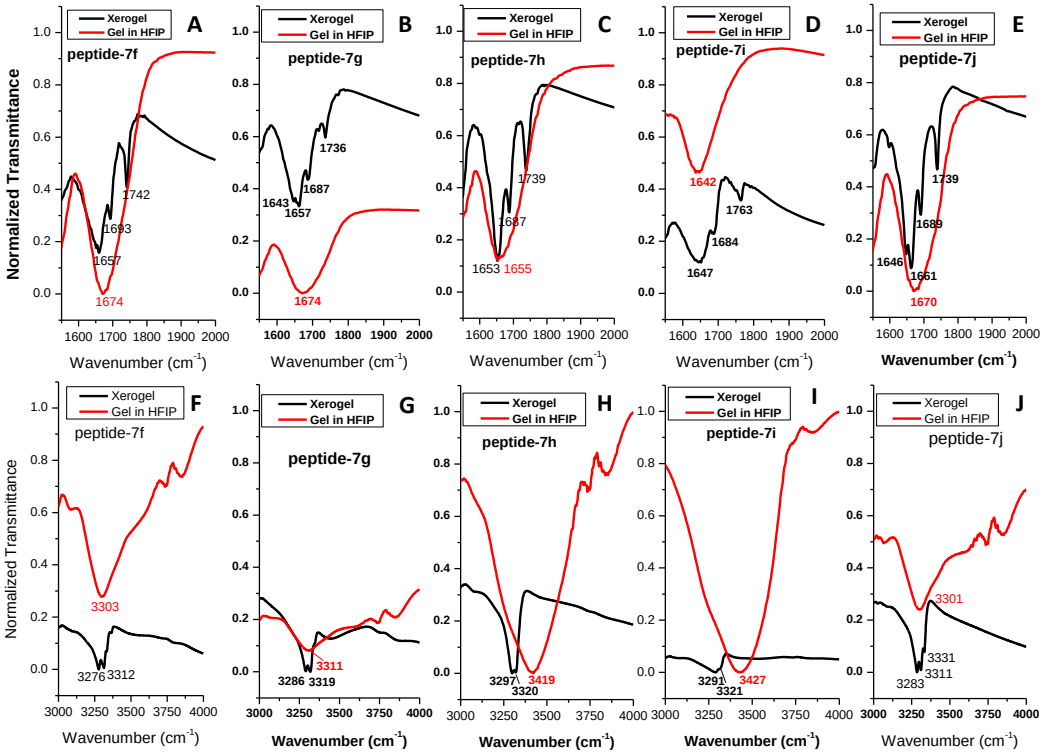


Figure 4A.17. FT-IR of **7f-7j** APA peptide xerogels and gels dissolved in HFIP. Amide-I, carbonyl region (A-E), Amide-A, amide NH region (F-J).

4A.4 Conclusion

In conclusion, three new 4-aminopyrazolonyl amino acid (APA) derivatives were successfully synthesized from the same precursor and their structures were confirmed by single crystal X-ray analyses. One of new amino acid analogue, *O*-alkylalted pyrazolonyl amino acid derivative, is linked at N-terminal α-peptides and β-α-hybrid peptides. Importantly the crystal structure of pyrazolonyl linked peptides, at least in two derivatives obtained by single crystal X-ray analysis

confirmed the distinctive type of supramolecular self-assembled structures due to pyrazolonyl residue. More interestingly, pyrazolonylated β-/α-hybrid peptides form new peptide organogels in co-solvents systems hexane:EtOAc with sonication. The surface morphology of those gels are demonstrated by SEM-imaging technique. The role of aminopyrazolonyl residues are accomplished in conformational and structural changes of pyrazolonyl-linked peptides by CD spectral analyses. The self-assembled structures of pyrazolonylated peptides organogel are determined as β-turn and helix type of structure by FT-IR of dried organogels. So far no studies are available about 4-aminopyrazolone linked amino acids and their organogels. Hence these studies are promising for tuning secondary structures/organogel formation of peptides by linking aminopyrazolonyl amino acid analogue at N-terminal of peptides.

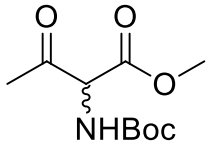
4A.5 Experimental section

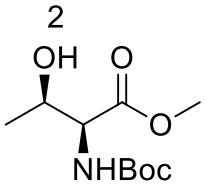
Materials

All required materials were obtained from commercial suppliers and used without any further purification. Dimethylformamide was distilled over calcium hydride. Reactions were monitored by thin layer chromatography, visualized by UV and Ninhydrin. Column chromatography was performed in 230-400 mesh silica. FT-IR spectra were recorded on Thermo Scientific Nicolet iS5N FT-NIR spectrometer Mass spectra and HRMS were obtained from Bruker micrOTOF-Q II Spectrometer. ^{1}H NMR, ^{13}C NMR, were recorded on Bruker AV-400 or 700 MHz at 298 K. ^{1}H and ^{13}C NMR chemical shifts were recorded in ppm downfield from tetramethylsilane or residual solvent peak. Splitting patterns are abbreviated as: s, Singlet; d, doublet; dd, doublet of doublet; t, triplet; q, quartet; dq, doublet of quartet; m, multiplet.

Experimental procedure for compounds, **2-7**:

(2S,3R)-methyl-2-((tert-butoxycarbonyl) amino)-3 hydroxybutanoate. Threonine (5 g, 42 mmol) dissolved in methanol (25 mL) and cooled to 0 ^{o}C, thionyl chloride (3.7 mL, 50.4 mmol) was added dropwise and warmed to room temperature then refluxed for 16 h. The crude reaction mixture concentrated to dryness and washed with hexane twice and concentrated under reduced pressure to yeiled methyl ester. The crude methyl ester (7.1 g, 42 mmol) dissolved in CH_3CN:H_2O (85:15) (85 mL) and $N(Et)_3$ (12.8 mL, 92.4 mmol) and cooled to 0 ^{o}C. $(Boc)_2O$ is added dropwise and allowed to stirr at room temperature for 18 h. The crude reaction mixtureis concentrated and partitioned between H_2O (50 mL) and EtOAc (100 mL) and extracted three times. The combined organic extracts were washed with saturated aqueous NaCl (30 mL), dried over anhydrous Na_2SO_4 and concentrated under vacuum. The resulting crude viscous oil was purified via column chromatography with 40% EtOAc in hexanes to yield 8.08 g (82%) as oil. *Rf* = 0.32 (1:1.5 EtOAc/hexanes); ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.39 (d, *J* = 8.2 Hz, 1H), 4.25 (dd, *J* = 16.2, 7.7 Hz, 2H), 3.75 (s, 3H), 2.45 (s, 1H), 1.43 (s, 9H), 1.22 (d, *J* = 6.4 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 172.16, 156.29, 80.22, 68.16, 58.86, 52.59, 28.38, 19.96.

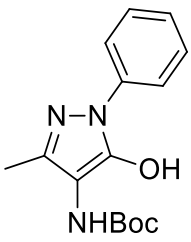




(S)-Methyl-2-((tert-butoxycarbonyl)amino)-3-oxobutanoate (2). Dess-Martin periodinane (1.8 g, 4.29 mmol) dissolved in DCM (20 mL) and stirred at room temperature for 20 min. N-Boc-Threoninemethylester (500 mg, 2.145 mmol) dissolved in DCM (5 mL), was added drop wise in 5 min. Water saturated DCM (20 mL)was added drop wise over 2 h and stirred at room temperature for an additional2 h. The reaction mixture evaporated to half of the volume and was triturated with ether (50 mL), followed by filtration through celite. Ether layer was washed with 1:1 mixture of saturated aqueous $Na_2S_2O_3$ and saturated aqueous $NaHCO_3$ (50 mL). The aqueous layer is further

extracted with ether (2*30 mL).The combined organic extracts were washed with saturated aqueous NaCl (30 mL), dried over anhydrous Na_2SO_4 and concentrated under vacuum. The resulting crude viscous oil was purified via chromatography with 10% EtOAcinhexanes to yield 0.421 g (85%) as pale yellow oil. R_f = 0.26 (1:9 EtOAc/hexanes); ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.71 (d, *J* = 5.8 Hz, 1H), 5.01 (d, *J* = 7.0 Hz, 1H), 3.78 (s, 3H), 2.32 (s, 3H), 1.40 (s, 9H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 198.89 (s), 167.02 (s), 154.88 (s), 80.58 (s), 64.06 (s), 53.16 (s), 28.18 (s), 27.75 (s). HRMS (ESI-TOF) m/z: $[M+Na]^+$ Calcd. for $C_{10}H_{17}NO_5Na^+$ 254.0999; Found 254.0999.

tert-butyl(5-methyl-3-oxo-2-phenyl-2,3-dihydro-1H-pyrazol-4-yl)carbamate (3). To a stirred solution of **2** (1.106 g, 4.78 mmol) in MeOH (40 mL) was added phenyl hydrazine (0.44 mL, 4.88 mmol) and heated at 80 ºC for 5 min. Then glacial acetic acid (0.82 mL, 14.3 mmol) was added drop wise and refluxed at 80 ºC for 20 h. The crude reaction mixture was evaporated to dryness, co evaporated twice with $CHCl_3$and precipitated with $CHCl_3$ and hexane to get a solid, which is further washed with hexane (3-5 times) and dried to yield 1.19 g (86.4 %) of **3** a**s** a pale yellow powder which is sufficiently pure to proceed to next step. R_f = 0.2 (3:7 EtOAc/hexanes); ^{1}H NMR (400 MHz, DMSO) δ 11.03 (br s, 1H), 7.90 (s, 1H), 7.70 (d, *J* = 7.9 Hz, 2H), 7.43 (t, *J* = 7.9 Hz, 2H), 7.21 (t, *J* = 7.3 Hz, 1H), 2.02 (s, 3H), 1.43 (s, 9H). ^{13}C NMR (101 MHz, DMSO) δ 154.67 (s), 145.93 (s), 128.89 (s), 124.91 (s), 119.78 (s), 78.42 (s), 28.15 (s), 11.46 (s); mp: 170-171 ºC.



3

4A.6 Reference

1. Fustero, S.; Sanchez-Rosello, M.; Barrio, P.; Simon-Fuentes, A., From 2000 to mid-2010: a fruitful decade for the synthesis of pyrazoles. *Chem. Rev.* **2011,** *111*, 6984-7034.

2. Varvounis, G., Chapter 2 Pyrazol-3-ones. Part IV: Synthesis and Applications. *Adv. Heterocycl. Chem.* **2009,** *98*, 143-224.

3. Zhou, J.; Huang, W.-J.; Jiang, G.-F., Synthesis of Chiral Pyrazolone and Spiropyrazolone Derivatives through Squaramide-Catalyzed Reaction of Pyrazolin-5-ones with o-Quinone Methides. *Org. Lett.* **2018,** *20*, 1158-1161.

4. Horton, D. A.; Bourne, G. T.; Smythe, M. L., The combinatorial synthesis of bicyclic privileged structures or privileged substructures. *Chem. Rev.* **2003,** *103*, 893-930.

5. Ariza, A.; Garcia-Martin, E.; Salas, M.; Montanez, M. I.; Mayorga, C.; Blanca-Lopez, N.; Andreu, I.; Perkins, J.; Blanca, M.; Agundez, J. A.; Torres, M. J., Pyrazolones metabolites are relevant for identifying selective anaphylaxis to metamizole. *Sci. Rep.* **2016,** *6*, 23845.

6. Hadi, V.; Koh, Y. H.; Sanchez, T. W.; Barrios, D.; Neamati, N.; Jung, K. W., Development of the next generation of HIV-1 integrase inhibitors: pyrazolone as a novel inhibitor scaffold. *Bioorg. Med. Chem. Lett.* **2010,** *20*, 6854-7.

7. Li, Y.; Liu, Y.; Wang, H.; Xiong, X.; Wei, P.; Li, F., Synthesis, crystal structure, vibration spectral, and DFT studies of 4-aminoantipyrine and its derivatives. *Molecules* **2013,** *18*, 877-893.

8. Kurumbail, R. G.; Stevens, A. M.; Gierse, J. K.; McDonald, J. J.; Stegeman, R. A.; Pak, J. Y.; Gildehaus, D.; Miyashiro, J. M.; Penning, T. D.; Seibert, K.; Isakson, P. C.; Stallings, W. C., Structural basis for selective inhibition of cyclooxygenase-2 by anti-inflammatory agents. *Nature* **1996,** *384*, 644-648.

9. Picot, D.; Loll, P. J.; Garavito, R. M., The X-ray crystal structure of the membrane protein prostaglandin H2 synthase-1. *Nature* **1994,** *367*, 243-249.

10. Hohlfeld, T.; Zimmermann, N.; Weber, A. A.; Jessen, G.; Weber, H.; Schror, K.; Holtje, H. D.; Ebel, R., Pyrazolinone analgesics prevent the antiplatelet effect of aspirin and preserve human platelet thromboxane synthesis. *J. Thromb. Haemost.* **2008,** *6*, 166-173.

11. Lin, R.; Chiu, G.; Yu, Y.; Connolly, P. J.; Li, S.; Lu, Y.; Adams, M.; Fuentes-Pesquera, A. R.; Emanuel, S. L.; Greenberger, L. M., Design, synthesis, and evaluation of 3,4-disubstituted pyrazole analogues as anti-tumor CDK inhibitors. *Bioorg. Med. Chem. Lett.* **2007,** *17*, 4557-4561.

12. Ghorab, M. M.; El-Gazzar, M. G.; Alsaid, M. S., Synthesis, characterization and anti-breast cancer activity of new 4-aminoantipyrine-based heterocycles. *Int. J. Mol. Sci.* **2014,** *15*, 7539-53.

13. Arshad, N.; Ahmad, M.; Ashraf, M. Z.; Nadeem, H., Spectroscopic, electrochemical DNA binding and in vivo anti-inflammatory studies on newly synthesized Schiff bases of 4-aminophenazone. *J. Photochem. Photobiol.* **2014,** *138*, 331-346.

14. Anupama, B.; Sunita, M.; Shiva Leela, D.; Ushaiah, B.; Gyana Kumari, C., Synthesis, spectral characterization, DNA binding studies and antimicrobial activity of Co(II), Ni(II), Zn(II), Fe(III) and VO(IV) complexes with 4-aminoantipyrine Schiff base of ortho-vanillin. *J. Fluoresc.* **2014,** *24*, 1067-76.

15. Farino, J.; Norwitz, G.; Boyko, W. J.; Keliher, P. N., Study of the behaviour of various phenolic compounds in the 4-aminoantipyrine and ultraviolet-ratio spectrophotometric methods without and with distillation. *Talanta* **1981,** *28*, 705-708.

16. Kim, H. Y.; Lee, H. J.; Chang, S. K., Reaction-based colorimetric signaling of Cu(2+) ions by oxidative coupling of phenols with 4-aminoantipyrine. *Talanta* **2015,** *132*, 625-629.

17. Asp, N. G., Improved method for the assay of phenylglycosidase activity with a 4-aminoantipyrine reagent. *Anal. Biochem.* **1971,** *40*, 281-286.

18. Zeng, Z.; Tian, L.; Li, Z.; Jia, L.; Zhang, X.; Xia, M.; Hu, Y., Whole-cell method for phenol detection based on the color reaction of phenol with 4-aminoantipyrine catalyzed by CotA laccase on endospore surfaces. *Biosens. Bioelectron.* **2015,** *69*, 162-166.

19. Giordanetto, F.; Tyrchan, C.; Ulander, J., Intramolecular hydrogen bond expectations in medicinal chemistry. *ACS Med. Chem. Lett.* **2017,** *8*, 139-142.

20. Arbačiauskienė, E.; Krikštolaitytė, S.; Mitrulevičienė, A.; Bieliauskas, A.; Martynaitis, V.; Bechmann, M.; Roller, A.; Šačkus, A.; Holzer, W., On the Tautomerism of N-Substituted Pyrazolones: 1, 2-Dihydro-3H-pyrazol-3-ones versus 1H-Pyrazol-3-ols. *Molecules* **2018,** *23*, 129.

21. Schrader, T.; Kirsten, C., Intermolecular stabilisation of the β-sheet conformation in dipeptides. *Chem. Commun.* **1996**, 2089-2090.

22. Kirsten, C. N.; Schrader, T. H., Intermolecular β-sheet stabilization with aminopyrazoles. *J. Am. Chem. Soc.* **1997,** *119*, 12061-12068.

23. Hellmert, M.; Müller-Schiffmann, A.; Peters, M. S.; Korth, C.; Schrader, T., Hybridization of an Aβ-specific antibody fragment with aminopyrazole-based β-sheet ligands displays striking enhancement of target affinity. *Org. Biomol. Chem.* **2015,** *13*, 2974-2979.

24. Gilfillan, L.; Artschwager, R.; Harkiss, A. H.; Liskamp, R. M.; Sutherland, A., Synthesis of pyrazole containing α-amino acids via a highly regioselective condensation/aza-Michael reaction of β-aryl α, β-unsaturated ketones. *Org. Biomol. Chem.* **2015,** *13*, 4514-4523.

25. Giri, R. S.; Mandal, B., Boc-Val-Val-OMe (Aβ 39–40) and Boc-Ile-Ala-OMe (Aβ 41–42) crystallize in a parallel β-sheet arrangement but generate a different morphology. *CrystEngComm* **2018,** *20*, 4441-4448.

26. Balachandra, C.; Sharma, N. K., Synthesis and conformational analysis of new troponyl aromatic amino acid. *Tetrahedron* **2014,** *70*, 7464-7469.

27. Chatterjee, S.; Maitra, U., Hierarchical self-assembly of photoluminescent CdS nanoparticles into a bile acid derived organogel: morphological and photophysical properties. *Phys. Chem. Chem. Phys.* **2017,** *19*, 17726-17734.

28. Bardelang, D.; Camerel, F.; Margeson, J. C.; Leek, D. M.; Schmutz, M.; Zaman, M. B.; Yu, K.; Soldatov, D. V.; Ziessel, R.; Ratcliffe, C. I., Unusual sculpting of dipeptide particles by ultrasound induces gelation. *J. Am. Chem. Soc.* **2008,** *130*, 3313-3315.

29. Maity, S.; Das, P.; Reches, M., Inversion of supramolecular chirality by sonication-induced organogelation. *Sci. Rep.* **2015,** *5*, 16365.

30. Maity, S.; Kumar, P.; Haldar, D., Sonication-induced instant amyloid-like fibril formation and organogelation by a tripeptide. *Soft Matter* **2011,** *7*, 5239-5245.

31. Cravotto, G.; Cintas, P., Molecular self-assembly and patterning induced by sound waves. The case of gelation. *Chem. Soc. Rev.* **2009,** *38*, 2684-2697.

32. Li, Y.; Wang, T.; Liu, M., Ultrasound induced formation of organogel from a glutamic dendron. *Tetrahedron* **2007,** *63*, 7468-7473.

33. Deber, M. P. L.; Hickman, D. T.; Nand, D.; Baldus, M.; Pfeifer, A.; Muhs, A., Engineering amyloid-like assemblies from unstructured peptides via site-specific lipid conjugation. *PloS one* **2014,** *9*, e105641.

34. Cerpa, R.; Cohen, F. E.; Kuntz, I. D., Conformational switching in designed peptides: the helix/sheet transition. *Fold. Des.* **1996,** *1*, 91-101.

35. Awasthi, S. K.; Shankaramma, S.; Raghothama, S.; Balaram, P., Solvent-induced β-hairpin to helix conformational transition in a designed peptide. *Biopolymers* **2001,** *58*, 465-476.

36. Sonnichsen, F.; Van Eyk, J.; Hodges, R.; Sykes, B., Effect of trifluoroethanol on protein secondary structure: an NMR and CD study using a synthetic actin peptide. *Biochemistry* **1992,** *31*, 8790-8798.

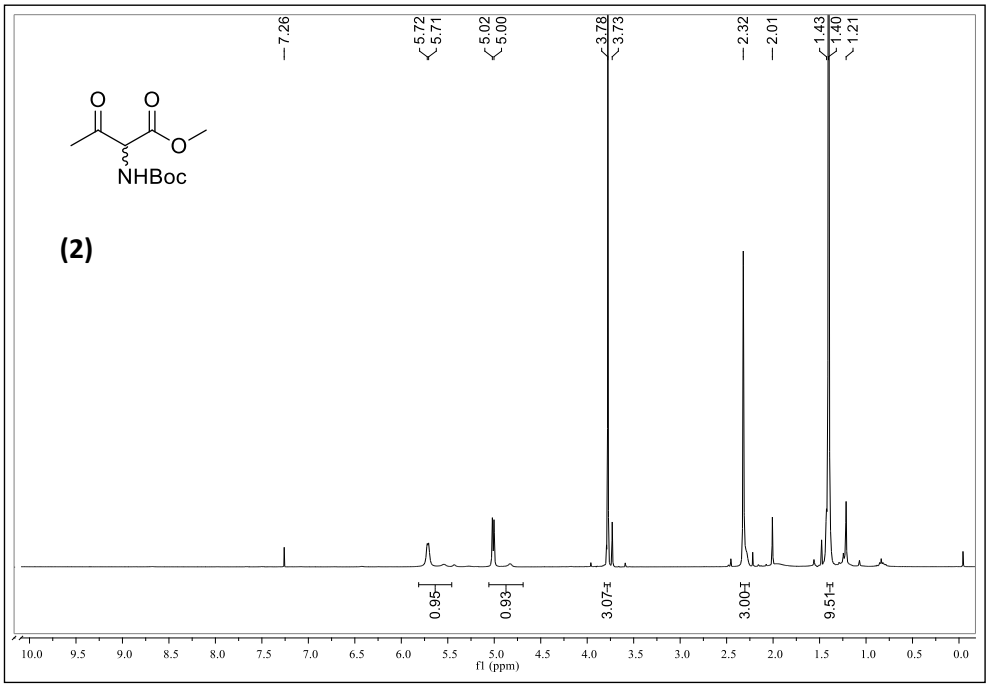
37. Vass, E.; Hollósi, M.; Besson, F.; Buchet, R., Vibrational spectroscopic detection of beta- and gamma-turns in synthetic and natural peptides and proteins. *Chem. Rev.* **2003,** *103*, 1917-1954.

38. Jackson, M.; Mantsch, H. H., The use and misuse of FTIR spectroscopy in the determination of protein structure. *Crit. Rev. Biochem. Mol. Biol.* **1995,** *30*, 95-120.

39. Yuran, S.; Razvag, Y.; Reches, M., Coassembly of aromatic dipeptides into biomolecular necklaces. *ACS Nano* **2012,** *6*, 9559-9566.

4A.7 Appendix

1. ^{1}H-/ ^{13}C spectra of (2)



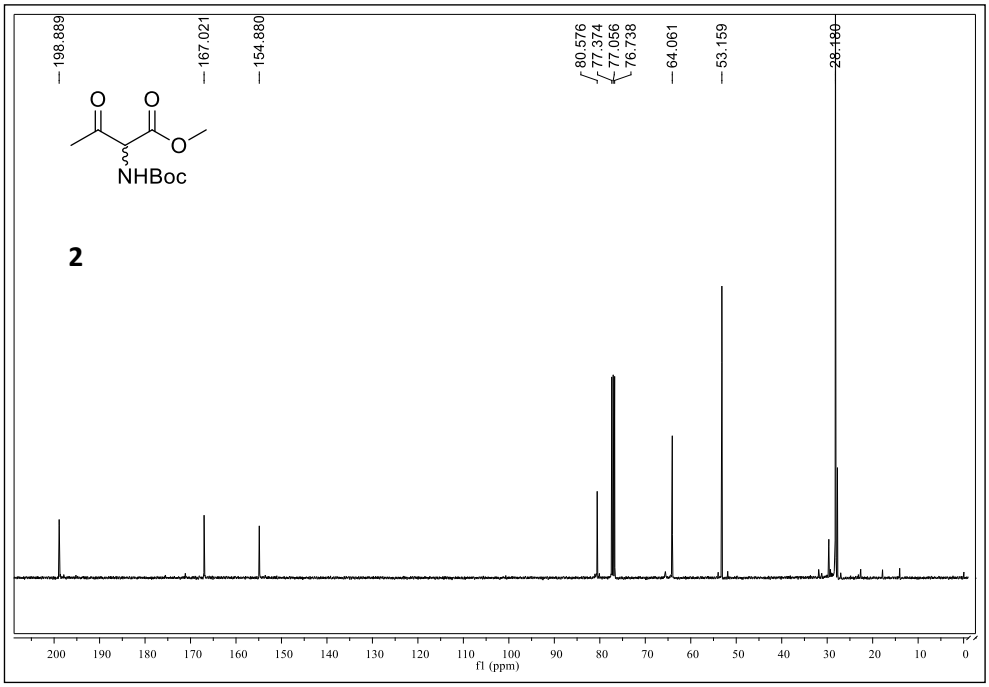
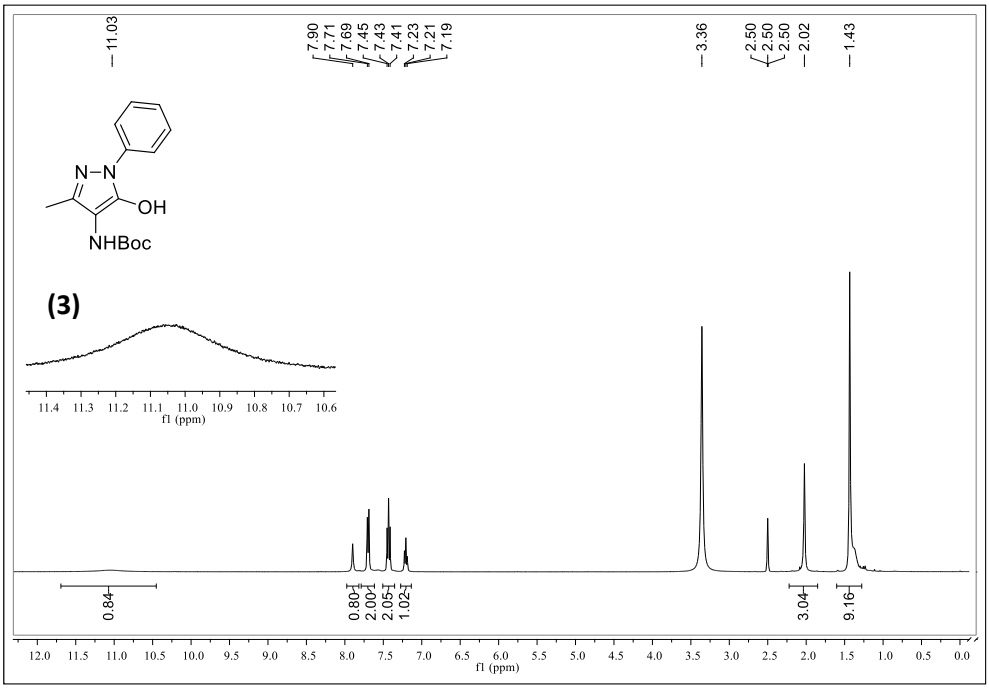


Figure S1. ^{1}H and ^{13}C NMR spectra of (**2**) in $CDCl_3$

2. ^{1}H-/ ^{13}C-/ESI-MS/HRMS spectra of (3)



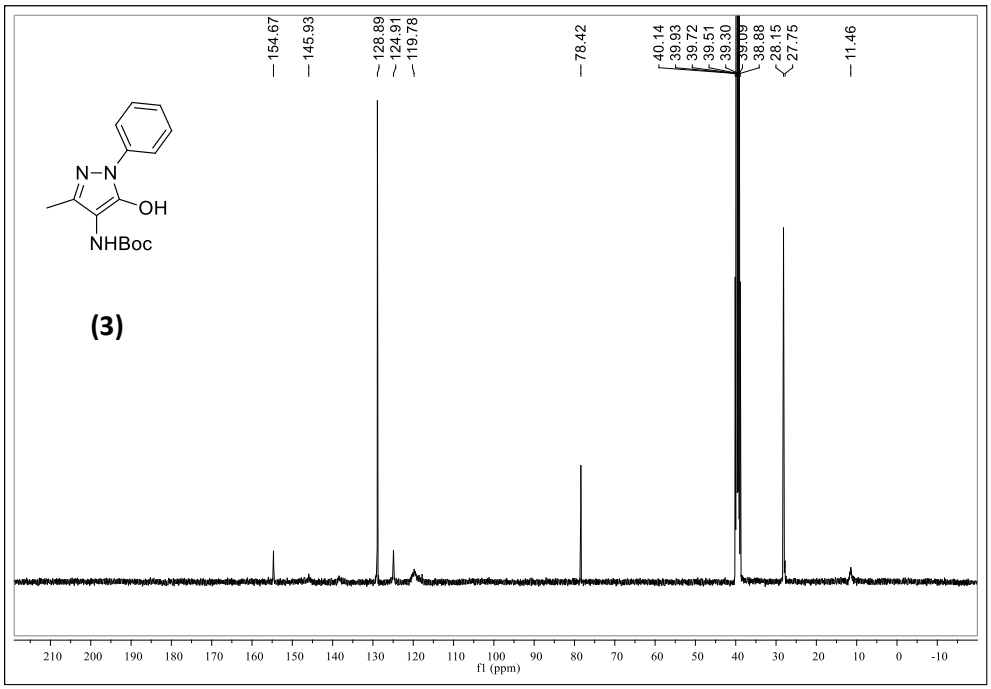


Figure S2. ^{1}H and ^{13}C NMR spectra of (**3**) in DMSO-d_6 (compound (**3**) can exist in tatautomeric forms it is well known that in polar solvent like DMSO-d_6 the enol form dominated).

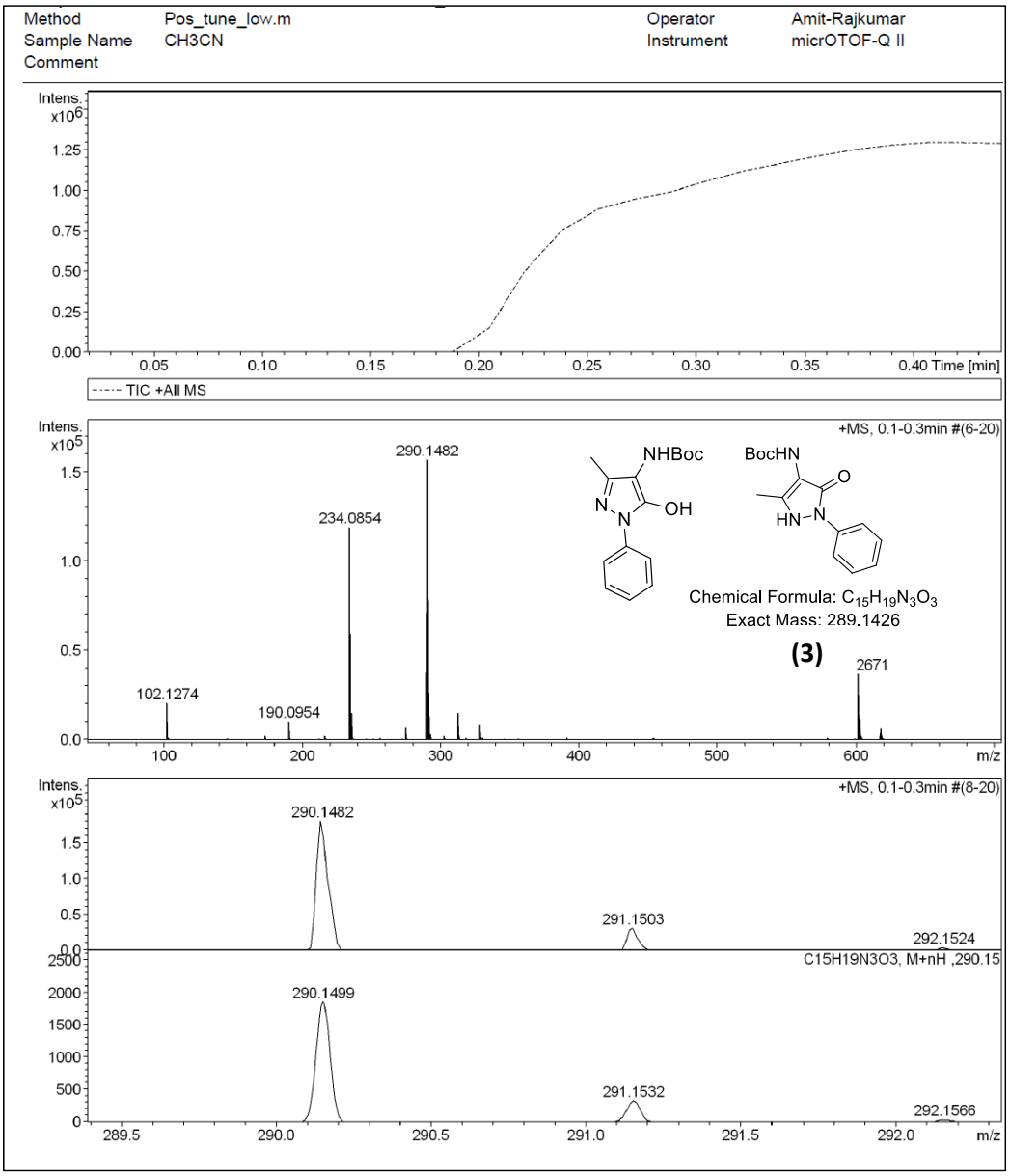


Figure S3. ESI-MS/HRMS spectra of (**3**) (compound (**3**) can exist in tatautomeric forms)

3.^{1}H-/^{13}C-/ ESI-MS/HRMS spectra of (4)

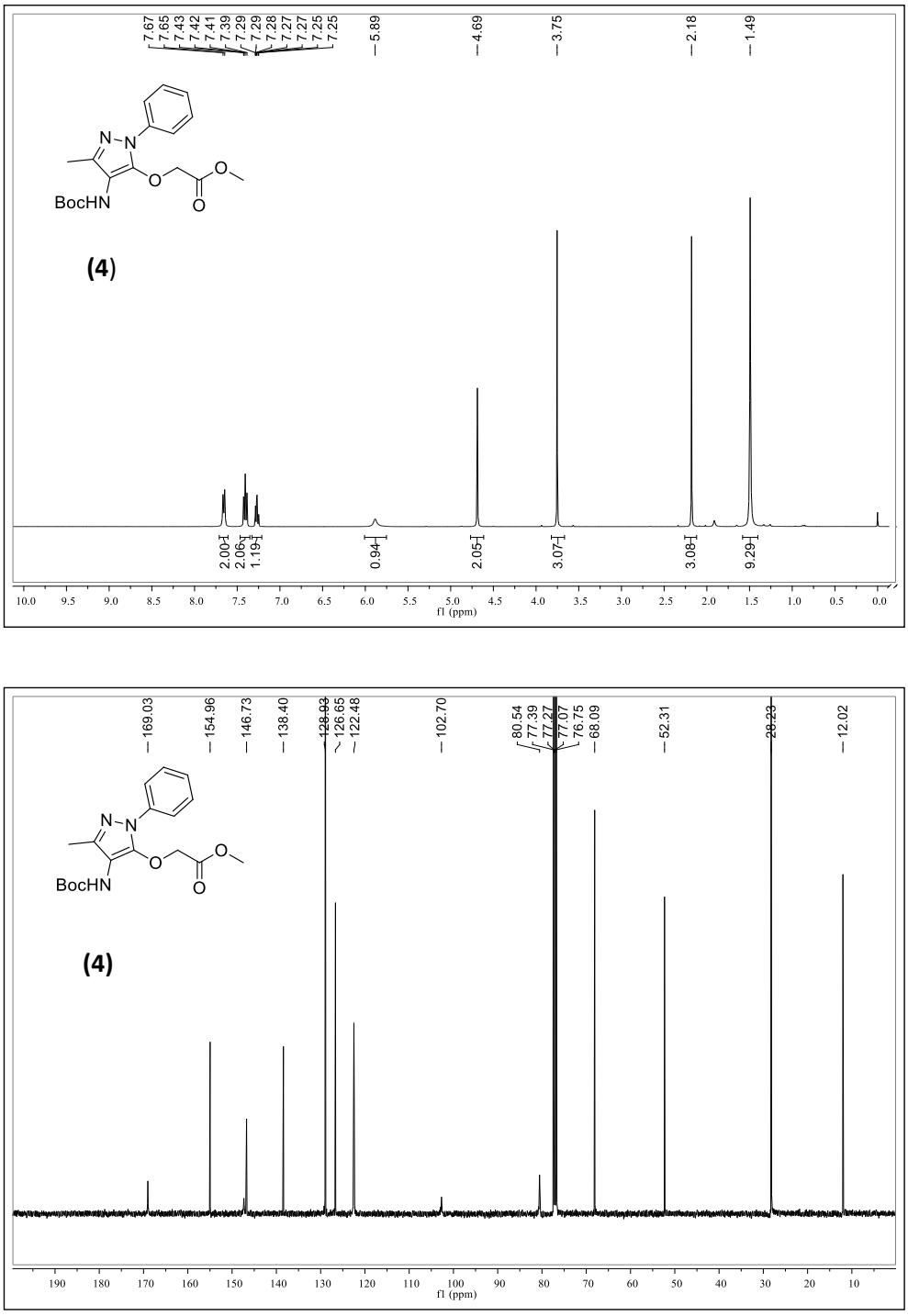


Figure S4. ^{1}H and ^{13}C NMR spectra of (**4**) in $CDCl_3$

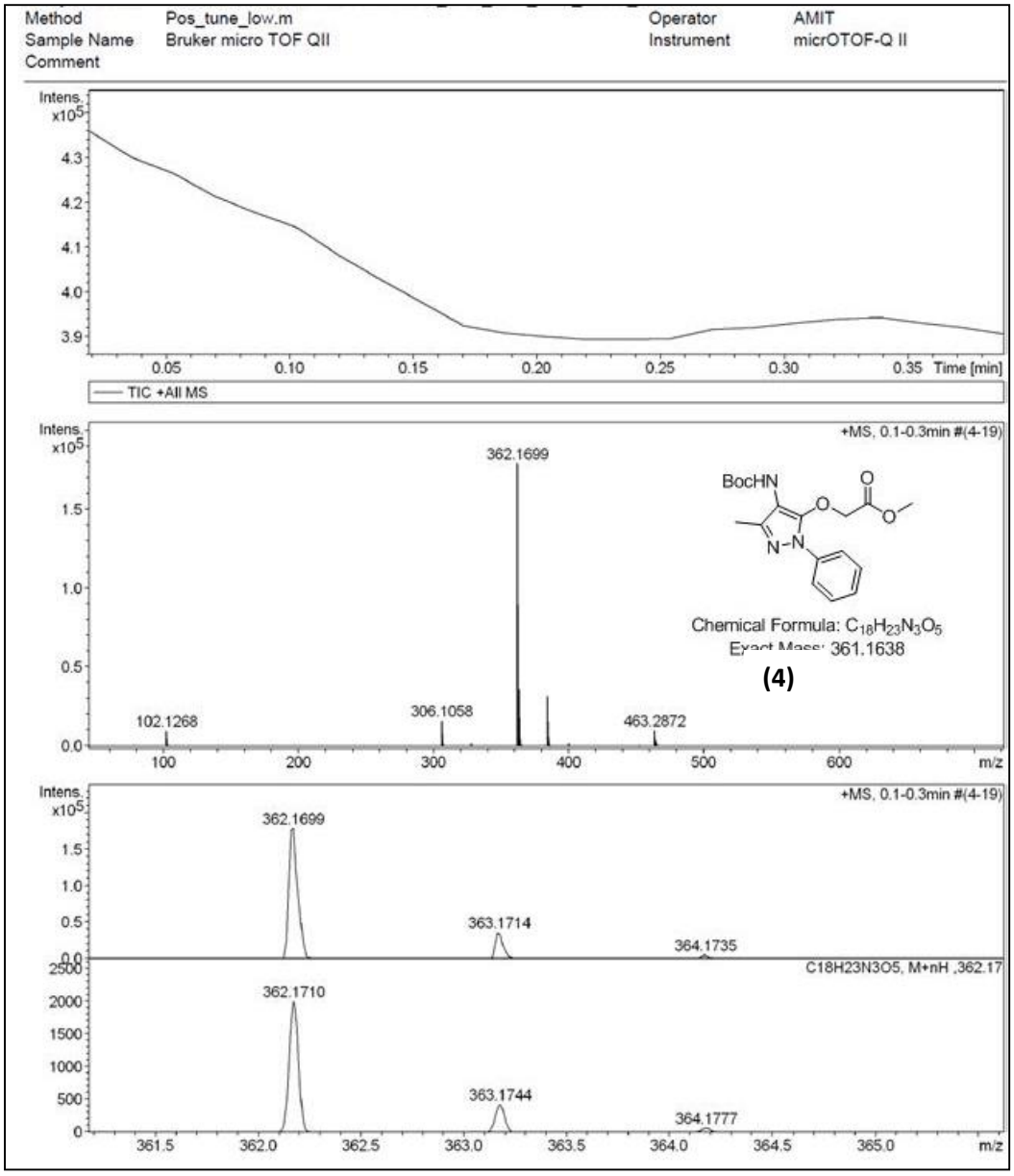


Figure S5. ESI-MS/HRMS spectra of (**4**)

4.^{1}H-/^{13}C-/ ESI-MS/HRMS spectra of (5)

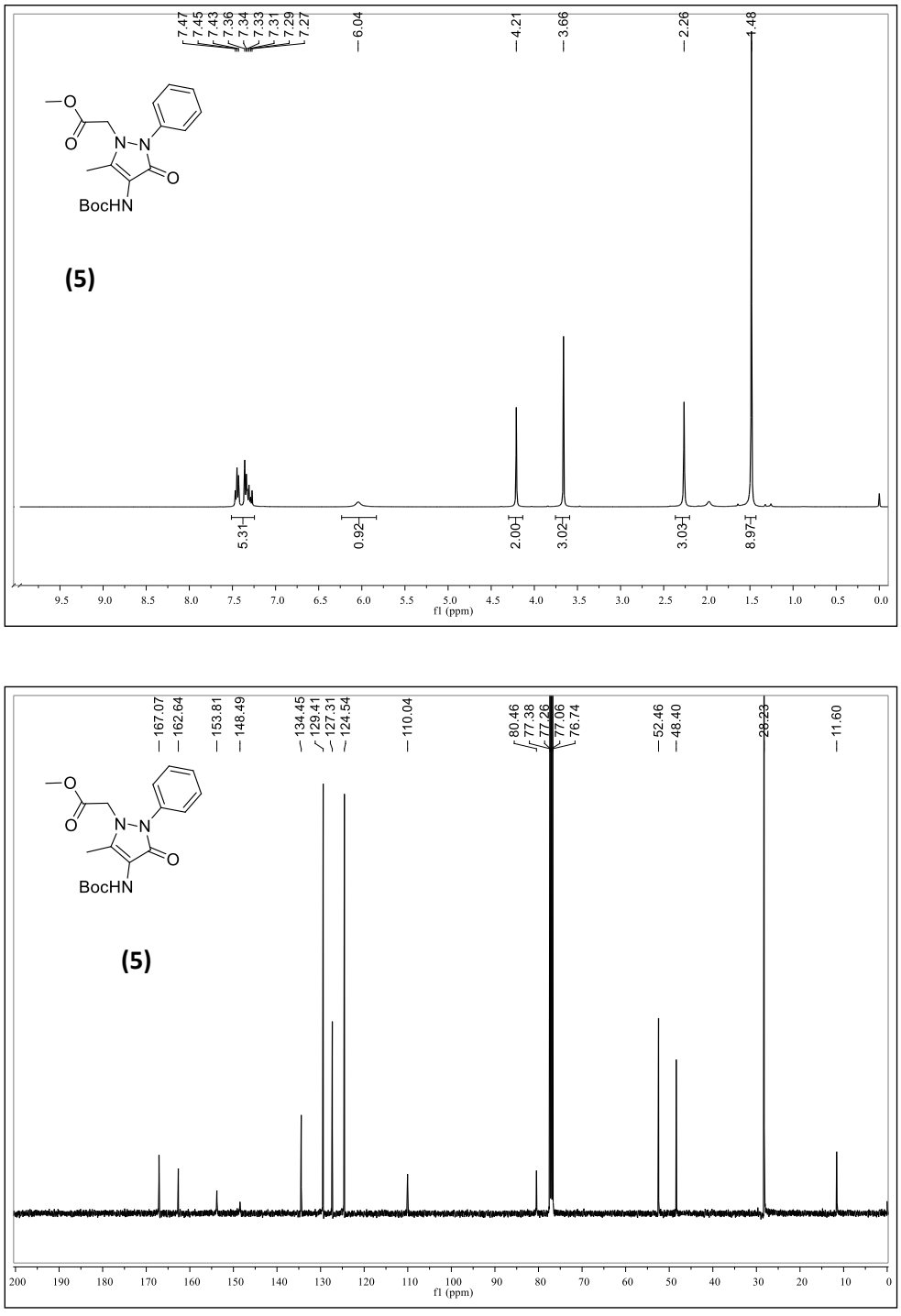


Figure S6. ^{1}H and ^{13}C NMR spectra of (**5**) in $CDCl_3$

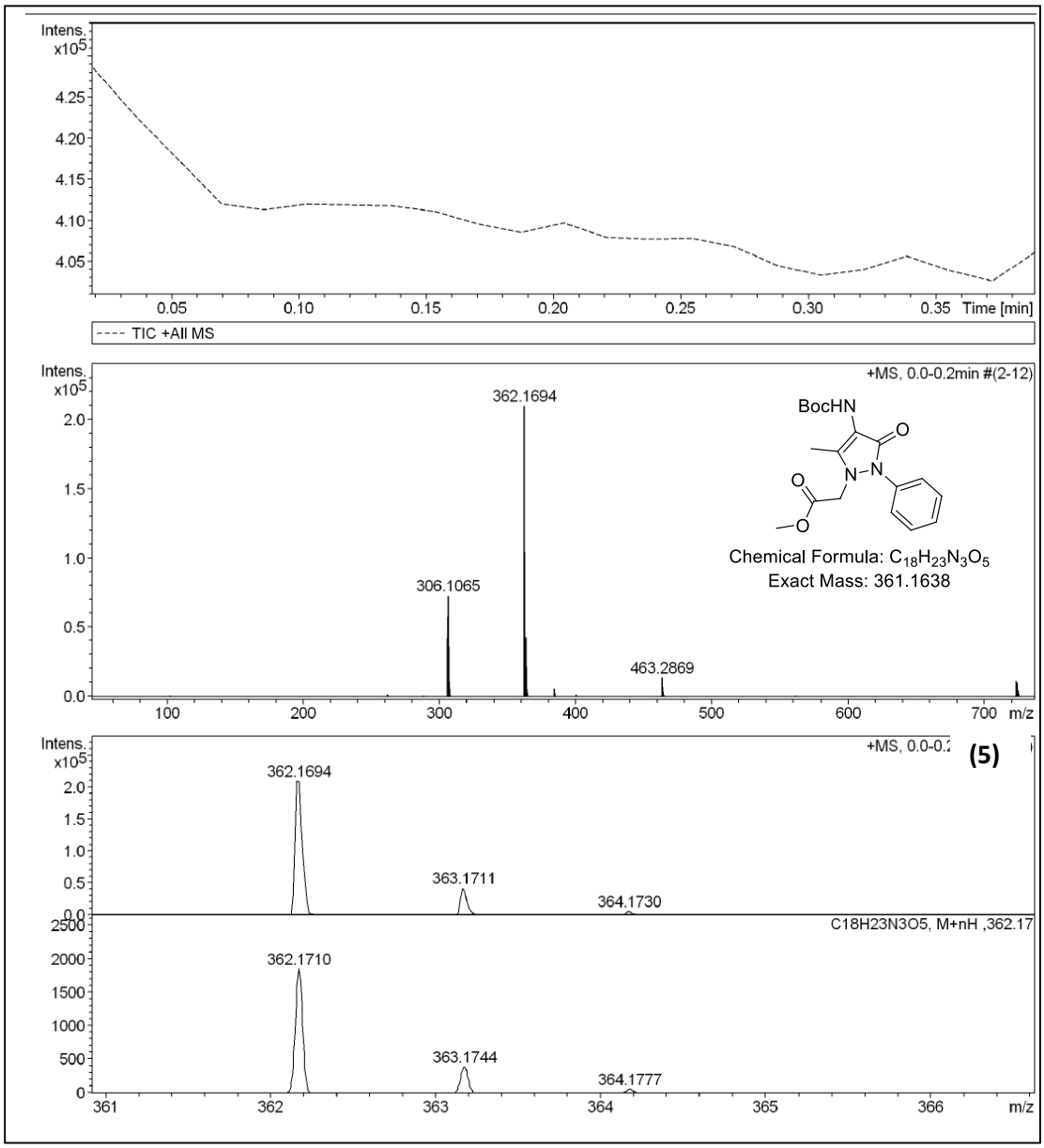


Figure S7. ESI-MS/HRMS spectra of (**5**)

5.^{1}H-/^{13}C-/ ESI-MS/HRMS spectra of (6)

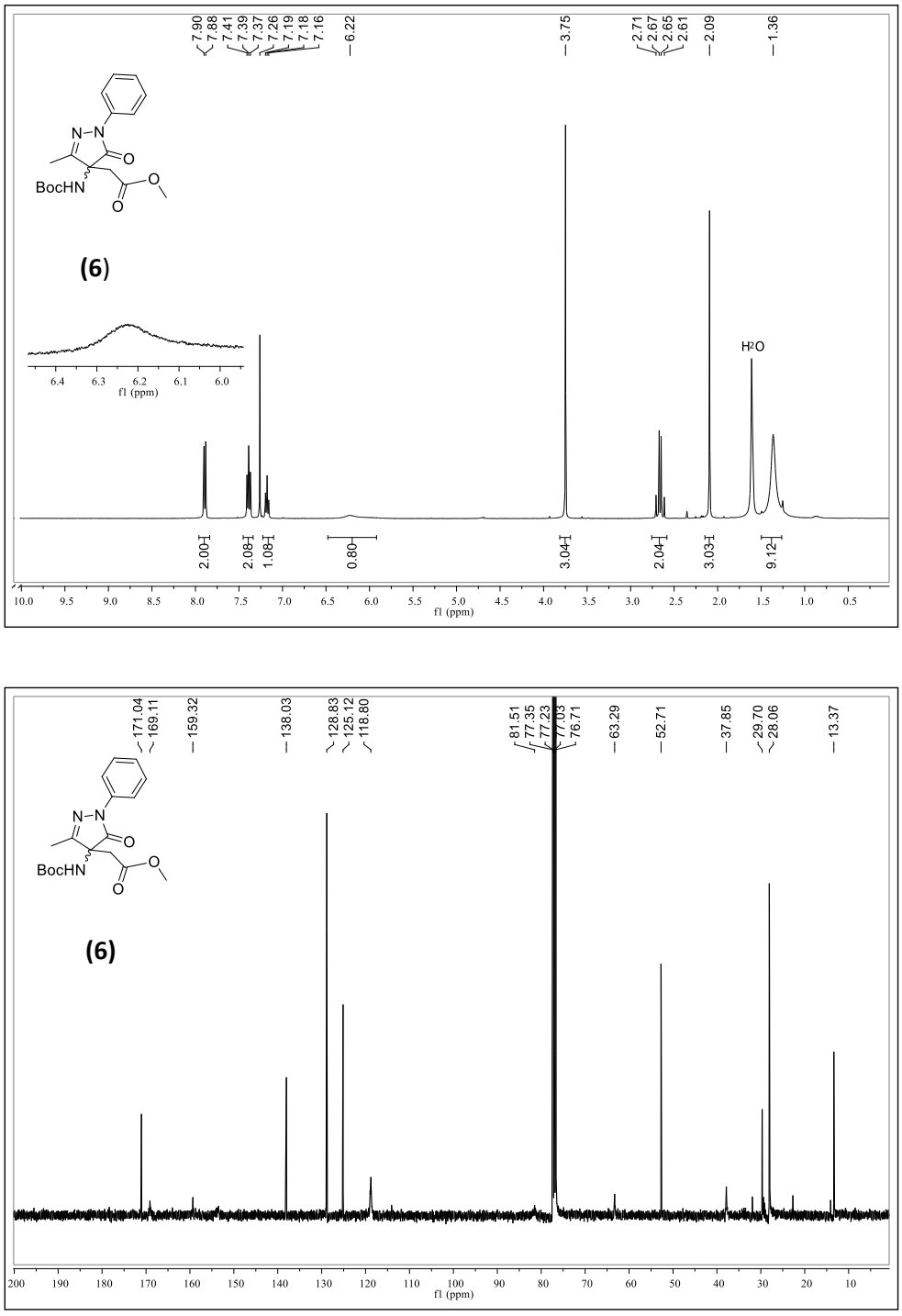


Figure S8. ^{1}H and ^{13}C NMR spectra of (**6**) in $CDCl_3$

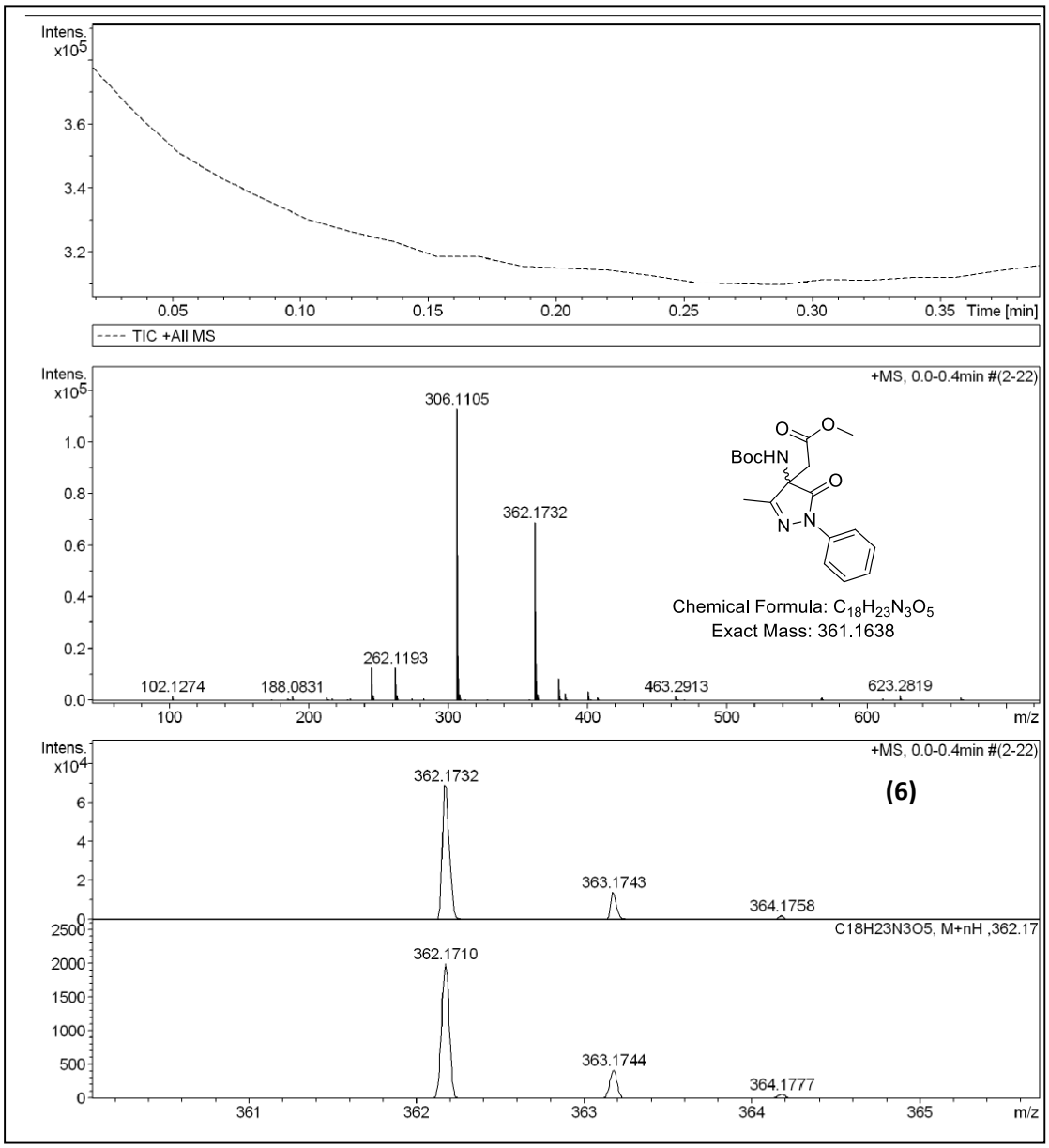


Figure S9. ESI-MS/HRMS spectra of (**6)**

6.Overlapped Proton NMR of alkylation

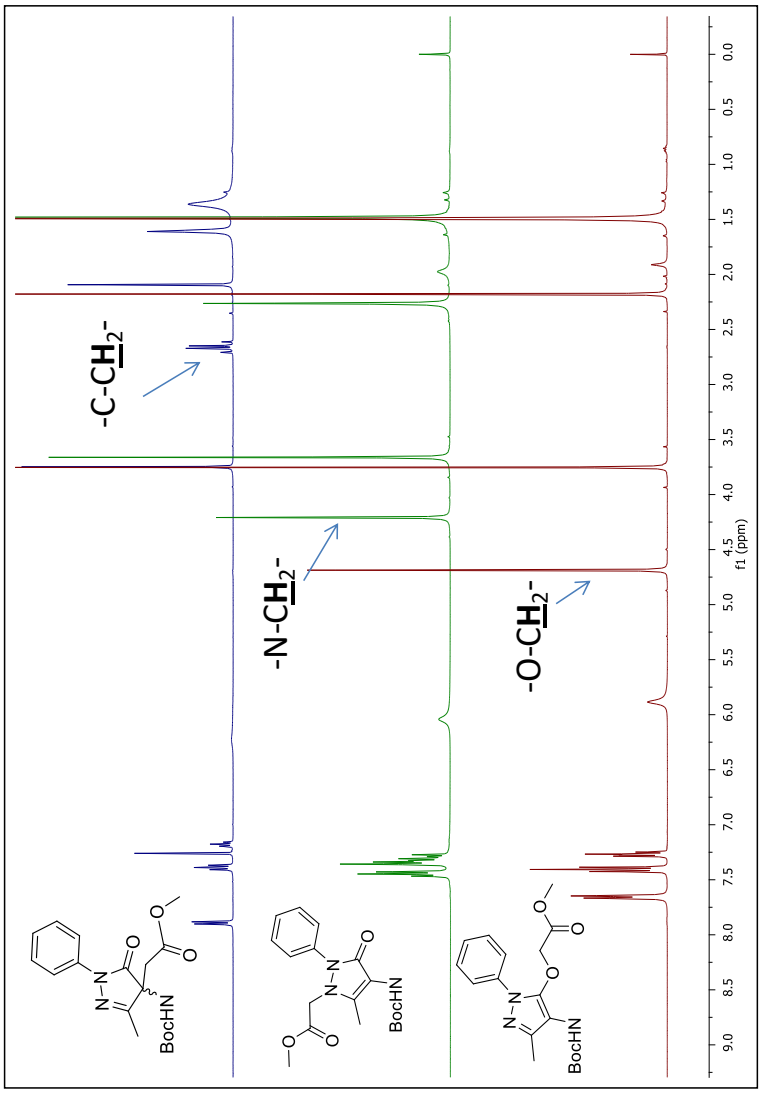


Figure S10. Overlapped ^{1}H NMR of (**4,5 & 6)**

7.^{1}H-/^{13}C-/ ESI-MS/HRMS spectra of (7a)

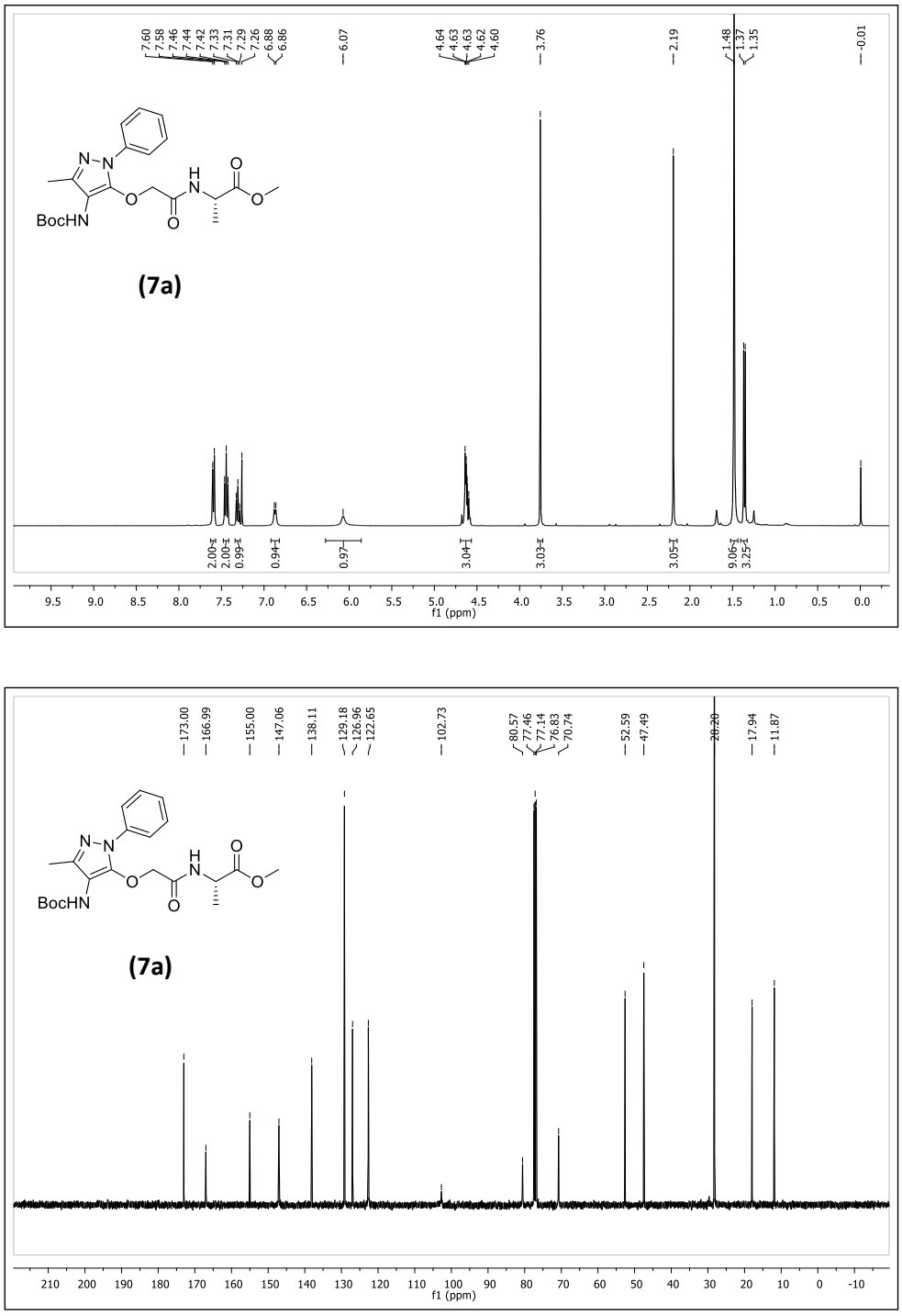


Figure S11. ^{1}H and ^{13}C NMR spectra of (**7a)** in $CDCl_3$

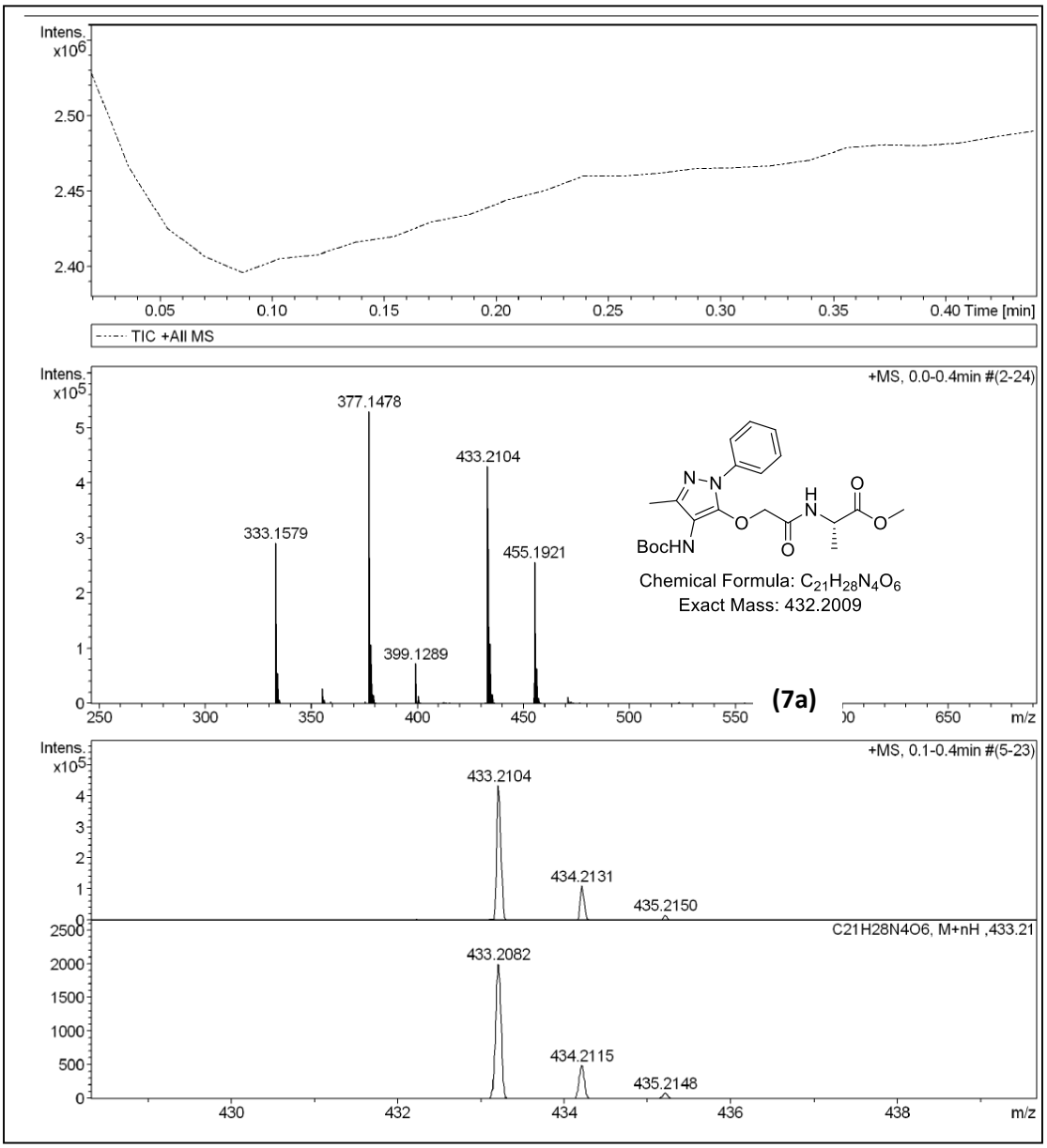


Figure S12. ESI-MS/HRMS spectra of (**7a)**

8.^{1}H-/^{13}C-/ ESI-MS/HRMS spectra of (7b)

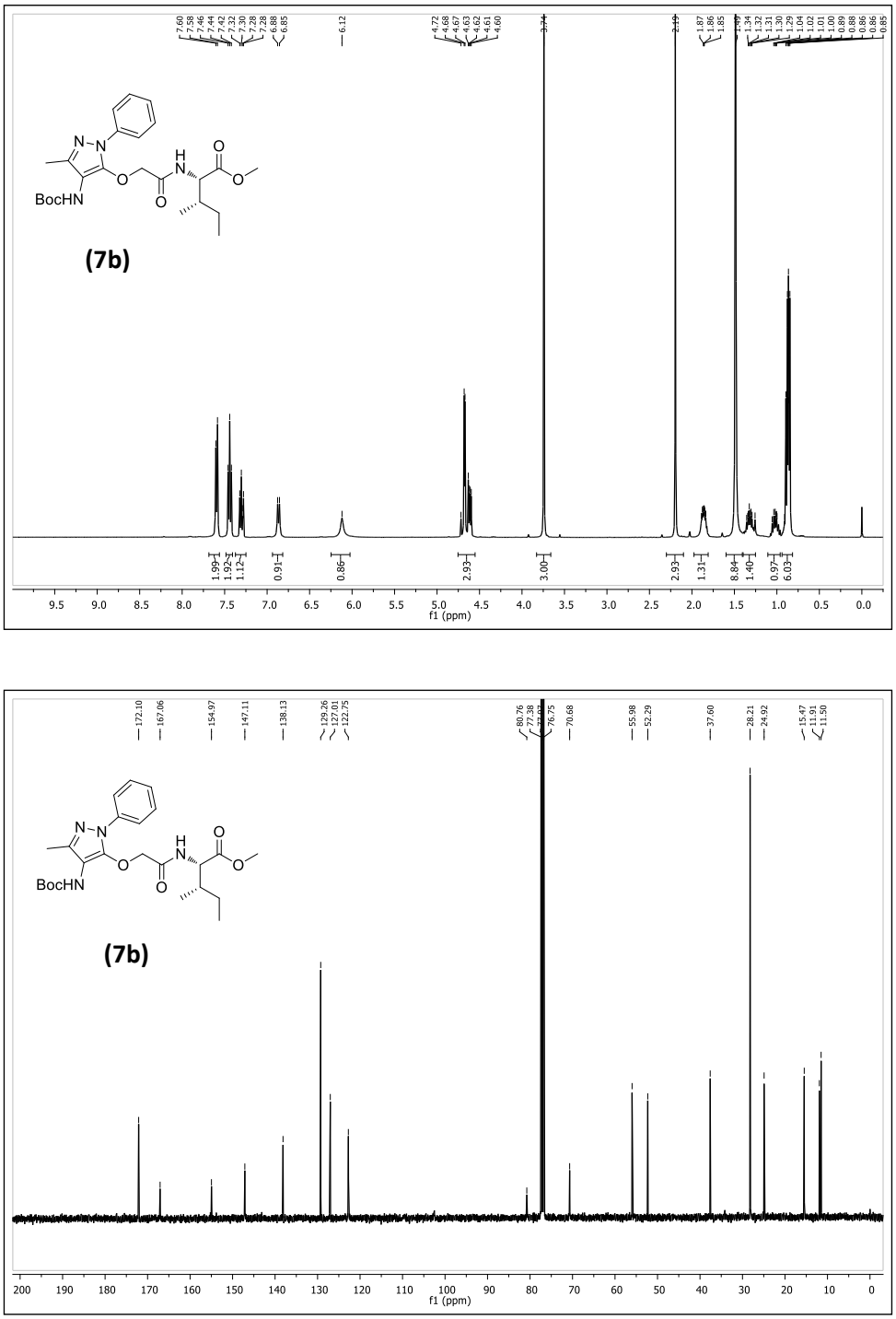


Figure 13. ^{1}H and ^{13}C NMR spectra of (**7b**) in $CDCl_3$

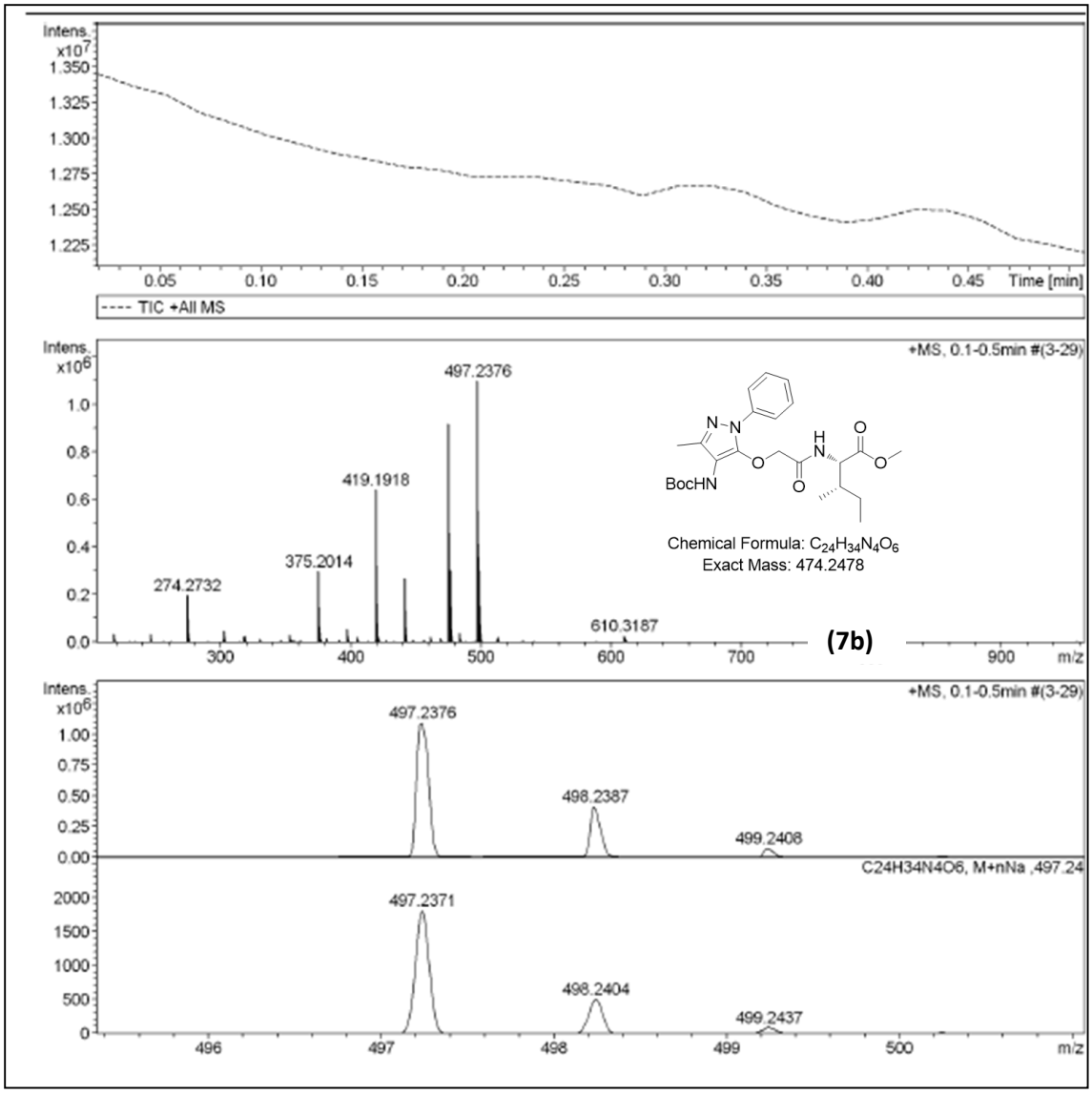
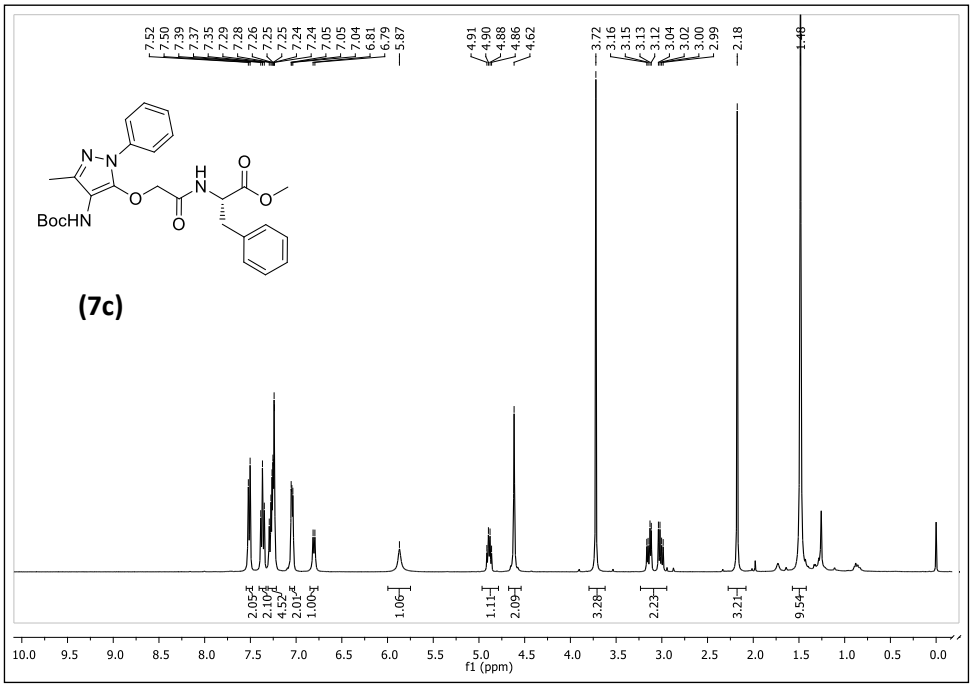


Figure S14. ESI-MS/HRMS spectra of (**7b)**

9.^{1}H-/^{13}C-/ COSY-/NOESY-/ESI-MS-/HRMS spectra of (7c)



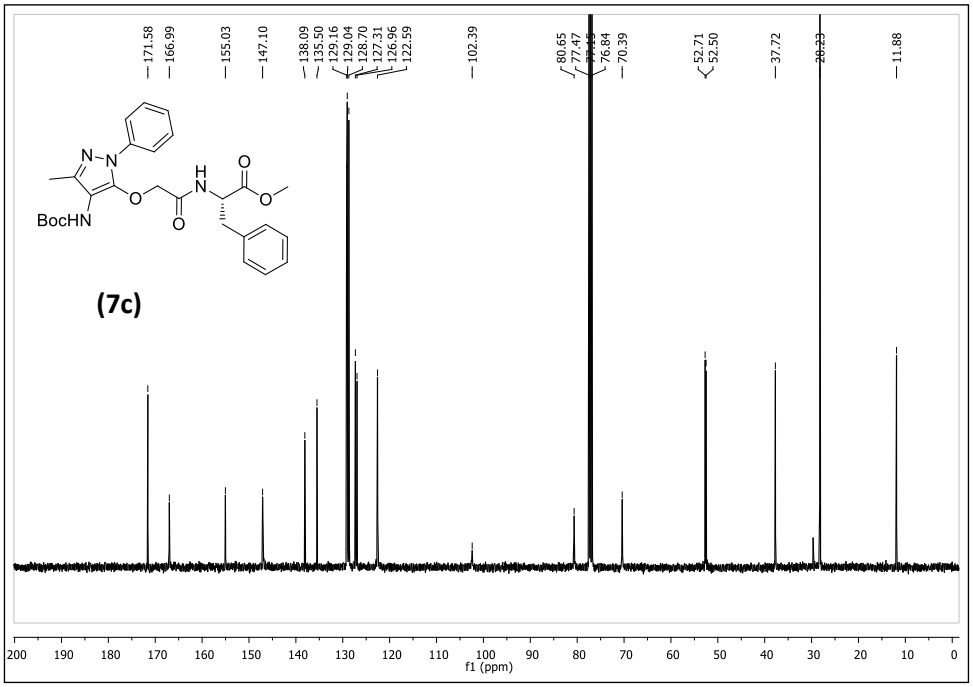


Figure S15. ^{1}H and ^{13}C NMR spectra of (**7c**) in $CDCl_3$

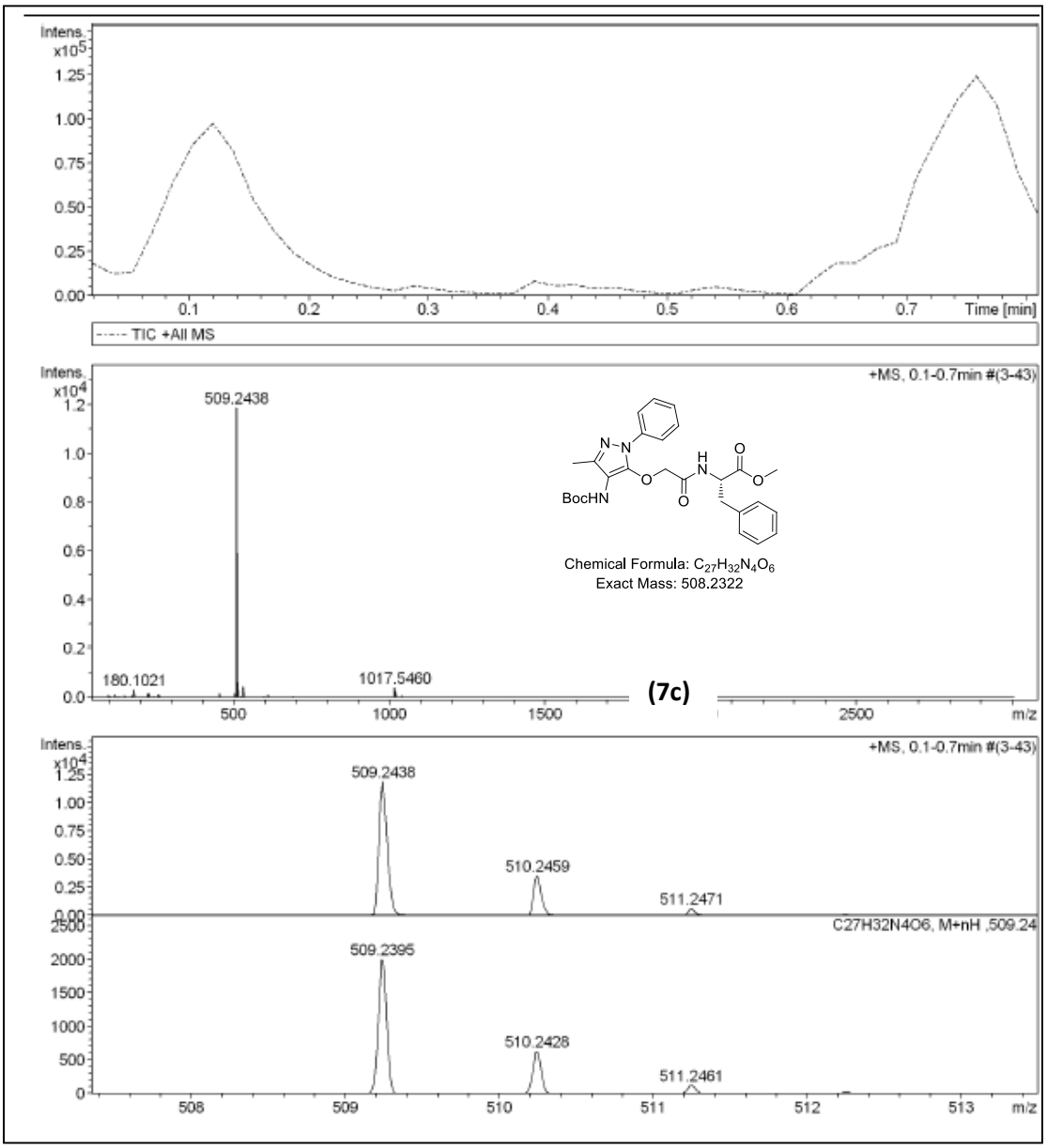


Figure S16. ESI-MS/HRMS spectra of (**7c**)

10.^{1}H-/^{13}C-/COSY-/NOESY-/ ESI-MS/HRMS spectra of (7d)

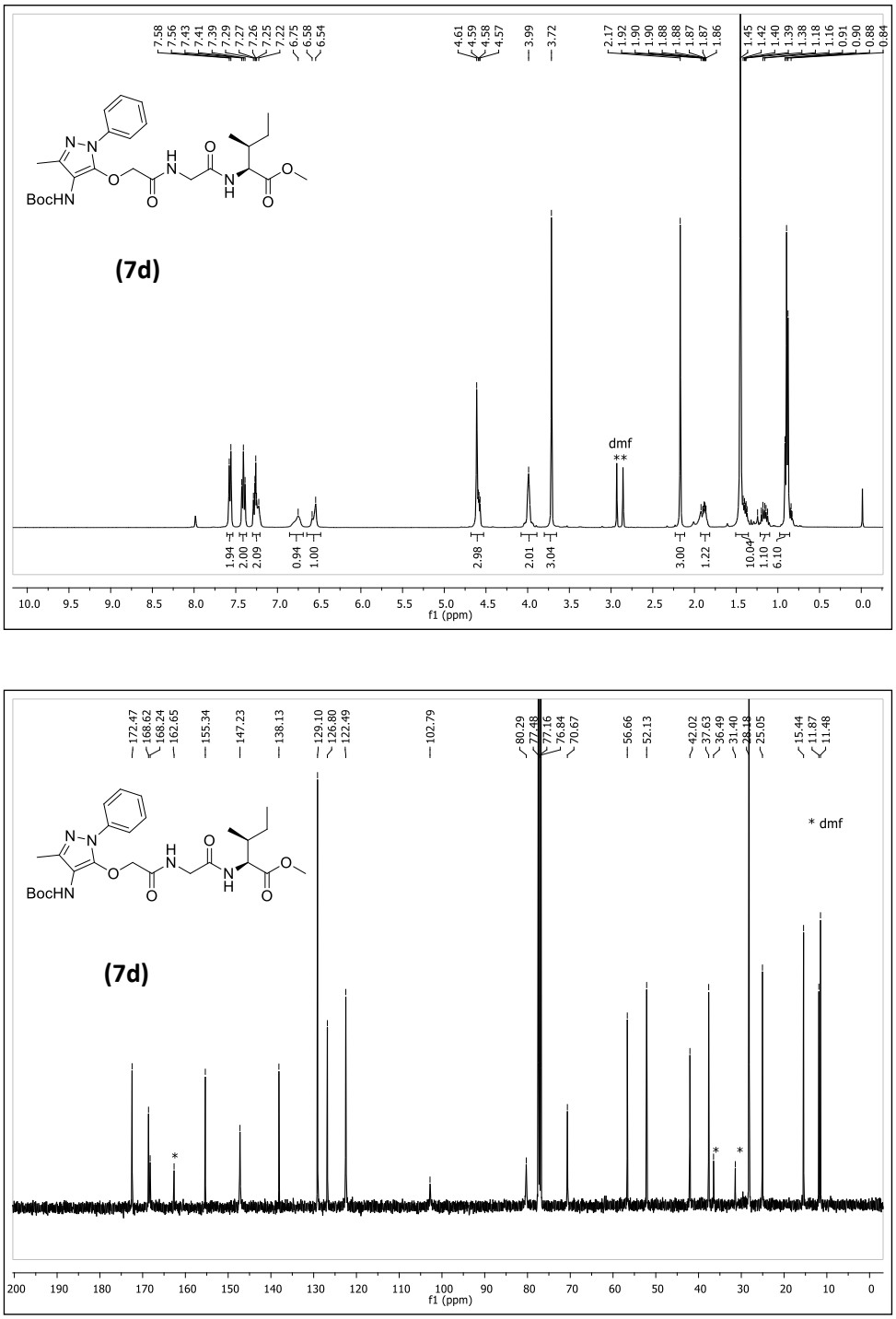


Figure 17. ^{1}H and ^{13}C NMR spectra of (**7d**) in $CDCl_3$

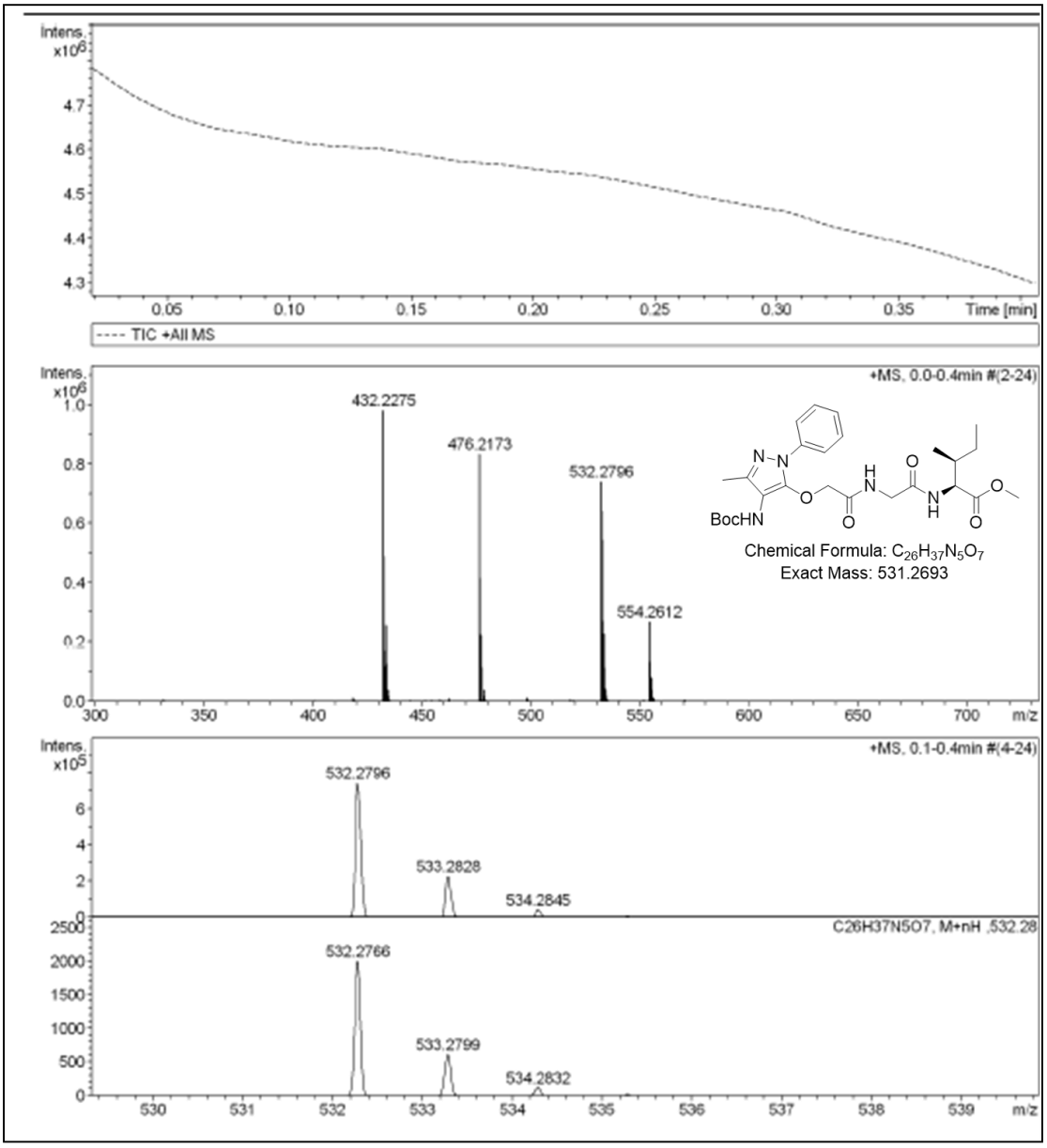


Figure S18. ESI-MS/HRMS spectra of (**7d**)

11.^{1}H-/^{13}C-/ ESI-MS/HRMS spectra of (7e)

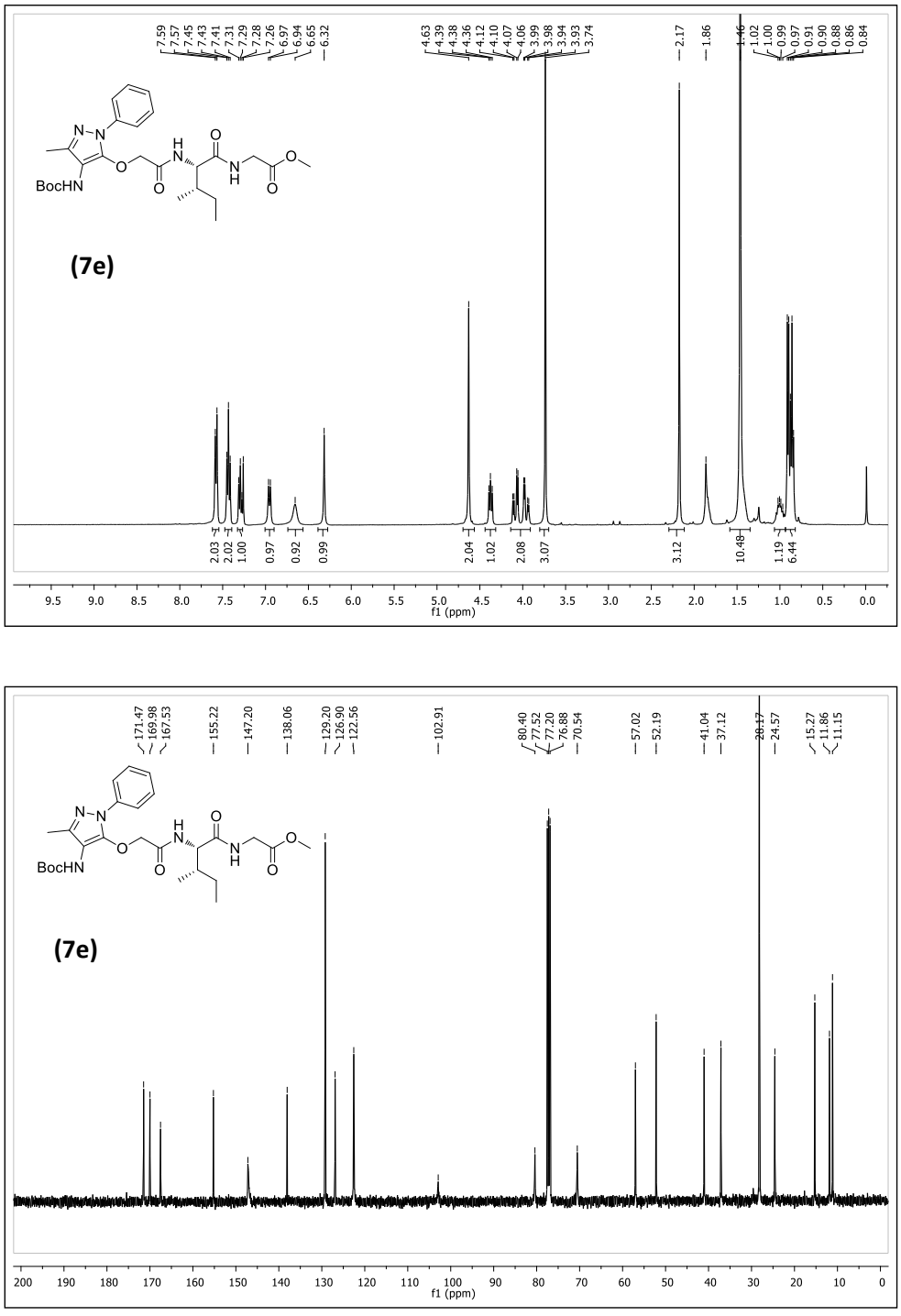


Figure S19. ^{1}H and ^{13}C NMR spectra of (**7e**) in $CDCl_3$

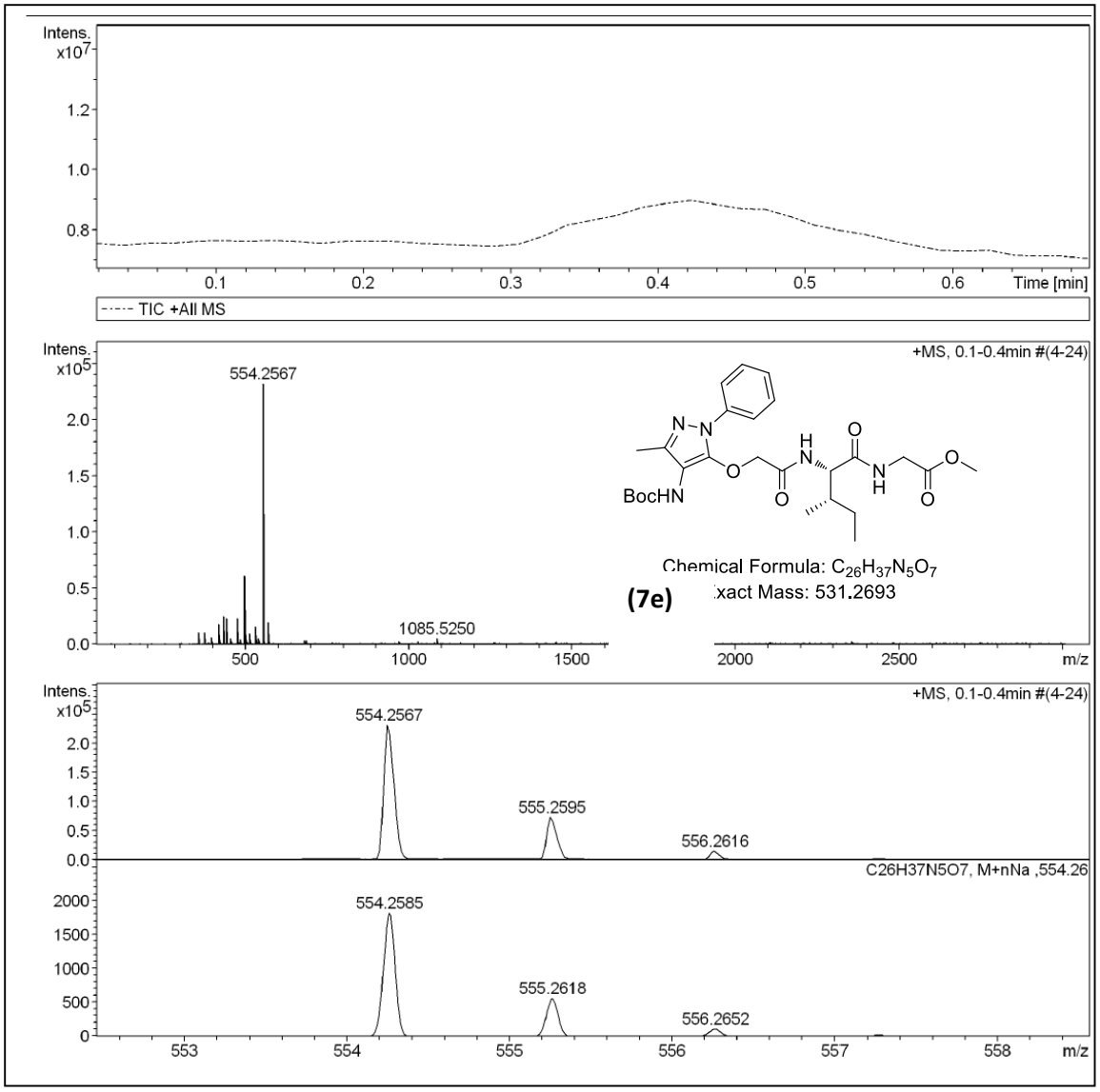
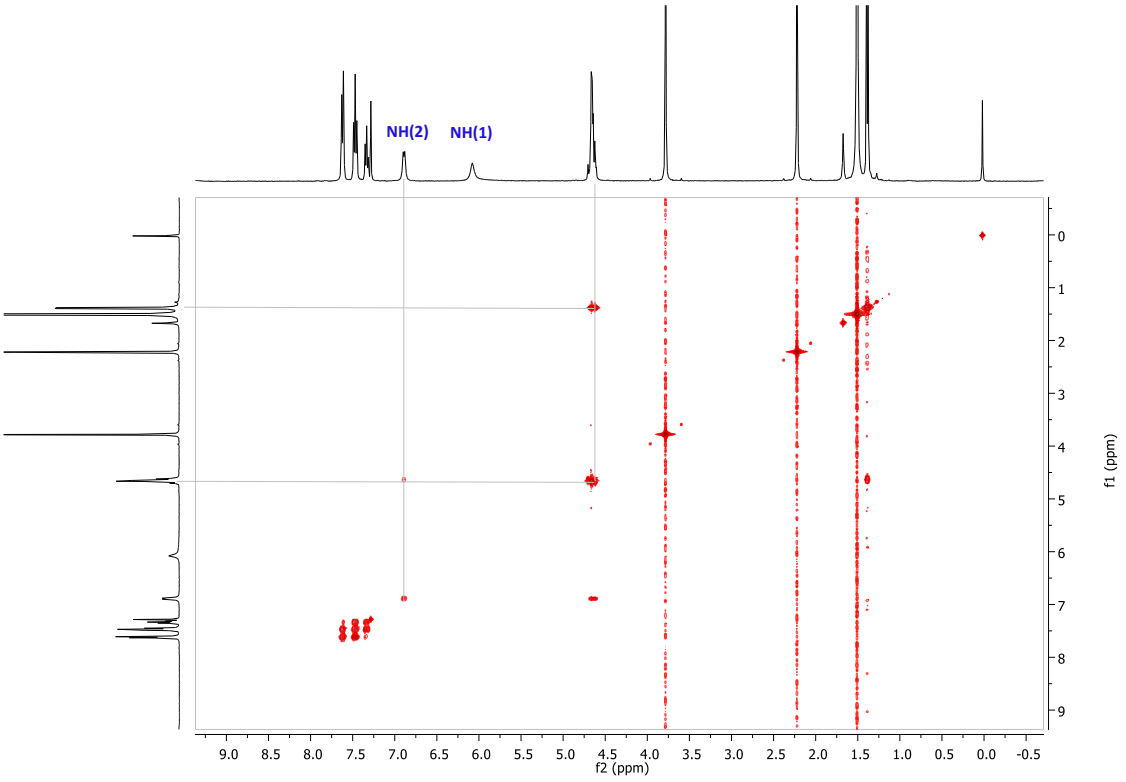


Figure S20. ESI-MS/HRMS spectra of (**7e**)

12.^{1}H-COSY & ^{1}H-NOESY spectra of (7a) in $CDCl_3$



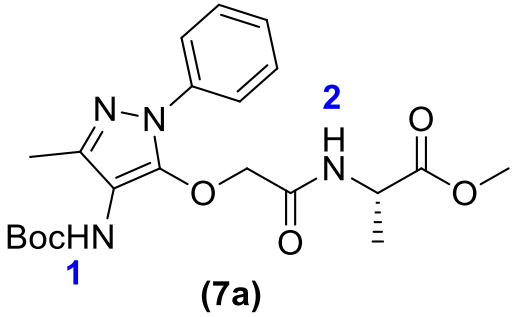


Figure S21. (1 of 3) ^{1}H-COSY spectra of (**7a**) in $CDCl_3$

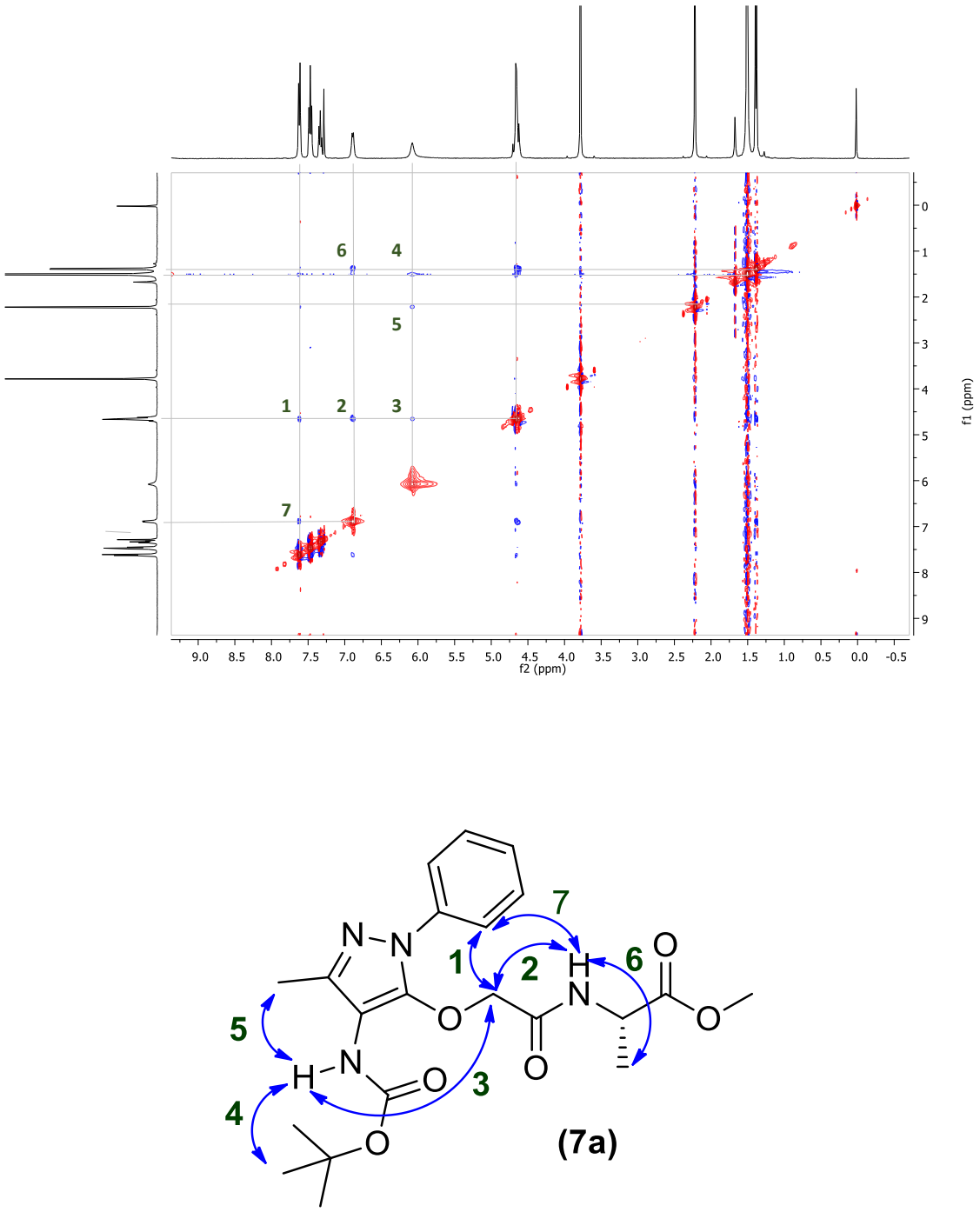


Figure S22. (2 of 3) ^{1}H-NOESY spectra of (**7a**) in $CDCl_3$

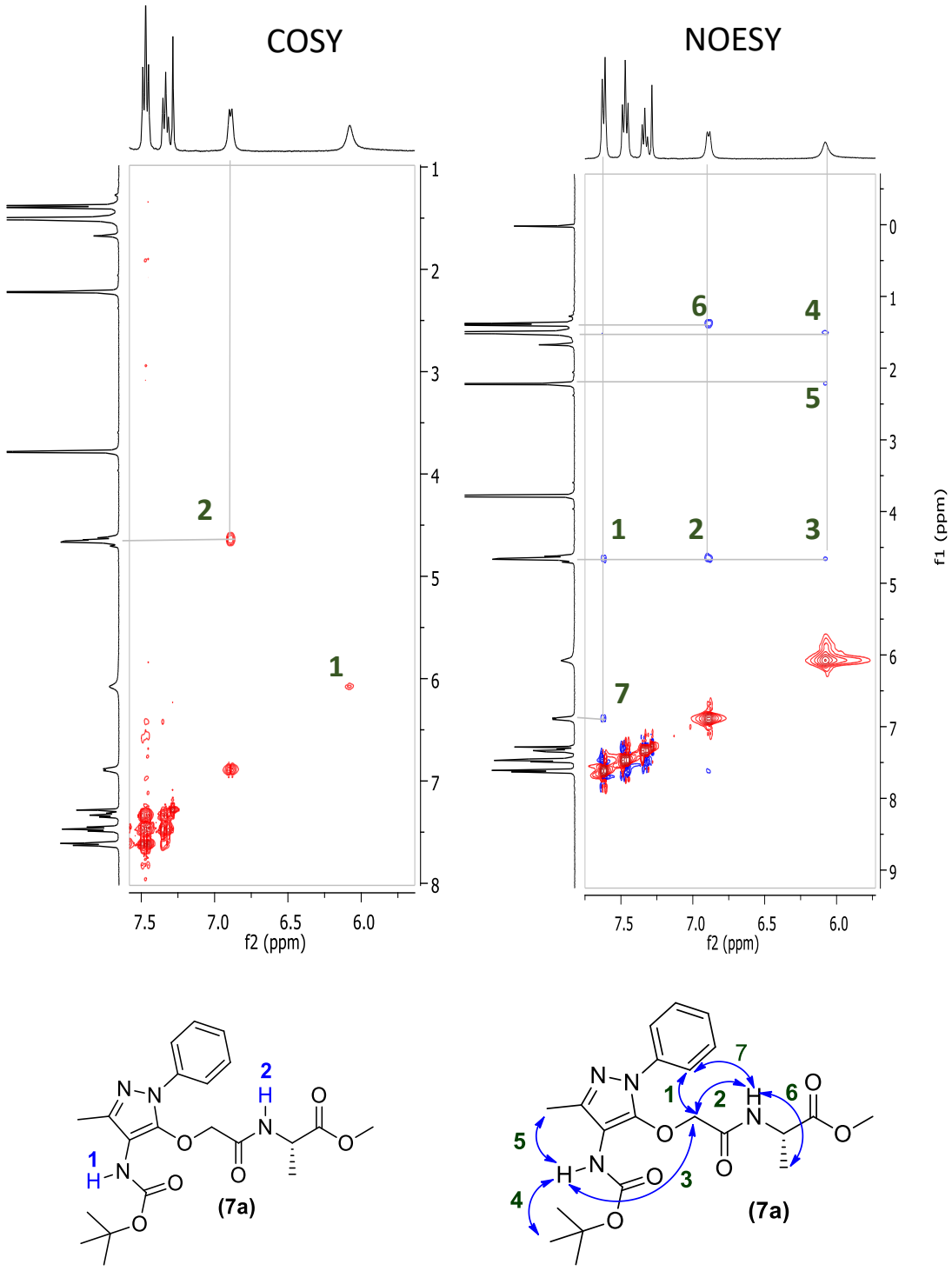


Figure S23. (3 of 3) ^{1}H-COSY & ^{1}H-NOESY expanded spectra in amide region of (**7a**) in $CDCl_3$.

13.^{1}H-COSY & ^{1}H-NOESY spectra of (7c) in $CDCl_3$

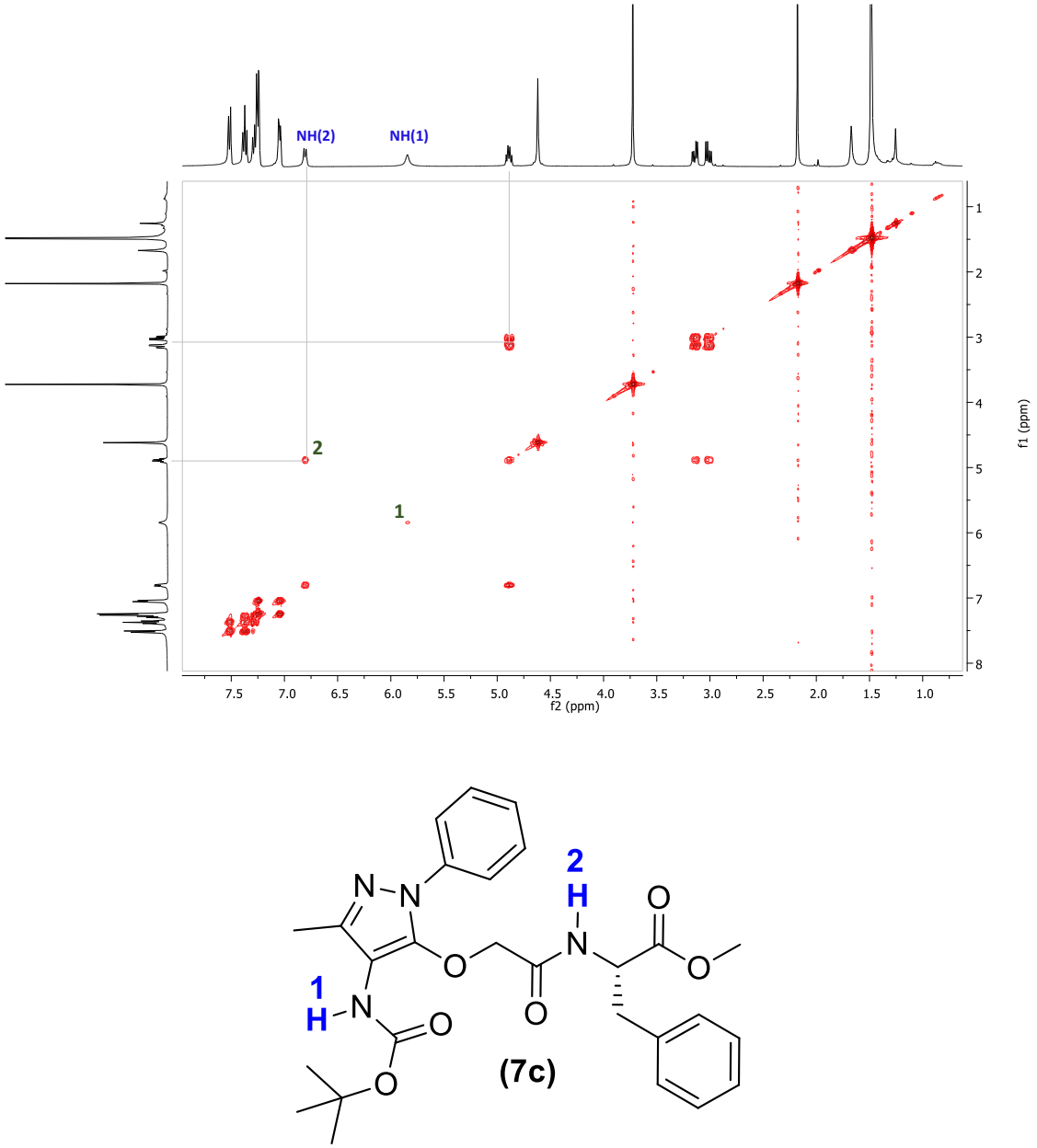


Figure S24. (1 of 3) ^{1}H-COSY spectra of (**7c**) in $CDCl_3$.

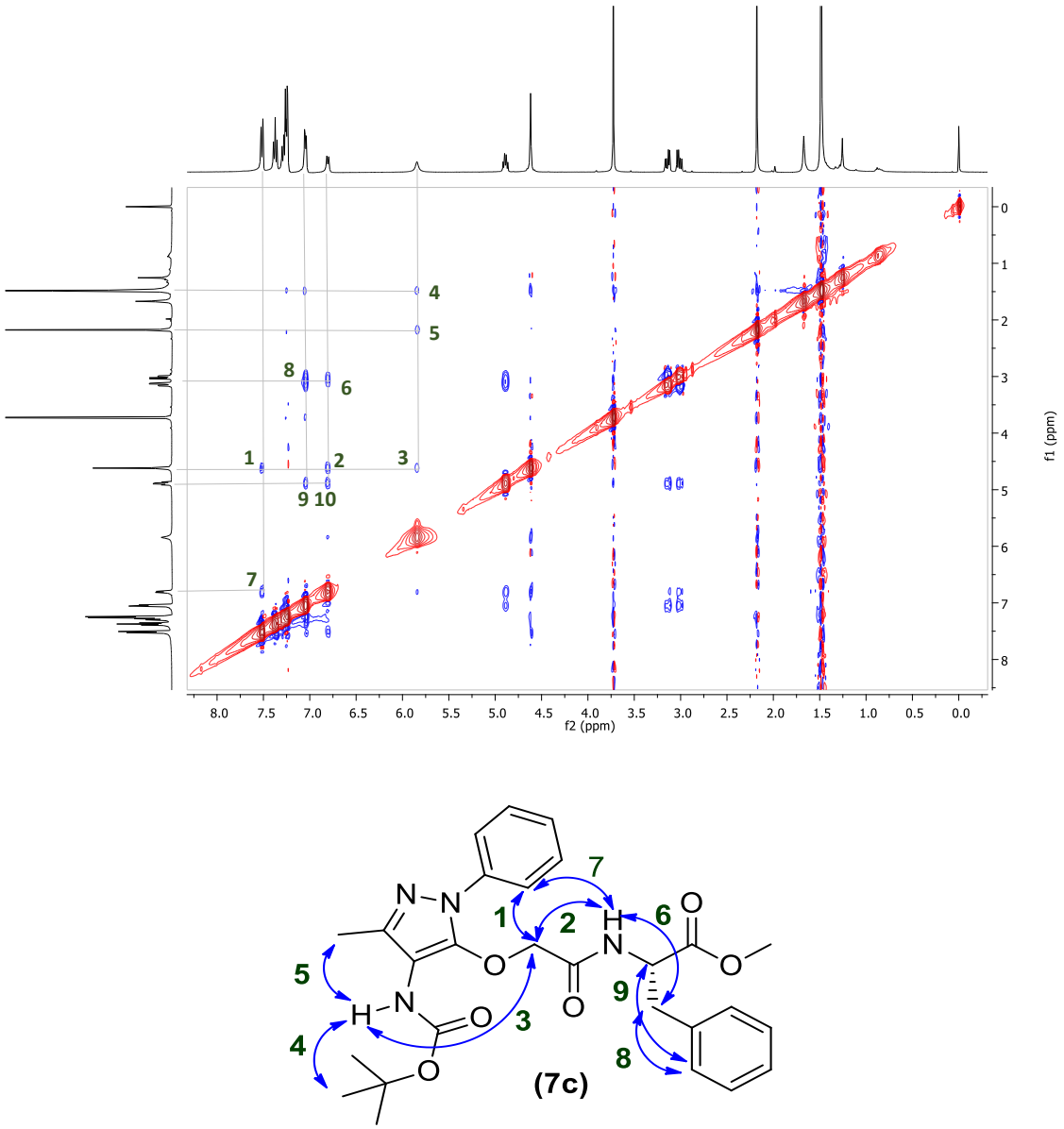


Figure S25. (2 of 3) ^{1}H-NOESY spectra of (**7c**) in $CDCl_3$.

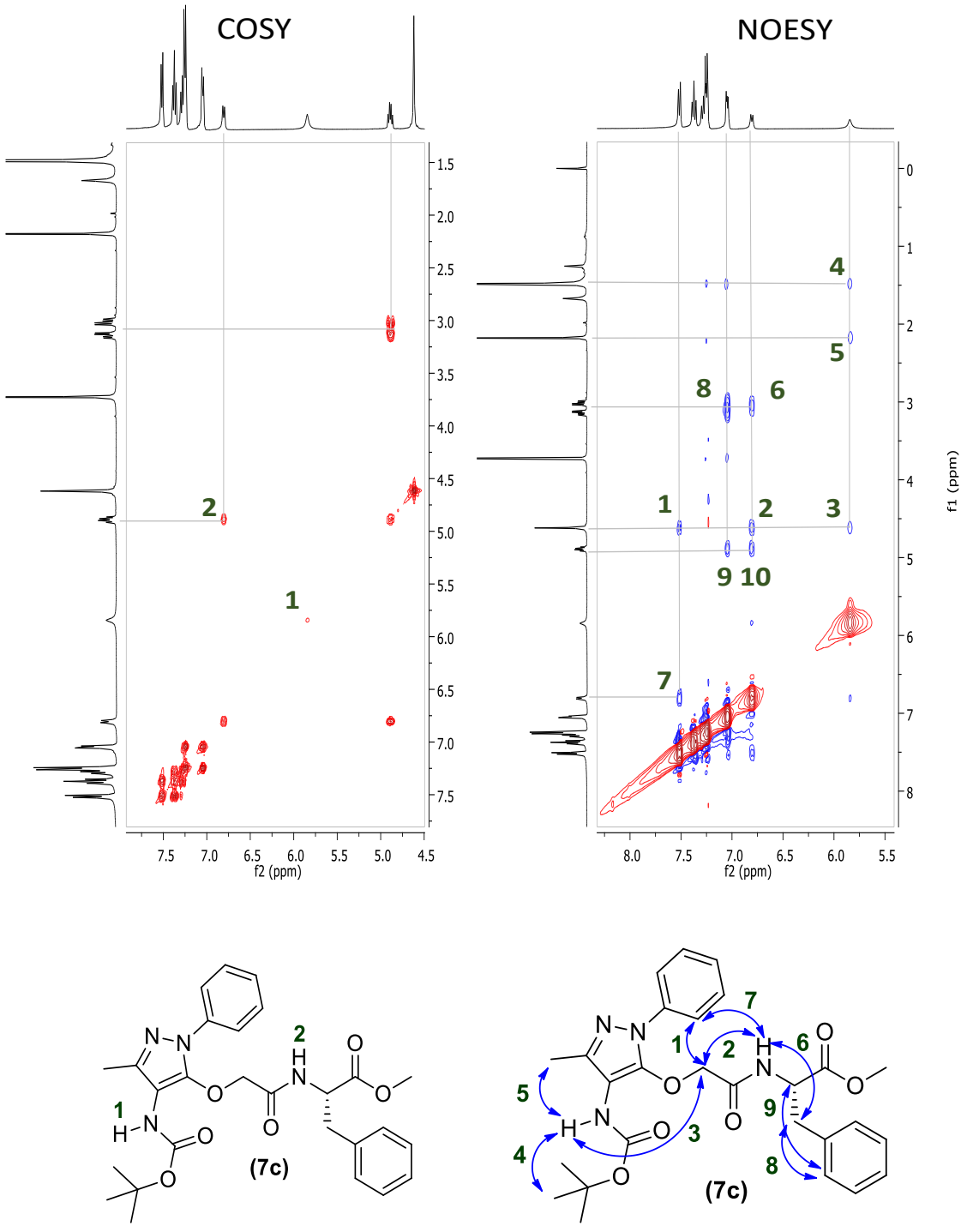


Figure S26. (3 of 3) ^{1}H-COSY & ^{1}H-NOESY expanded spectra in amide region of (**7c**) in $CDCl_3$.

14.^{1}H-COSY & ^{1}H-NOESY spectra of (7d) in $CDCl_3$

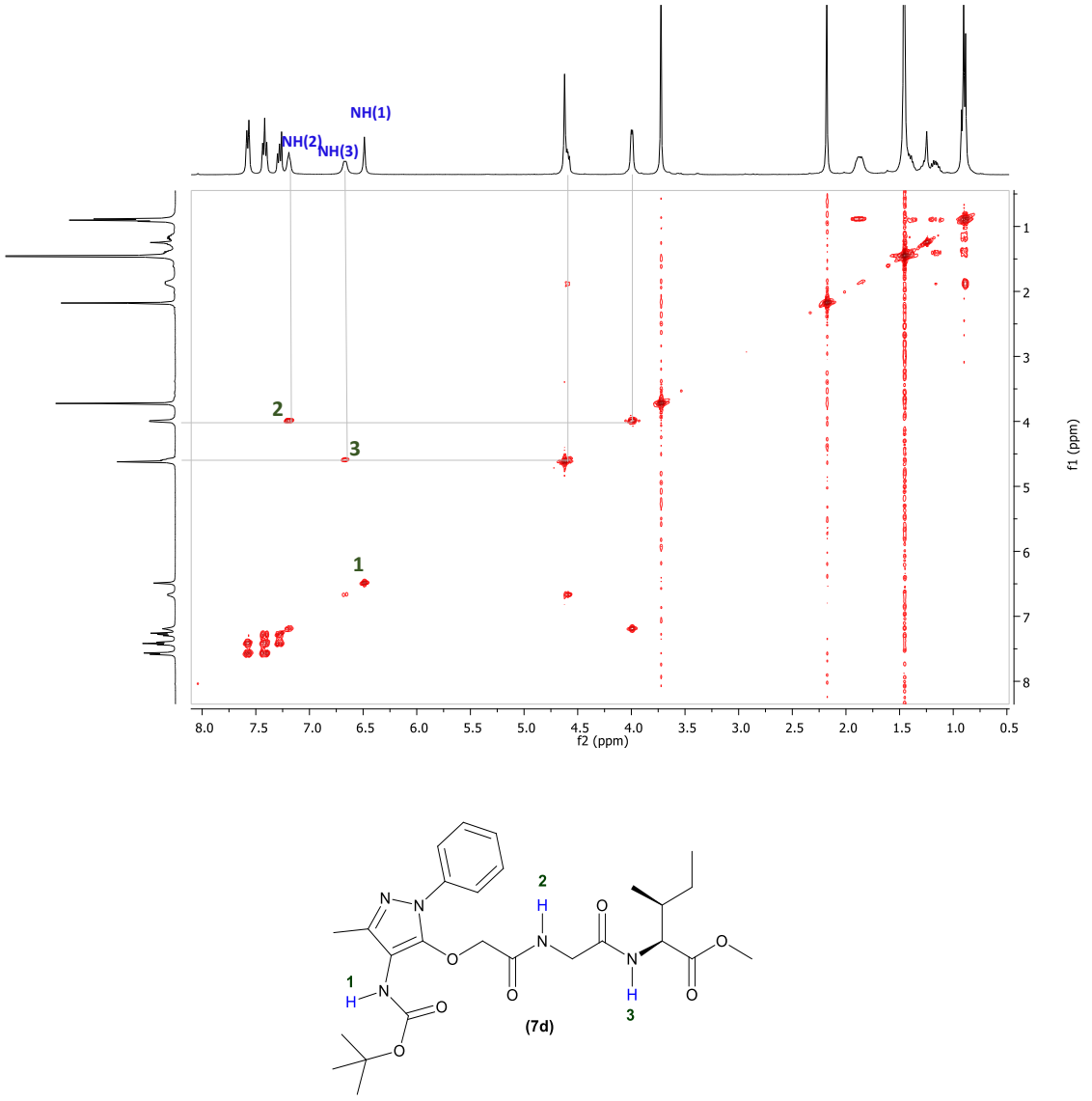


Figure S27. (1 of 3) ^{1}H-COSY spectra of (**7d**) in $CDCl_3$.

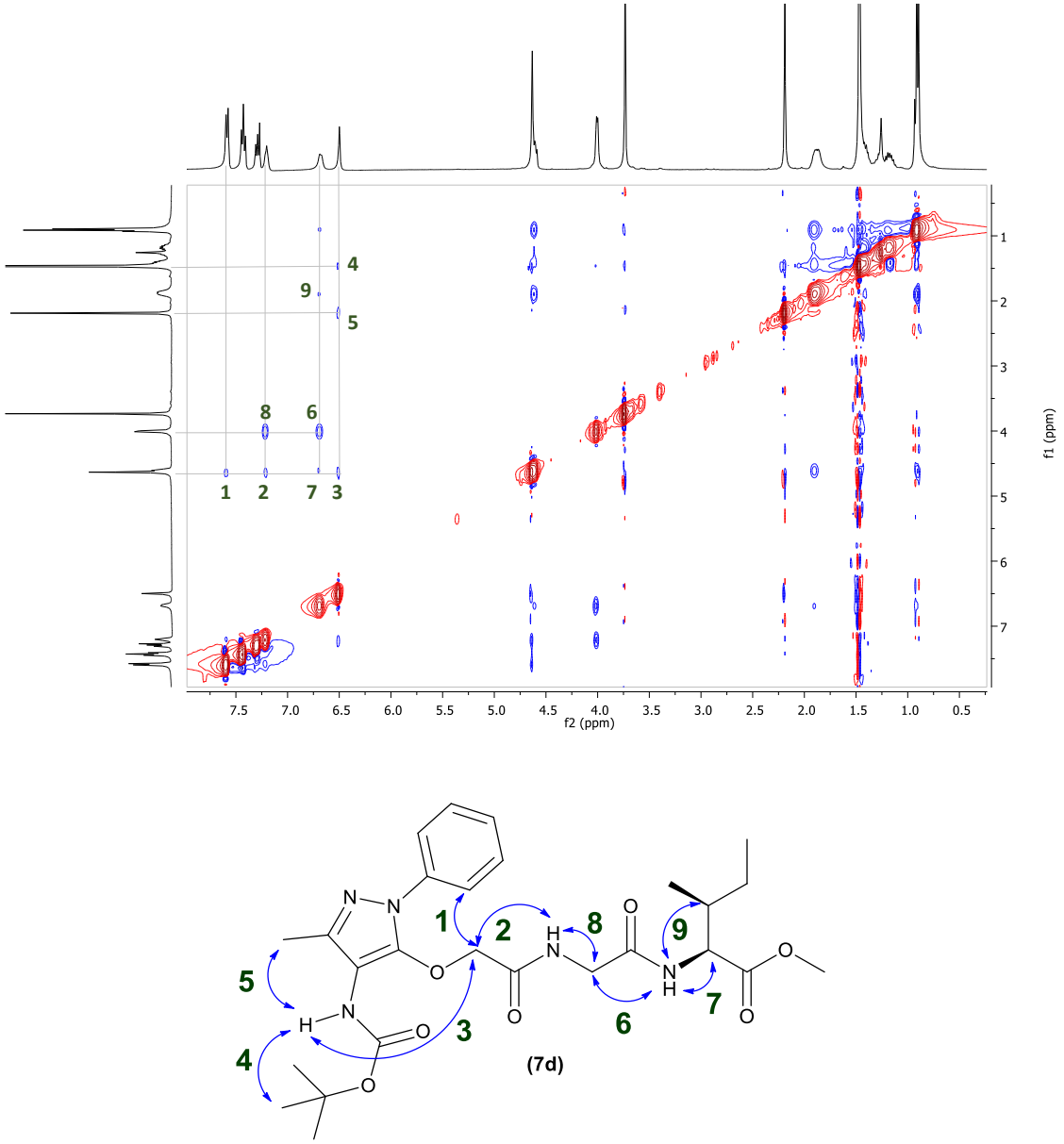


Figure S28. (2 of 3) ^{1}H-NOESY spectra of (**7d**) in $CDCl_3$.

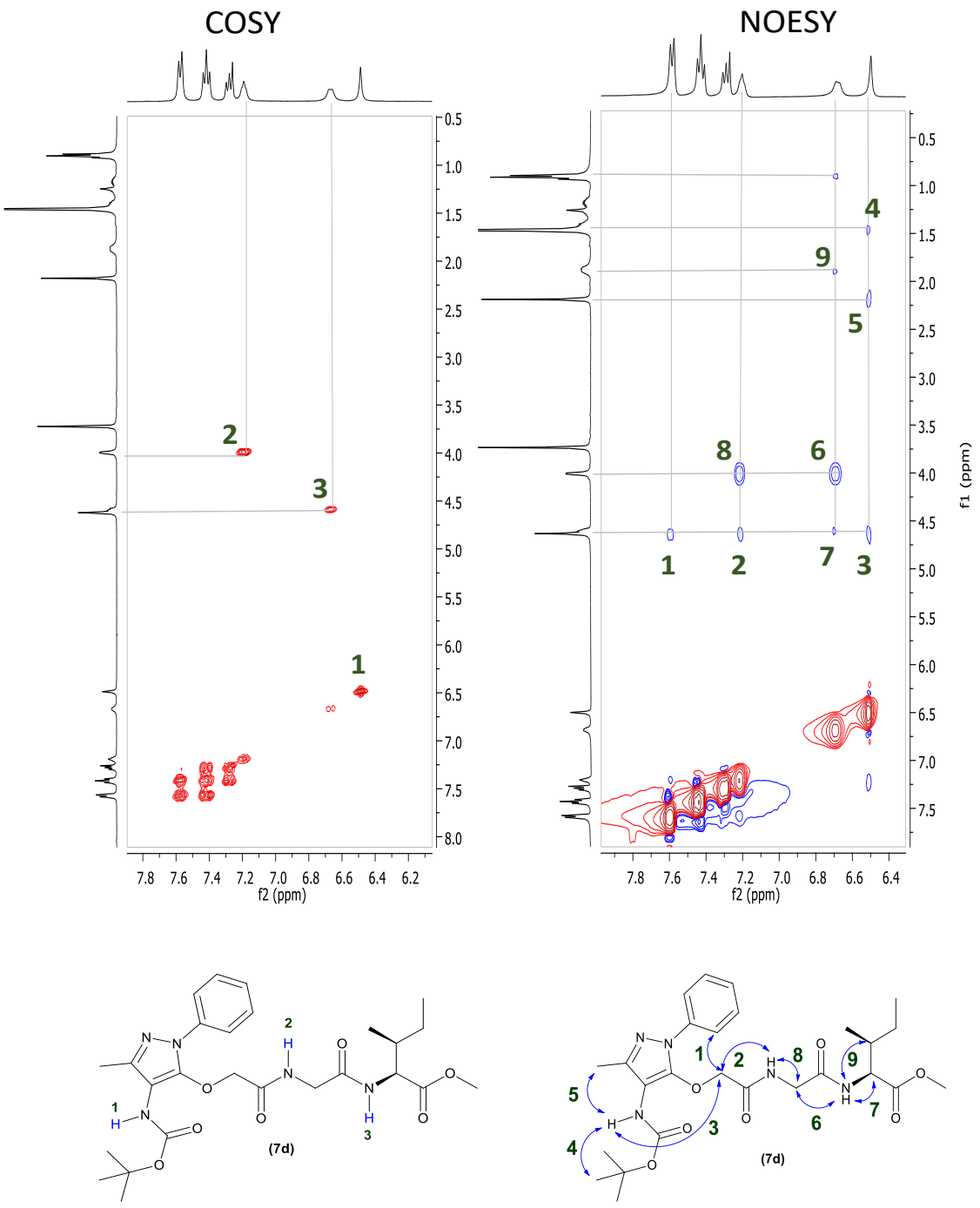


Figure S29. (3 of 3) Expanded spectra of ^{1}H-COSY & ^{1}H-NOESY in amide region of (**7d**) in $CDCl_3$.

15.^{1}H-COSY & ^{1}H-NOESY spectra of (7e) in $CDCl_3$

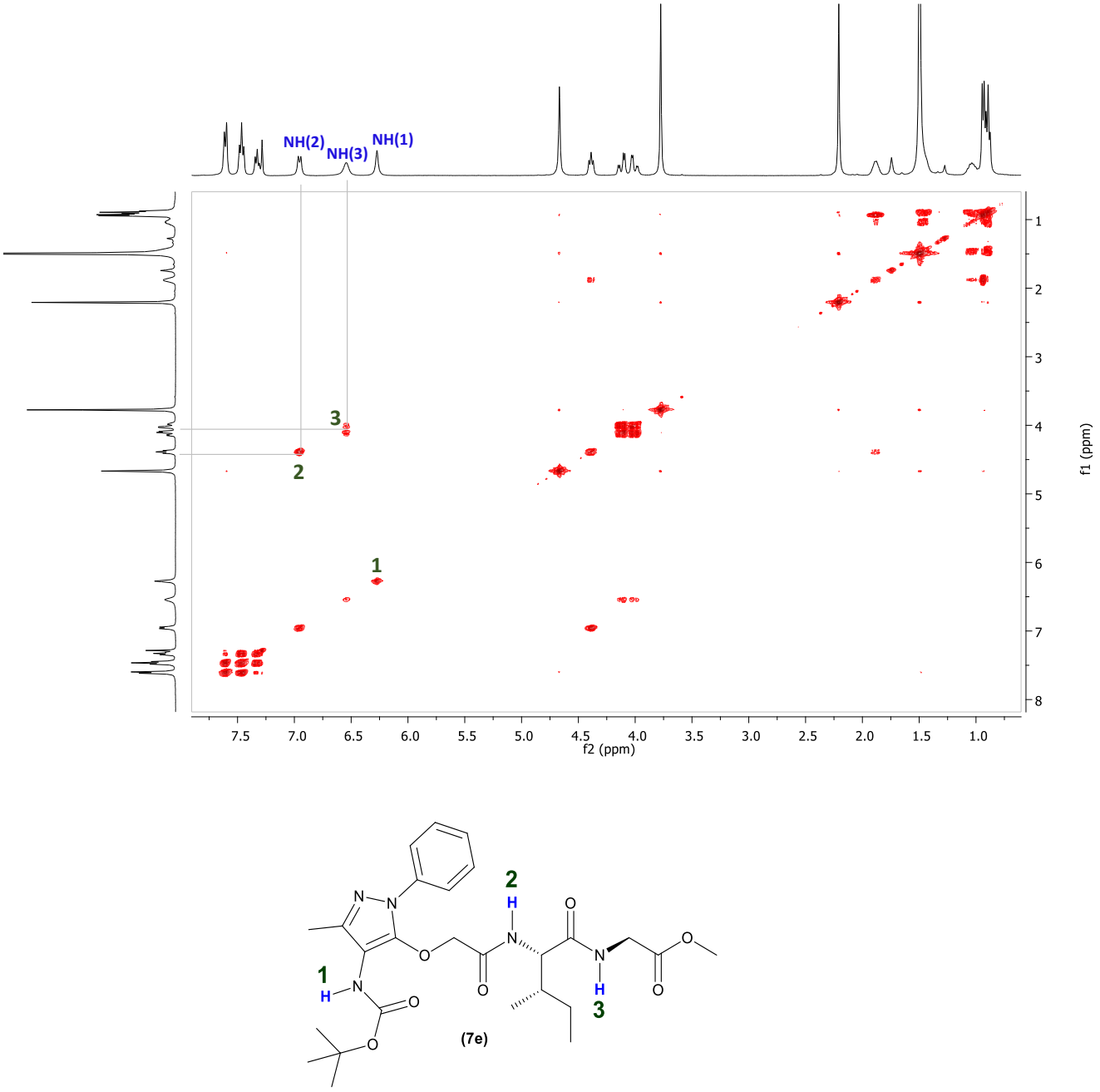


Figure S30. (1 of 3) ^{1}H-COSY spectra of (**7e)** in $CDCl_3$.

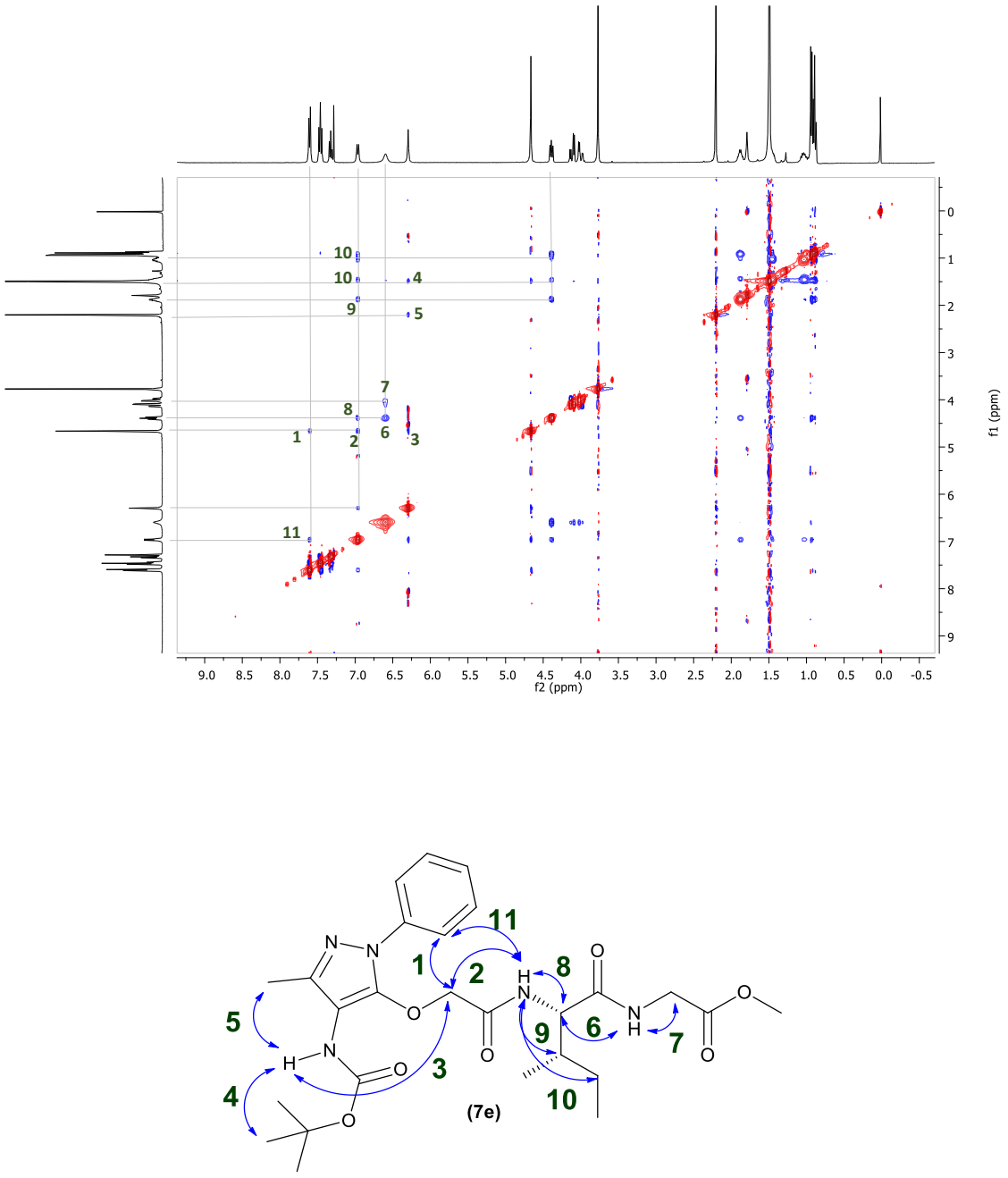


Figure S31. (2 of 3) ^{1}H-NOESY spectra of (**7e**) in $CDCl_3$.

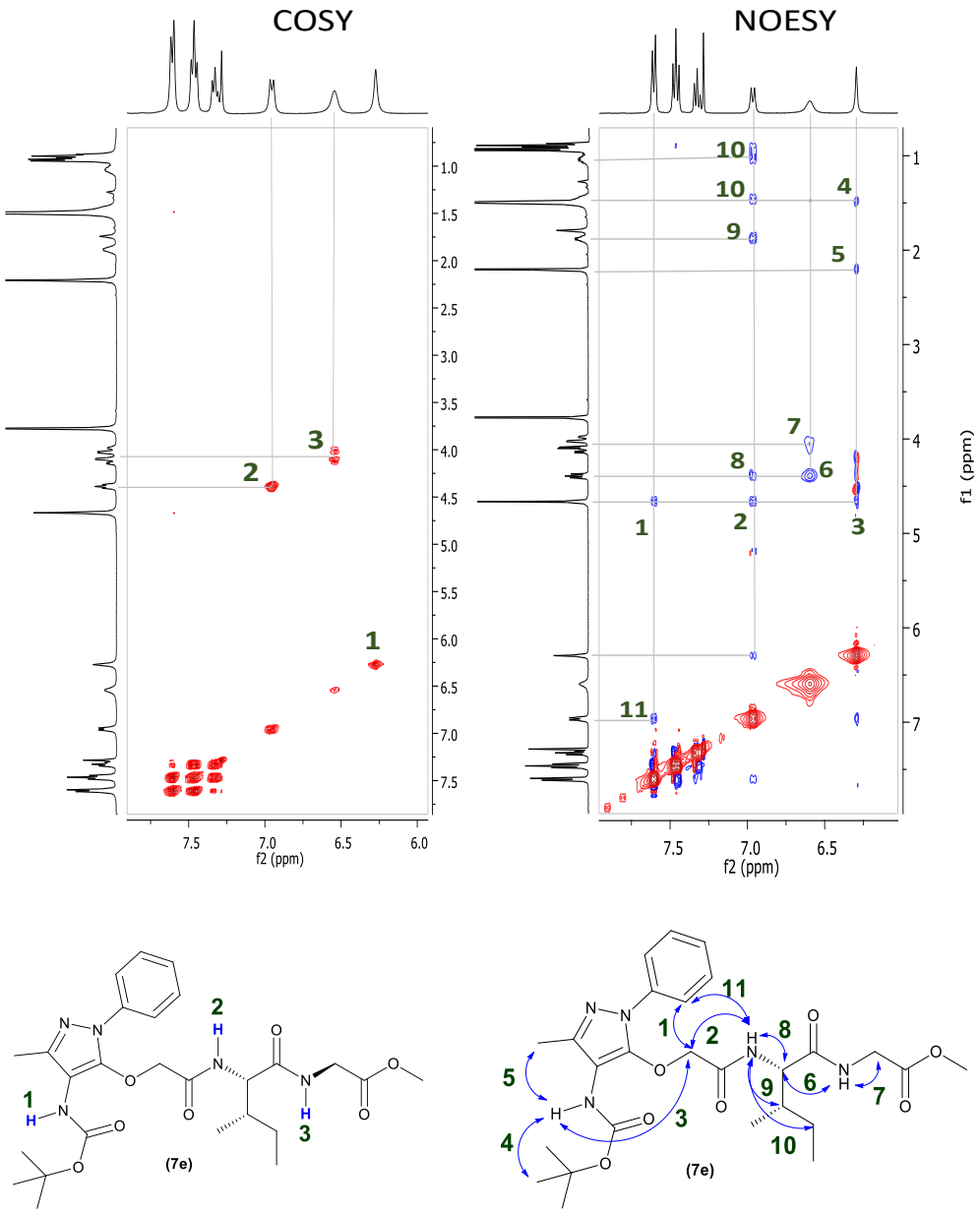


Figure S32. (3 of 3) Expanded spectra of ^{1}H-COSY & ^{1}H-NOESY in amide region of (**7e**) in $CDCl_3$.

16.DMSO-d_6 titration experiments by ^{1}H-NMR of peptides (7a-7e) in $CDCl_3$

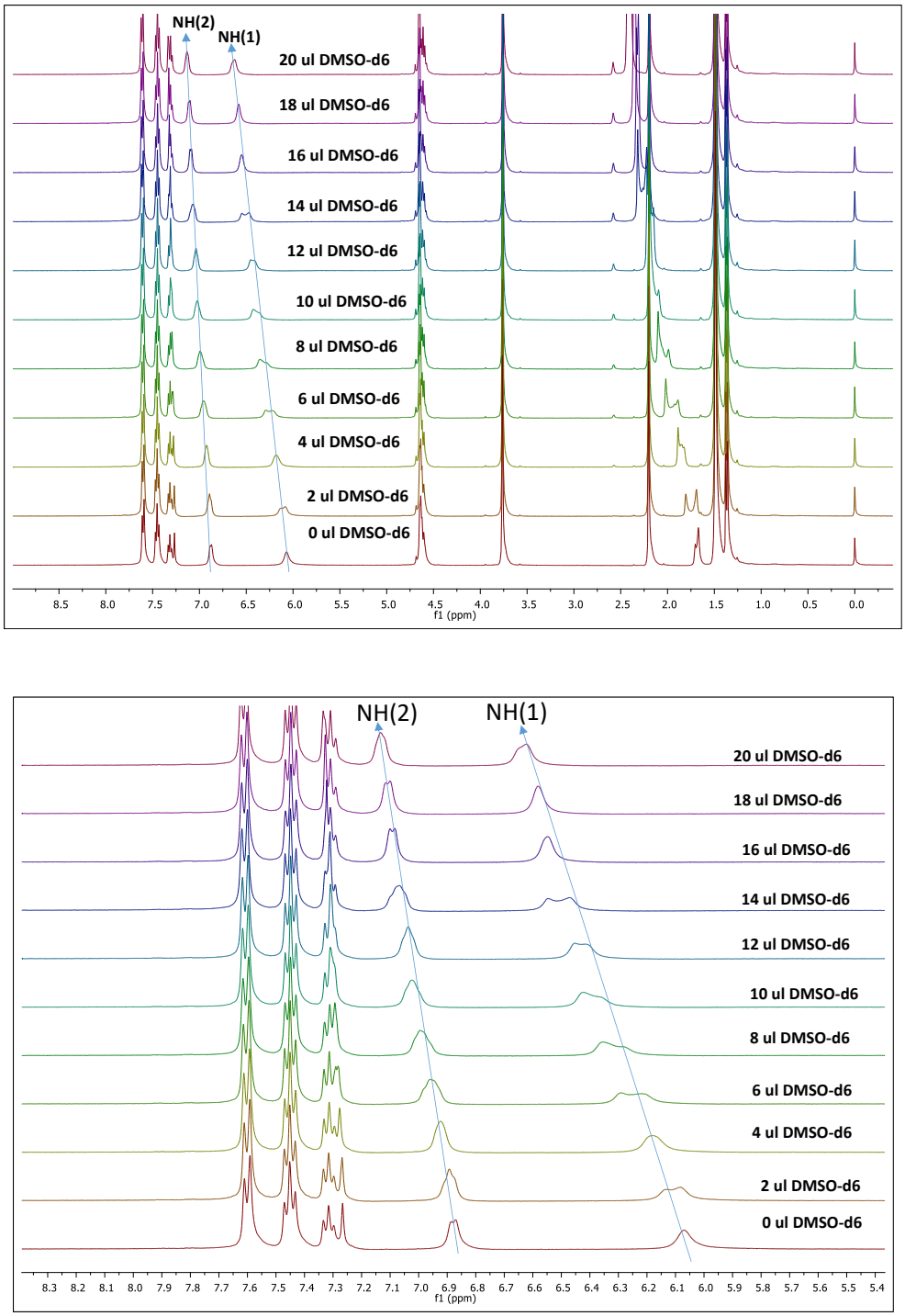
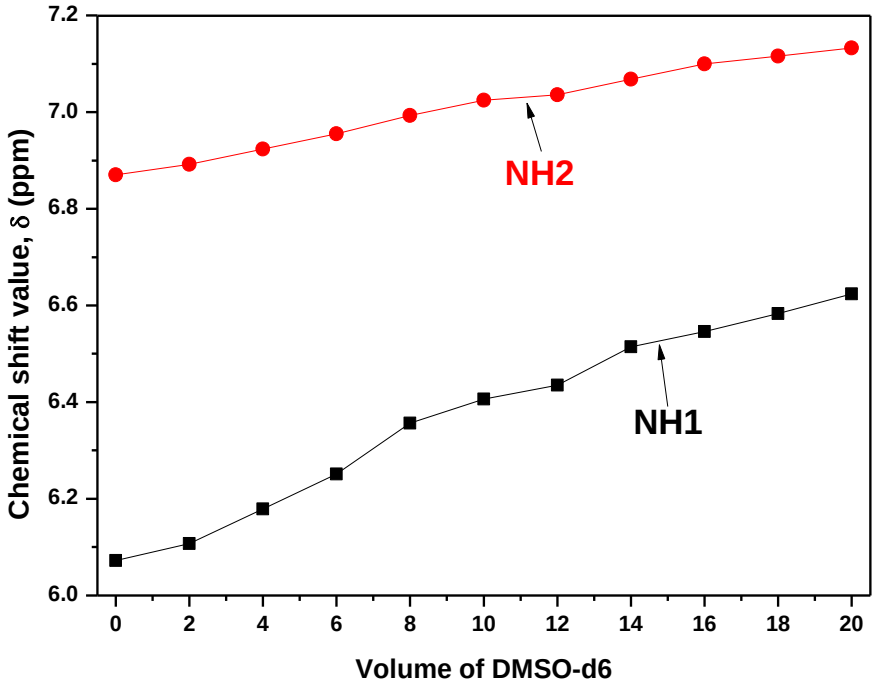


Figure S33. (1 of 2) ^{1}H-NMR of DMSO-d_6 titration of (**7a**) in $CDCl_3$ & its expanded spectral region.



	NH(1)			NH(2)		
Compound	**0 µL DMSO-d6**	**20 µL DMSO-d6**	**dδ (ppm)**	**0 µL DMSO-d6**	**20 µL DMSO-d6**	**dδ (ppm)**
7a	6.072	6.624	0.552	6.87	7.133	0.263

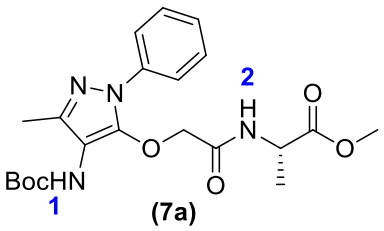
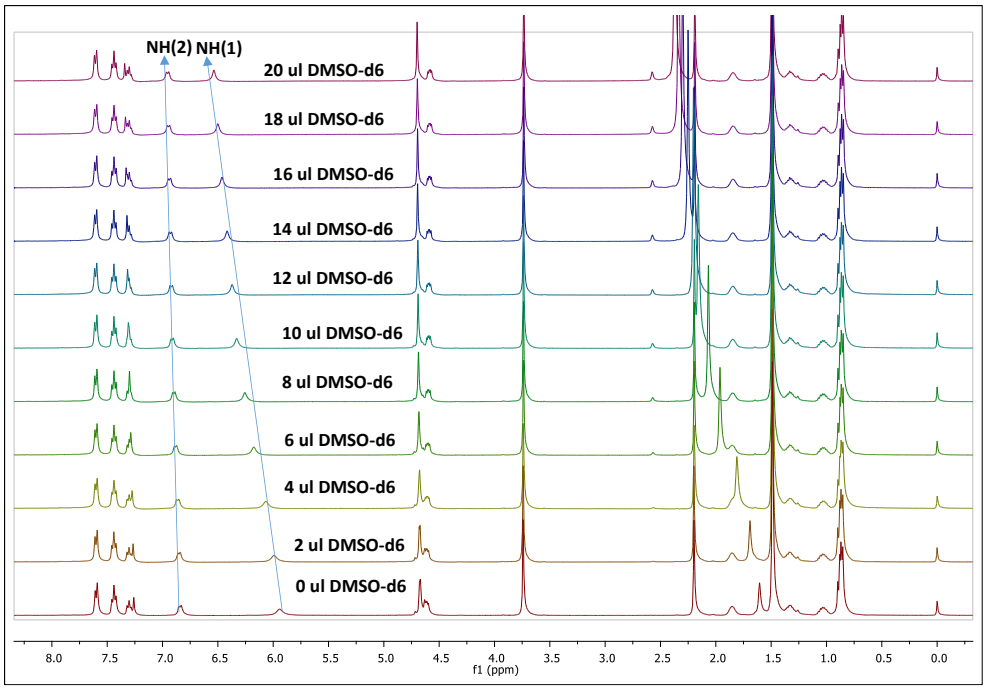


Figure S34. (2 of 2) DMSO-d_6 titration profile of (**7a**) in $CDCl_3$



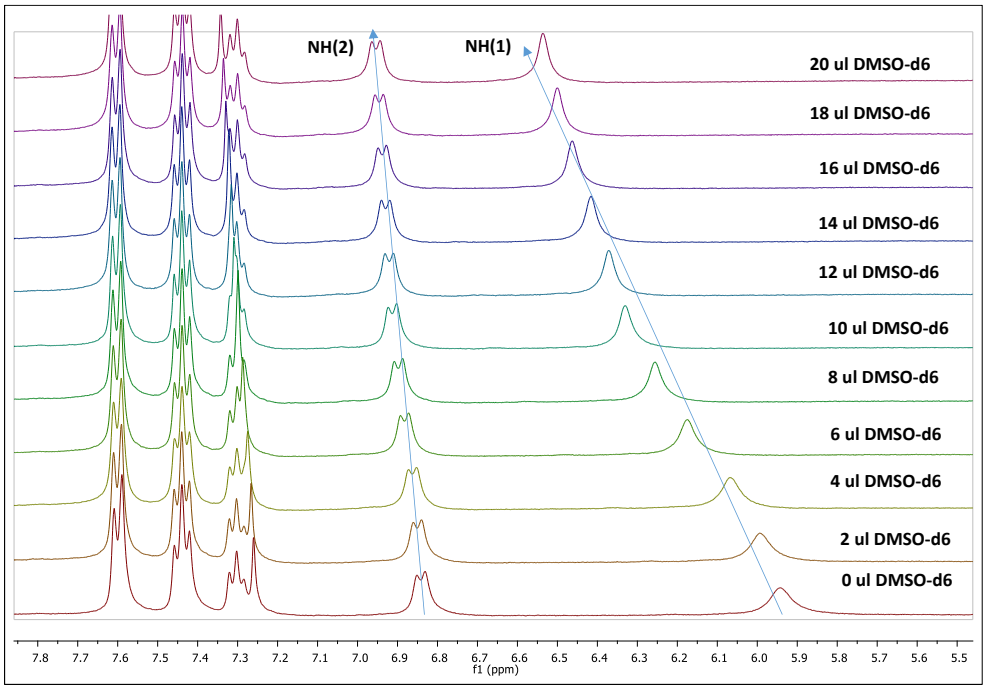
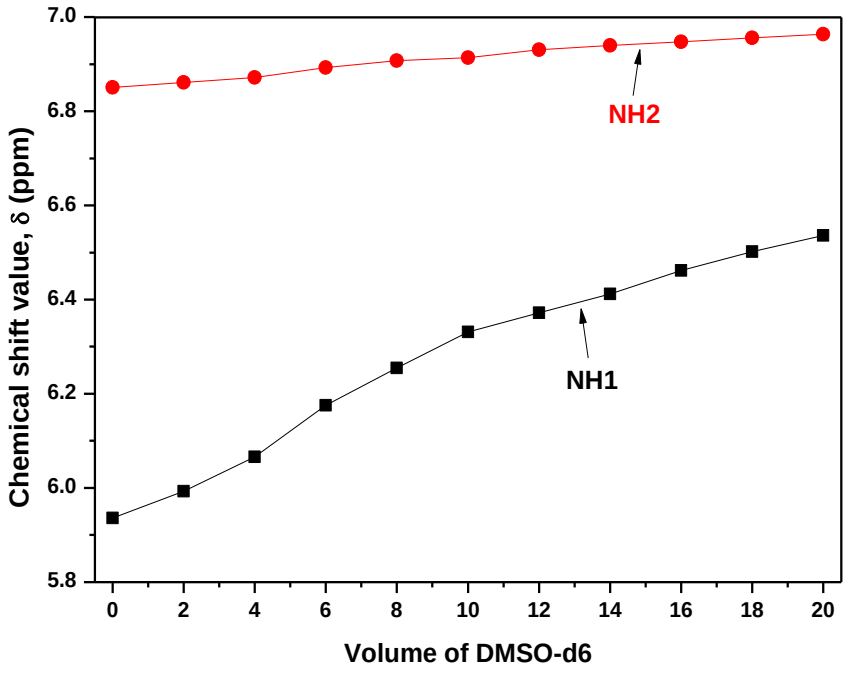


Figure S35. (1 of 2) ^{1}H-NMR of DMSO-d_6 titration of (**7b**) in $CDCl_3$ & its expanded spectral region.



	NH(1)			NH(2)		
Compound	**0 µL DMSO-d6**	**20 µL DMSO-d6**	**dδ (ppm)**	**0 µL DMSO-d6**	**20 µL DMSO-d6**	**dδ (ppm)**
7b	5.936	6.536	0.6	6.851	6.964	0.113

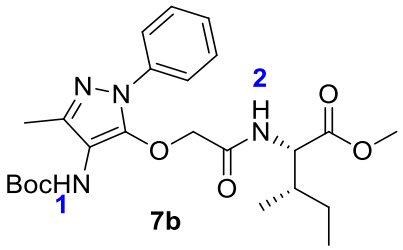


Figure S36. (2 of 2) DMSO-d_6 titration profile of (**7b**) in $CDCl_3$

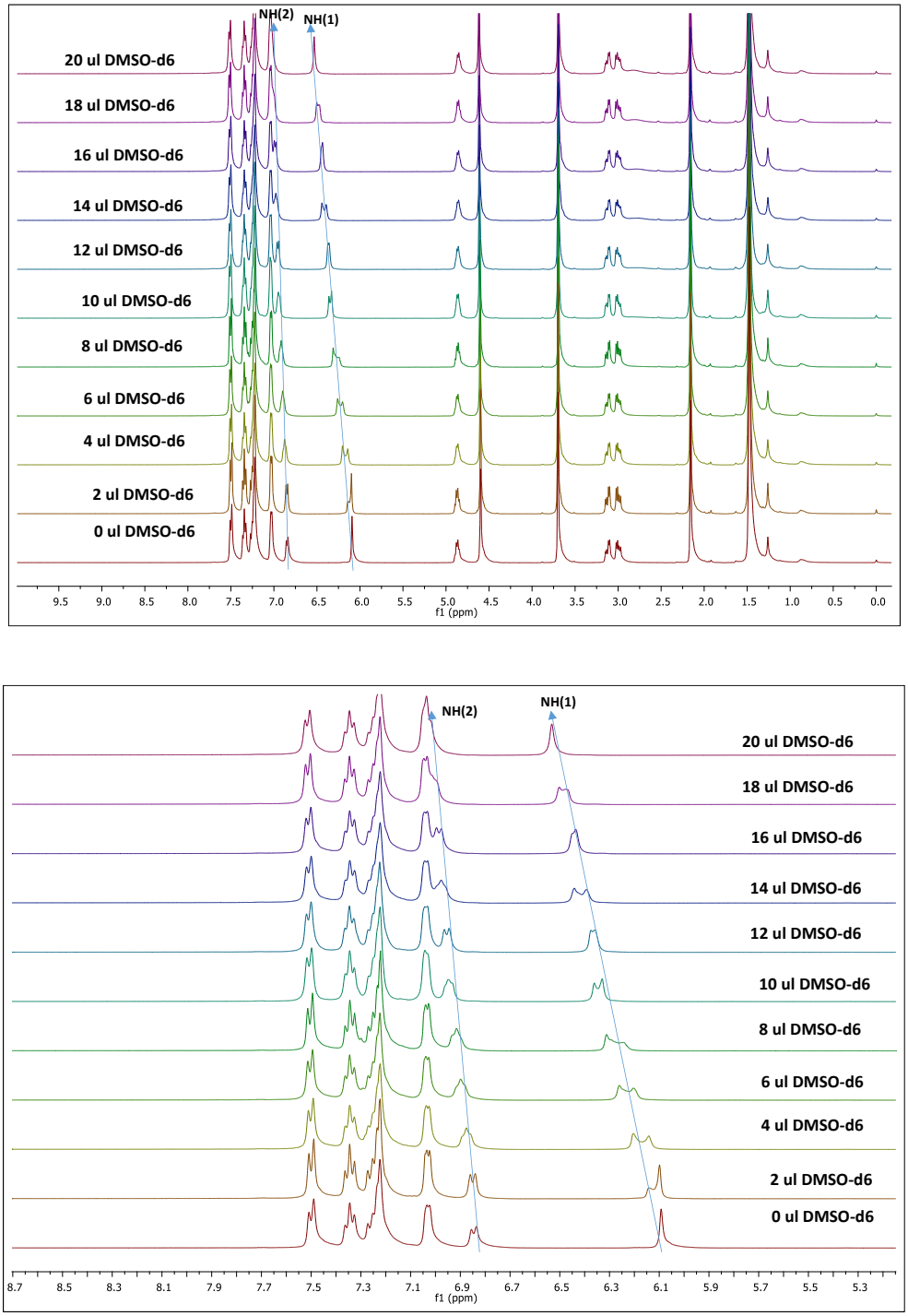
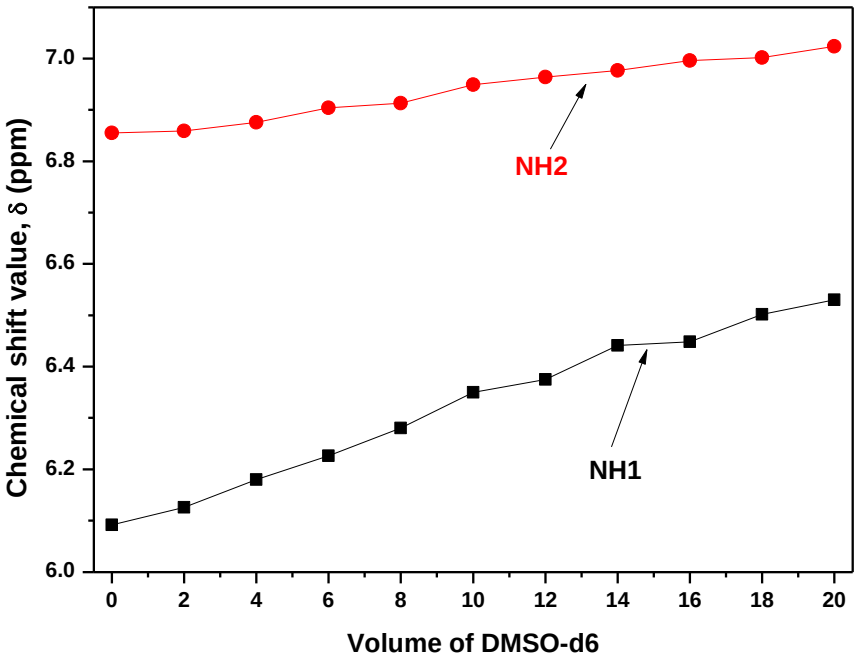


Figure S37. (1 of 2) ^{1}H-NMR of DMSO-d_6 titration of (**7c**) in $CDCl_3$ & its expanded spectral region.



	NH(1)			NH(2)		
Compound	**0 µL DMSO-d6**	**20 µL DMSO-d6**	**dδ (ppm)**	**0 µL DMSO-d6**	**20 µL DMSO-d6**	**dδ (ppm)**
7c	6.092	6.53	0.438	6.855	7.024	0.169

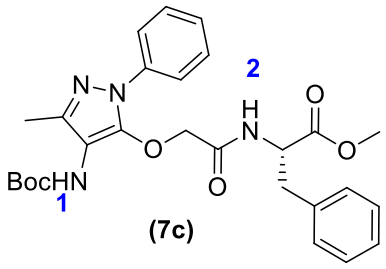


Figure S38. (2 of 2) DMSO-d_6 titration profile of (**7c**) in $CDCl_3$

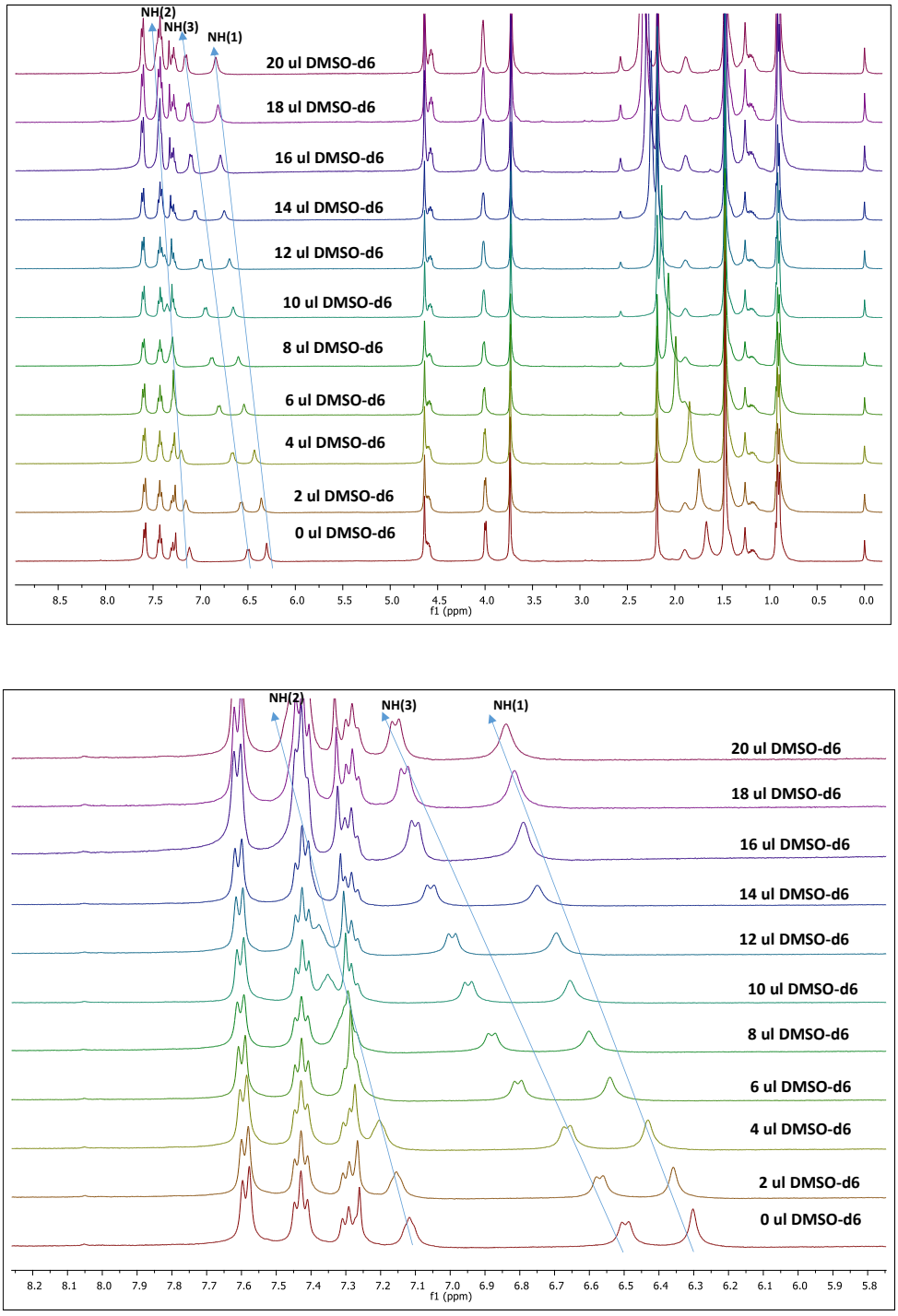
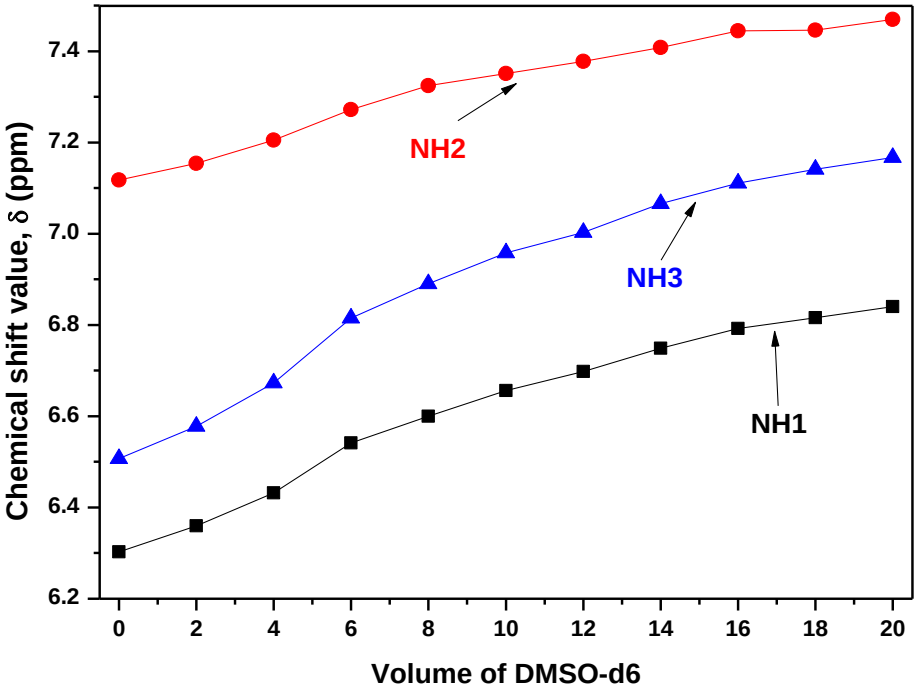


Figure S39. (1 of 2) ^{1}H-NMR of DMSO-d_6 titration of (**7d**) in $CDCl_3$ & its expanded spectral region.



	NH(1)			NH(2)			NH(3)		
Compound	**0 µL DMSO-d6**	**20 µL DMSO-d6**	**dδ (ppm)**	**0 µL DMSO-d6**	**20 µL DMSO-d6**	**dδ (ppm)**	**0 µL DMSO-d6**	**20 µL DMSO-d6**	**dδ (ppm)**
7d	6.302	6.84	0.538	7.118	7.47	0.352	6.507	7.167	0.66

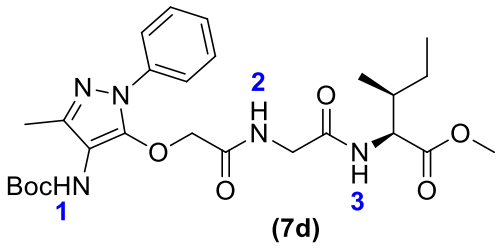


Figure S40. (2 of 2) DMSO-d_6 titration profile of (**7d**) in $CDCl_3$

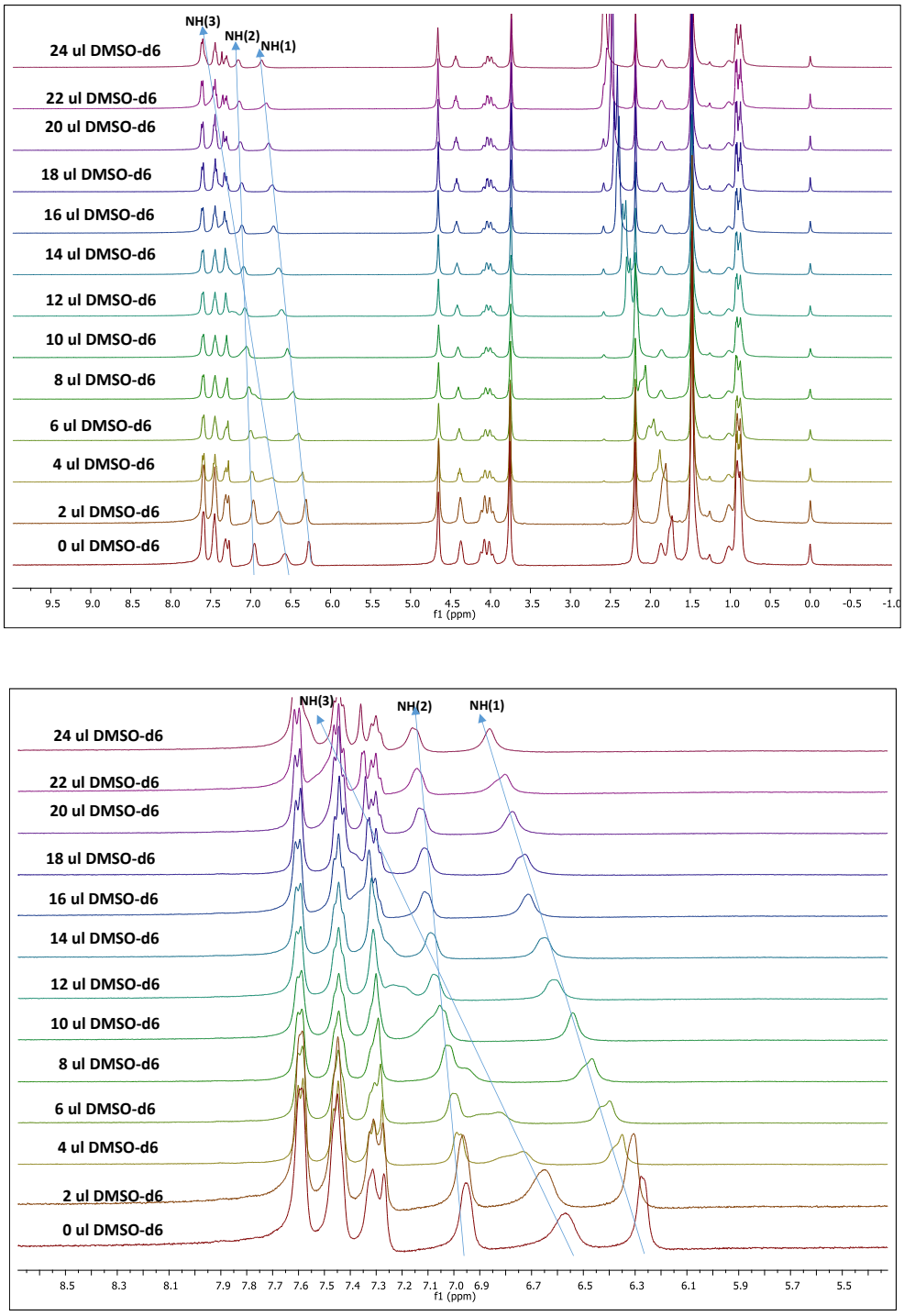
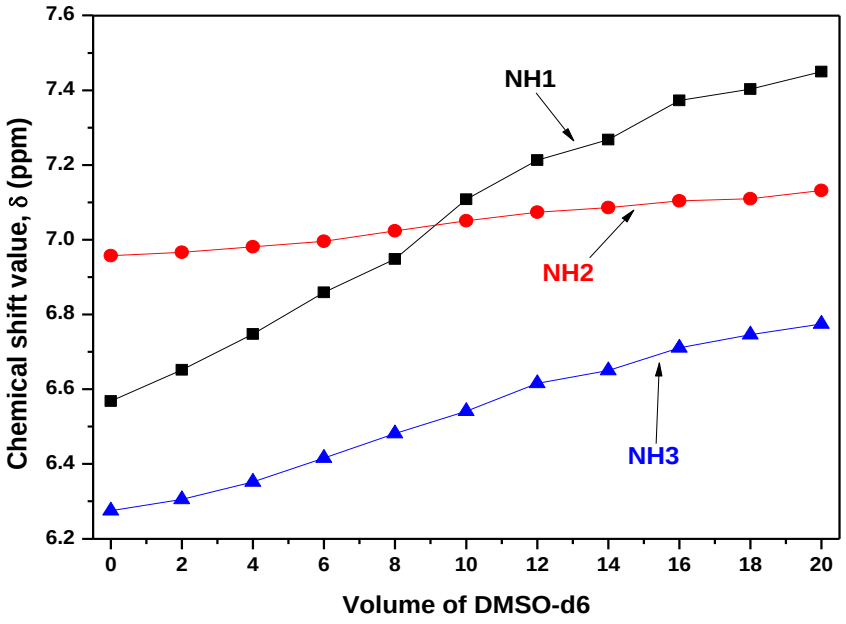


Figure S41. (1 of 2) ^{1}H-NMR of DMSO-d_6 titration of (**7e**) in $CDCl_3$ & its expanded spectral region.



	NH(1)			NH(2)			NH(3)		
Compound	**0 µL DMSO-d6**	**20 µL DMSO-d6**	**dδ (ppm)**	**0 µL DMSO-d6**	**20 µL DMSO-d6**	**dδ (ppm)**	**0 µL DMSO-d6**	**20 µL DMSO-d6**	**dδ (ppm)**
7e	6.568	7.4	0.832	6.957	7.132	0.175	6.275	6.775	0.5

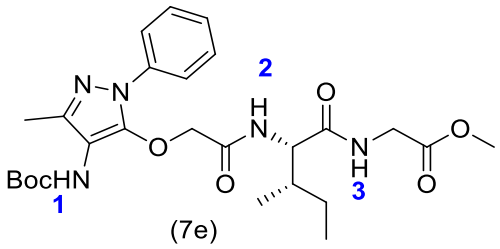


Figure S42. (2 of 2) DMSO-d_6 titration profile of (**7e**) in $CDCl_3$

Table.S1. Difference in Chemical shift values (dδ ppm) of N-H before and after adding DMSO-d_6 of APA peptides (**7a-7e**)

Compound	Chemical shift value (δ, ppm)								
	NH(1)			NH(2)			NH(3)		
	0 µL DMSO-d6	20 µL DMSO-d6	dδ (ppm)	0 µL DMSO-d6	20 µL DMSO-d6	dδ (ppm)	0 µL DMSO-d6	20 µL DMSO-d6	dδ (ppm)
7a	6.072	6.624	0.552	6.87	7.133	0.263	-	-	-
7b	5.936	6.536	0.6	6.851	6.964	0.113	-	-	-
7c	6.092	6.53	0.438	6.855	7.024	0.169	-	-	-
7d	6.302	6.84	0.538	7.118	7.47	0.352	6.507	7.167	0.66
7e	6.568	7.4	0.832	6.957	7.132	0.175	6.275	6.775	0.5

17.Crystallography

A suitable single crystal of each complex was carefully selected under a polarizing microscope. Single crystal structure determination by X–ray diffraction was performed on a Siemens SMART-CCD diffractometer equipped with a normal focus, 2.4 kW sealed-tube X-ray source (*Mo–Kα* radiation, λ = 0.71073Å) operating at 50 kV and 30 mA. Structures were solved by the direct method using SHELXT 2014 and refined on *F*2 by a full–matrix least-squares technique using the SHELXL 2014 programs package. An empirical absorption correction based on symmetry equivalent reflections was applied using SADABS. The graphic programs DIAMOND3.2 wasused to draw the structures. Non-hydrogen atoms were refined anisotropically. In the refinement, hydrogens were treated as riding atoms using the SHELXL default parameters. Details of crystal structure refinement parameters for **3, 4, 6, 5, 7a** and **7e** are given as Table S2.

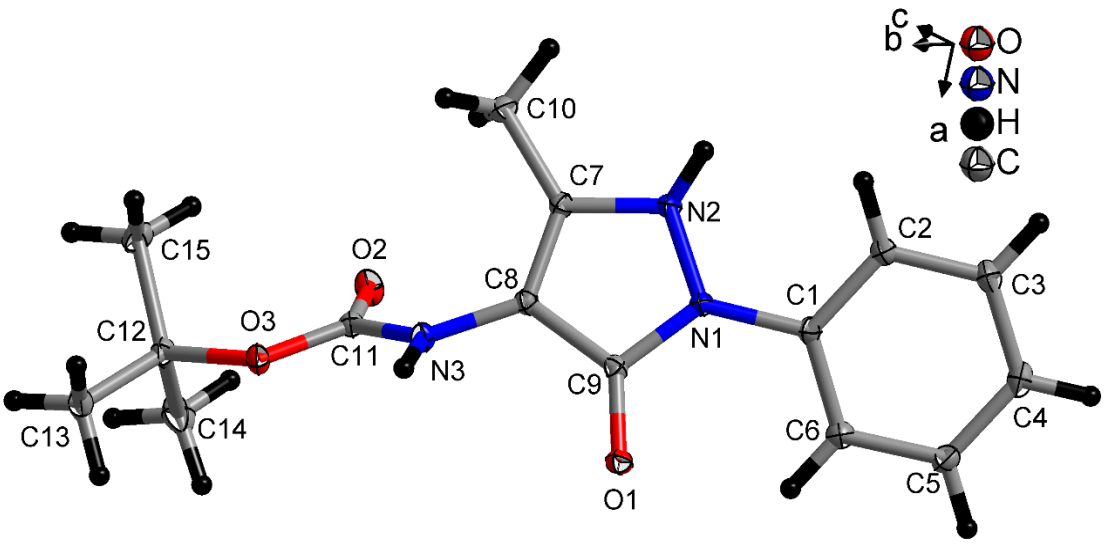


Figure S43. ORTEP diagram of compound (**3**)

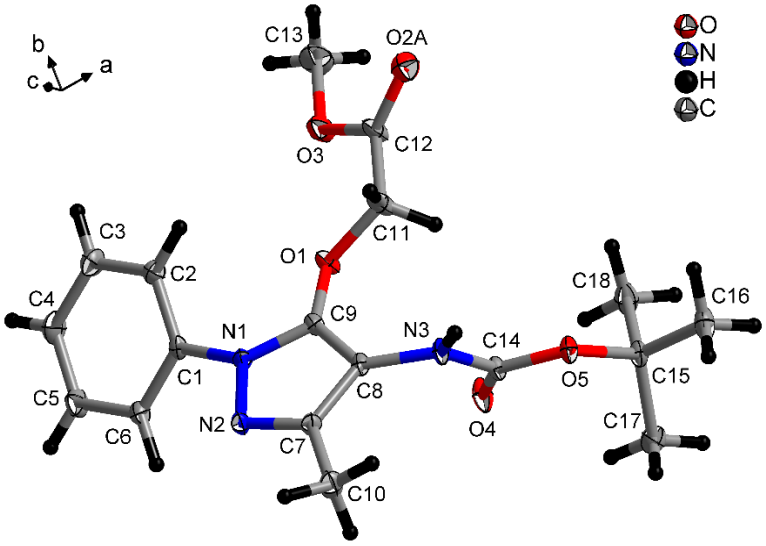


Figure S44. ORTEP diagram of *O*-alkylated APA-amino ester derivative (**4)**

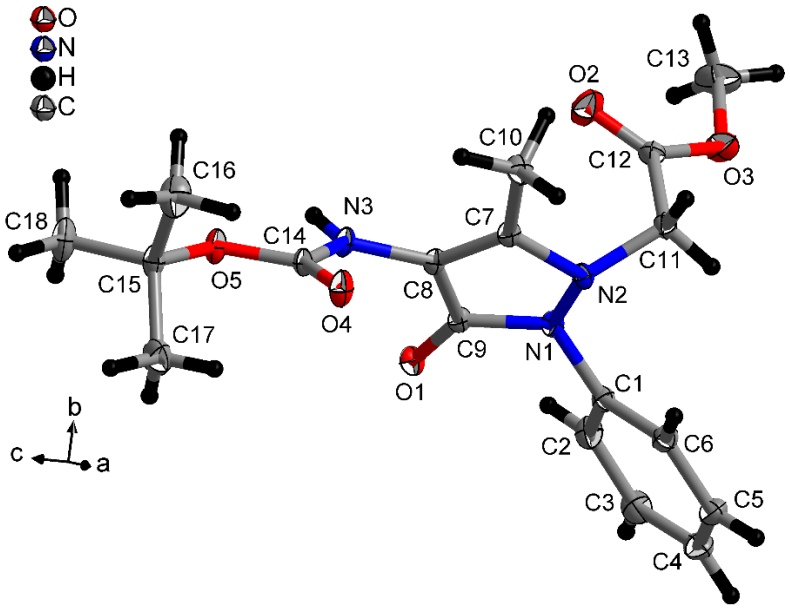


Figure S45. ORTEP diagram of *N*-alkylated APA-amino ester derivative (**5)**

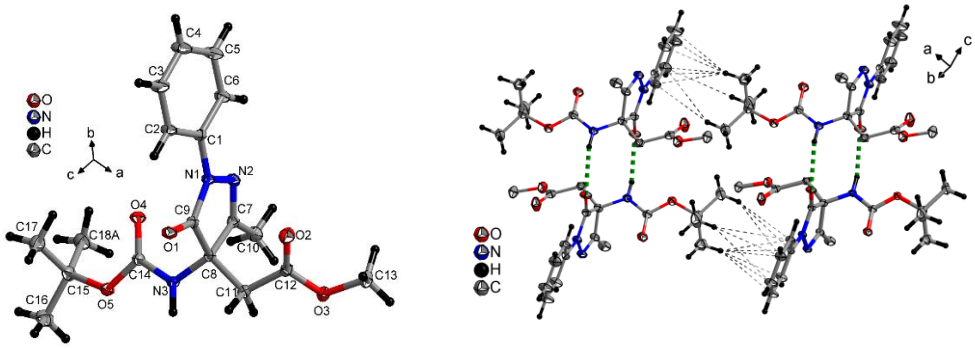


Figure S46. ORTEP diagram of *C*-alkylated APA-amino ester derivative (**6**) left**,** Non classical intermolecular hydrogen bonding, showing multiple *CH*-p hydrogen bonding (shown in grey dotted lines), and Classical intermolecular hydrogen bonding between *N-H---O* (shown in green dotted line) with a bond length ranging from 3.1-3.9 Å and 2.3 Å respectively. (Other hydrogen atoms which are not involved are omitted) (right)

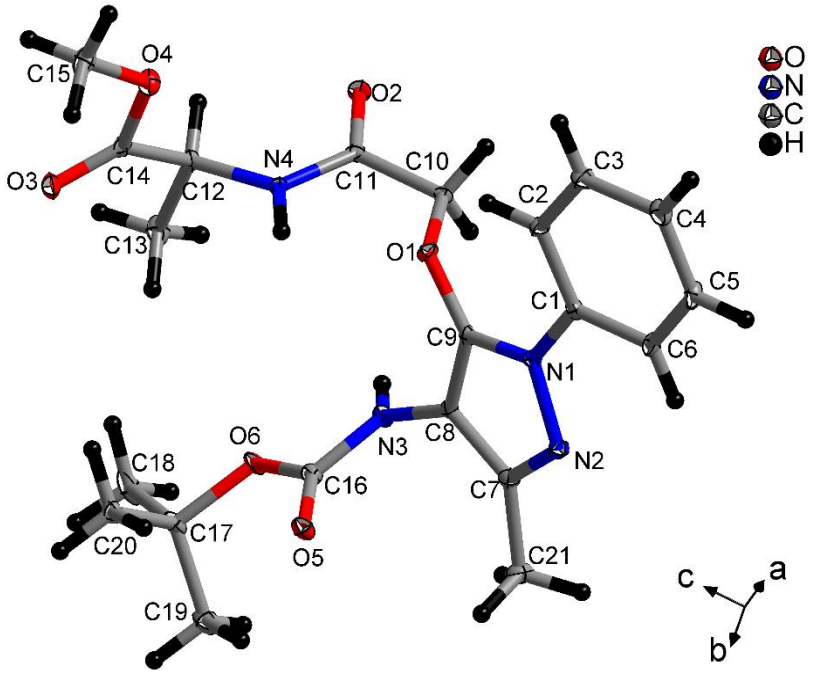


Figure S47. ORTEP diagram of APA dipeptide (**7a**)

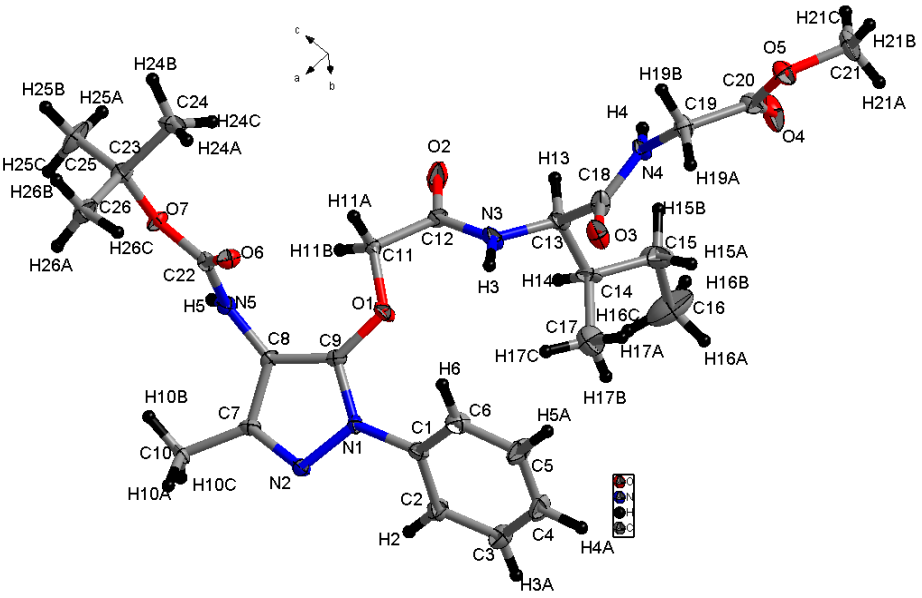


Figure S48. ORTEP diagram of APA tripeptide (7e**)**

Table S2 Crystallographic table

Parameter	**(3)**	**(4)**	**(6)**	**(5)**	**(7a)**	**(7e)**
Formula	C15 H19 N3 O3	C18 H23 N3 O5	C18 H23 N3 O5	C18 H23 N3 O5	C21 H28 N4 O6	C26 H37 N5 O7
*M*r	289.33	361.39	361.39	361.39	432.47	531.6
crystal system	Orthorhombic	Orthorhombic	Triclinic	Monoclinic	Triclinic	Monoclinic
space group	*P b c a*	*P b c a*	*P -1*	*P 21/c*	*P 1*	*P 21*
a/Å	11.7022(2)	10.8478(4)	9.1784(2)	10.6153(3)	9.9902(5)	11.734(3)
b/Å	9.9337(2)	17.2214(6)	10.7536(2)	14.1197(4)	10.5414(6)	8.9135(19)
c/Å	25.3769(5)	20.6904(8)	11.4779(3)	12.9649(4)	11.6835(6)	13.716(3)
α/°	90°.	90°.	105.538(3)°.	90°.	99.094(3)°.	90°.
β/°	90°.	90°.	97.2730(10)°.	96.941(2)°.	102.478(3)°.	94.354(11)°.
γ/°	90°.	90°.	113.8950(10)°.	90°.	109.046(4)°.	90°.
V/Å^3	2949.97(10)	3865.3(2)	961.72(4)	1929.00(10)	1100.20(10)	1430.4(5)
Z	8	8	2	4	2	2
ρ_{calc} (gcm^{-3})	1.303	1.242	1.248	1.244	1.305	1.234
μ/ mm^{-1} μ(cm^{-1})	0.092	0.092	0.092	0.092	0.097	0.091
θ range (deg)	3.482 to 29.119°.	2.427 to 25.495°.	2.213 to 25.500°.	2.141 to 30.528°.	1.843 to 25.500°.	1.741 to 24.989°.
reflections collected	47191	23680	14139	30152	15018	13036
Unique reflections[R(int)]	3960 [0.0509]	3588 [0.0482]	3577 [0.0206]	5899 [0.0398]	7043 [0.0514]	4890 [0.2282]
data/restraints/ parameters	3960 / 0 / 198	3588 / 0 / 241	3577 / 1 / 241	5899 / 0 / 240	7043 / 3 / 571	4890 / 5 / 347
GOF on F^2	1.017	1.045	1.042	1.023	1.024	0.866
*R*1 and *R*2 [$I>2\sigma(I)$]	0.0415, 0.1018	0.0535, 0.1308	0.0453, 0.1172	0.0499, 0.1232	0.0448, 0.0926	0.0970, 0.1438
*R*1 and *R*2 (all data)	0.0613, 0.1136	0.0837, 0.1528	0.0558, 0.1275	0.0893, 0.1432	0.0575, 0.1000	0.3054, 0.1883
largest residual peaks (e.Å^{-3})	0.223 and -0.200	0.487 and -0.357	0.338 and -0.257	0.310 and -0.241	0.163 and -0.205	0.200 and -0.177
CCDC no	**1831988**	**1831989**	**1831993**	**1831990**	**1831991**	**1831992**

18.UV Spectra of peptide (7c)

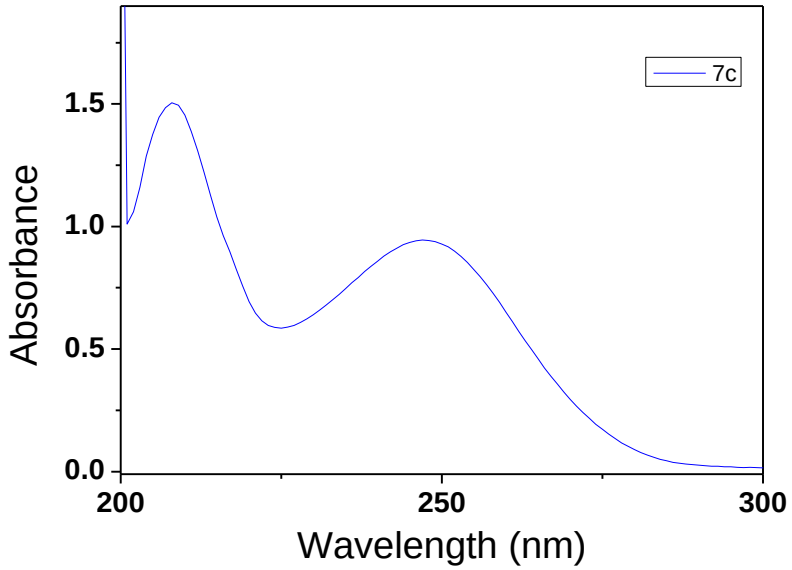


Figure S49. UV spectra of peptide (**7c**)

19.Circular Dichroism (CD) spectra.

CD spectra were recorded in degassed CH_3OH, AcCN, $CHCl_3$, and CF3CH2OH at 20 ^{0}C from 300-190 nm with peptide concentrations of 0.1 mM. CD data is collected with following parameters, Data pitch 2 nm, DIT 2 sec, bandwidth 2 nm, scanning speed 100 nm/min.

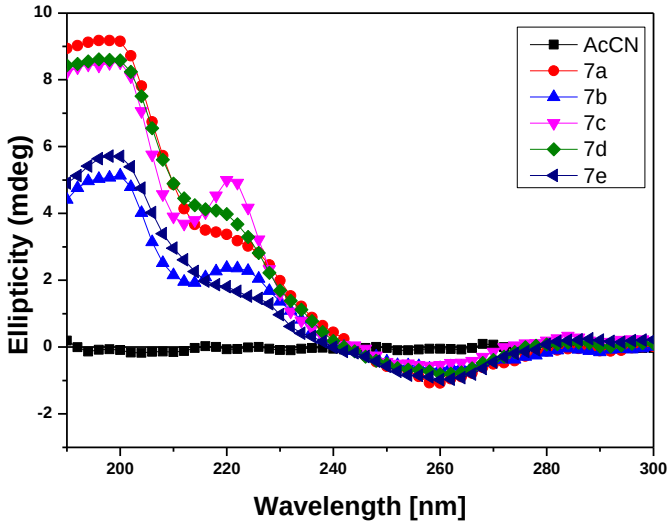


Figure S50. CD spectra of peptide (**7a-7e)** in Acetonitrile (AcCN)

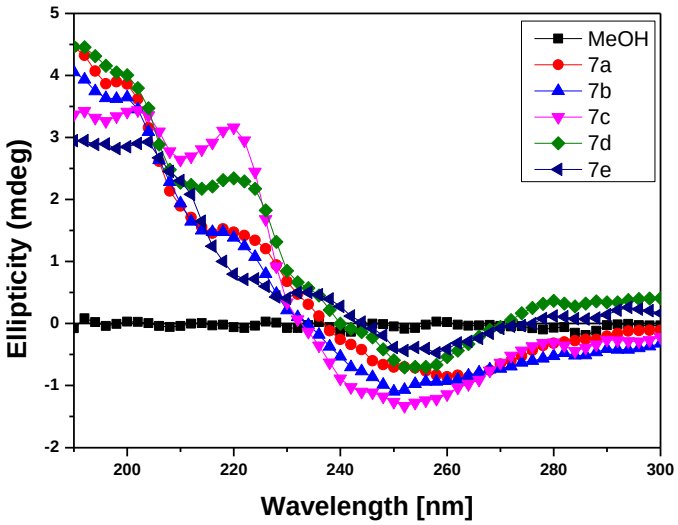
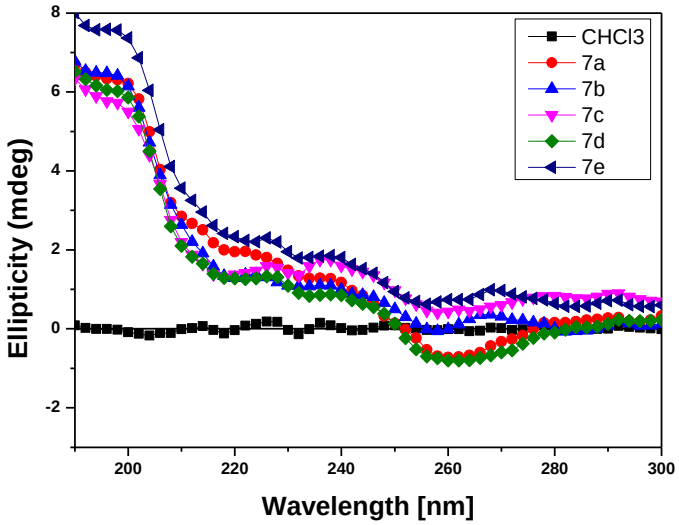


Figure S51. CD spectra of peptide (**7a-7e**) in Methanol (MeOH)



3

Figure S52. CD spectra of peptide (**7a-7e**) in Chloroform ($CHCl_3$)

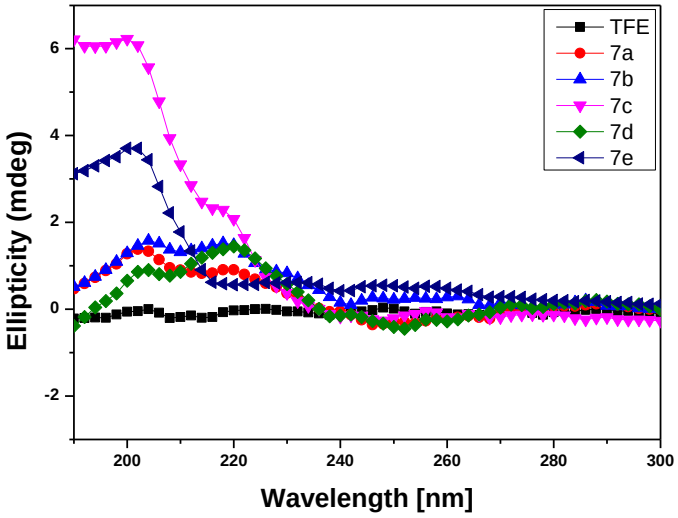


Figure S53. CD spectra of peptide (**7a-7e**) in Trifluoroethanol (TFE)

20.FT-IR spectra.

FT-IR spectra were recorded by dissolving respective compounds in degassed $CHCl_3$, drop casted on KBr plate and dried thoroughly. The samples were scanned from 500-4000 ($\tilde{\nu}$, cm^{-1}), the spectra are obtained from average of 64 scans.

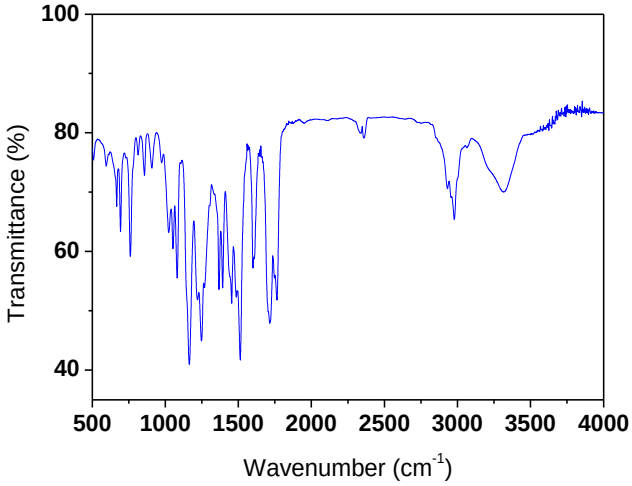


Figure S54. FT-IR spectra of (**4**)

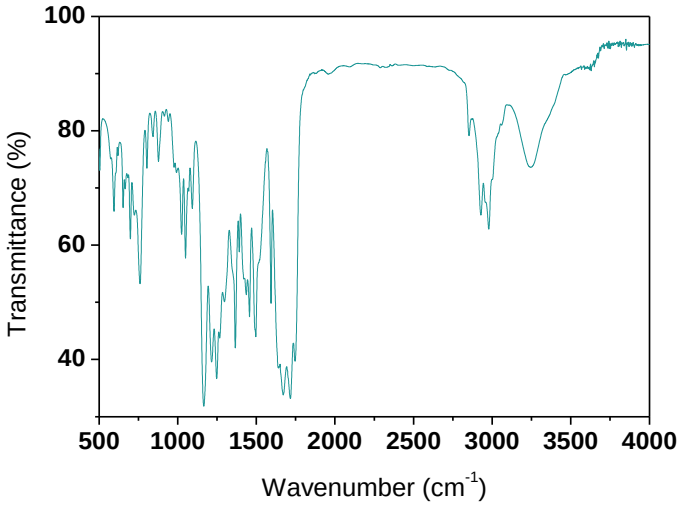


Figure S55. FT-IR spectra of (**5**)

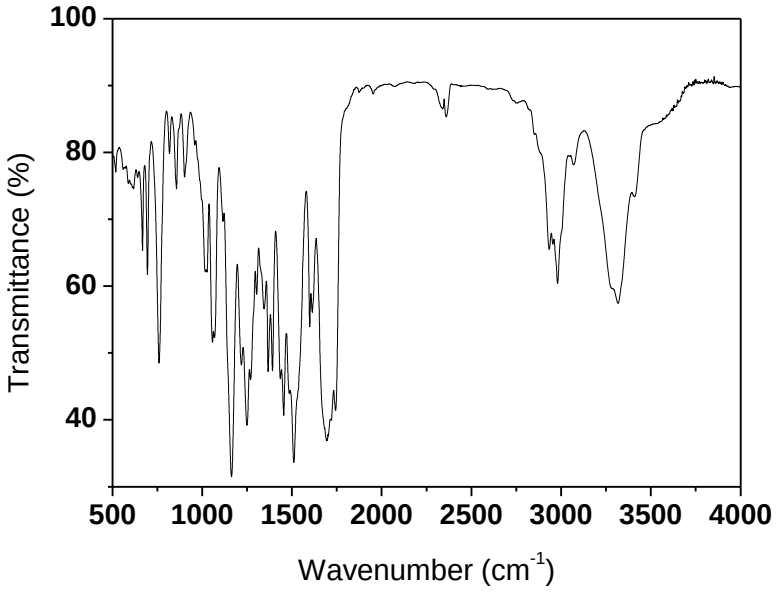


Figure S56. FT-IR spectra of (**7a**)

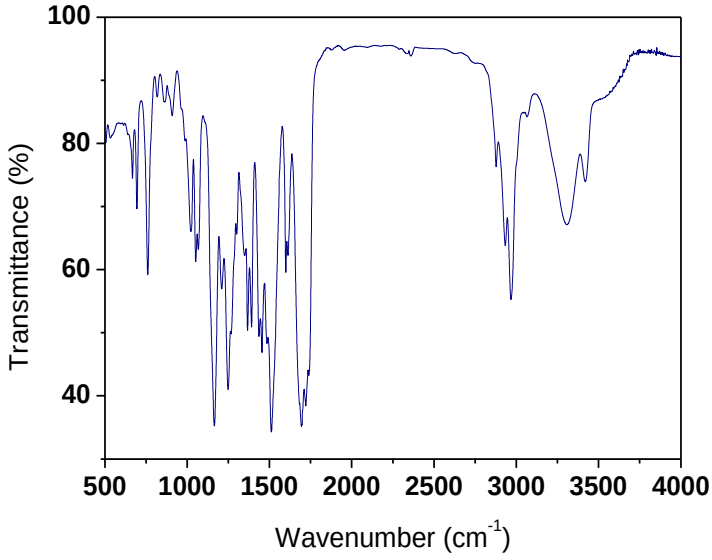


Figure S57. FT-IR spectra of (**7b**)

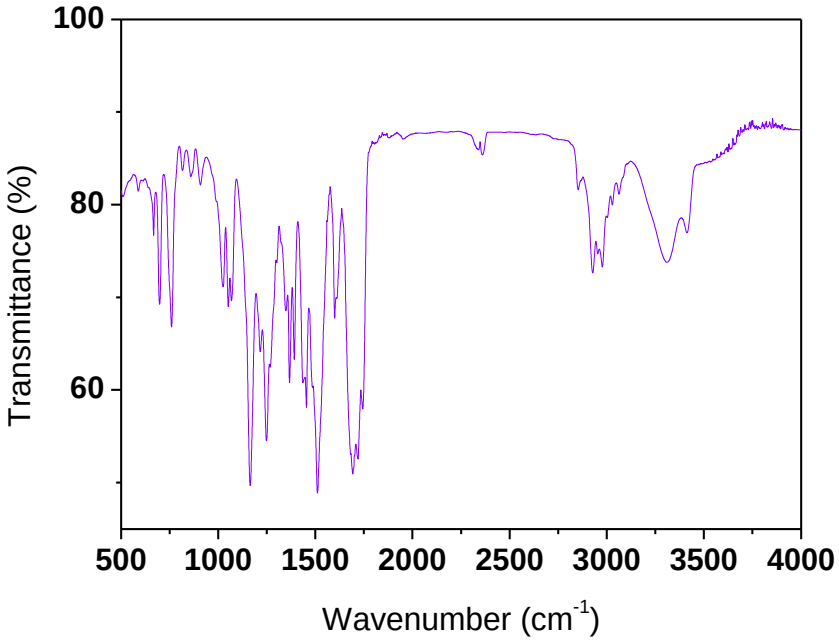


Figure S58. FT-IR spectra of (**7c**)

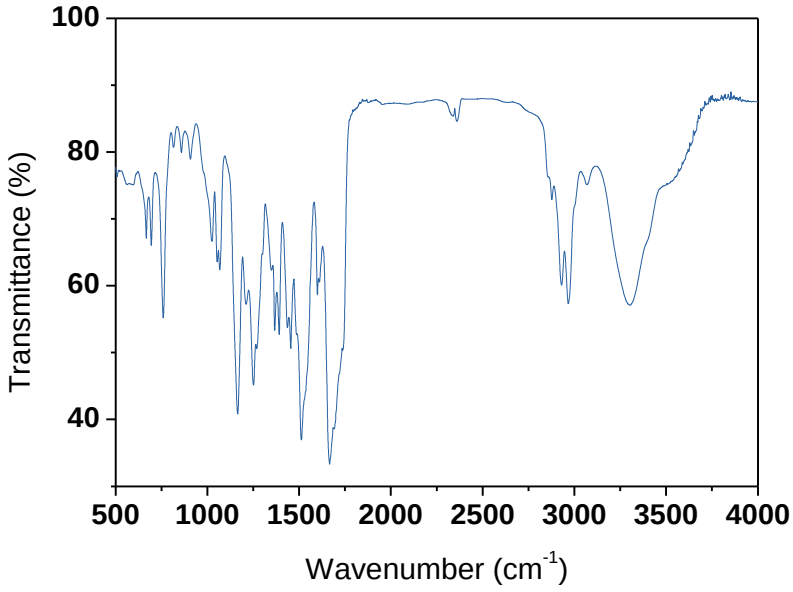


Figure S59. FT-IR spectra of (**7d**)

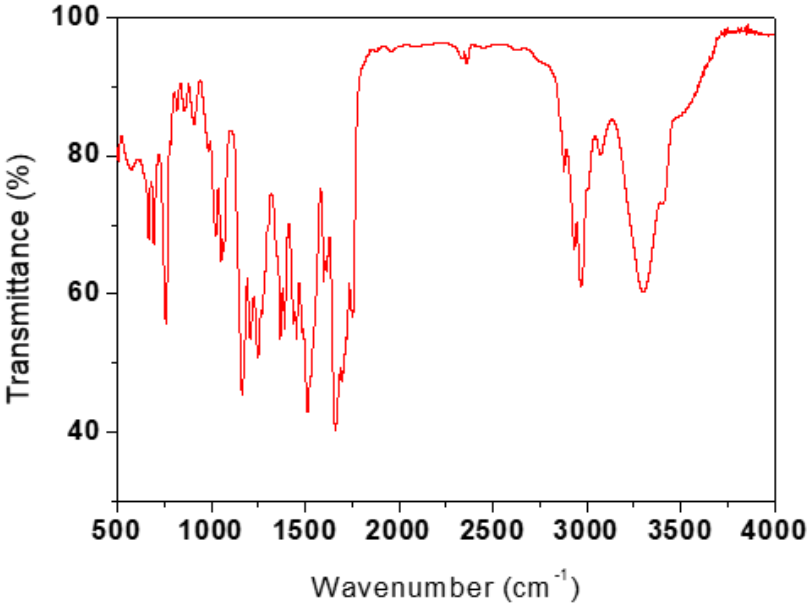


Figure S60. FT-IR spectra of (**7e**)

CHAPTER 4

PART-B

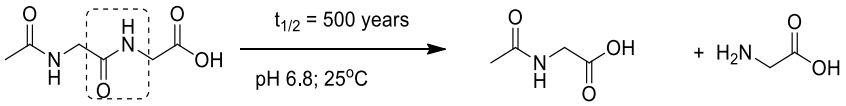
Cleavable Amide Bond: Mechanistic Insight of Cleavable *4-Aminopyrazolyloxy Acetamide/Acid/Ester* at low pH

Chapter 4B. Cleavable Amide Bond: Mechanistic Insight of Cleavable *4-Aminopyrazolyloxy Acetamide/Acid/Ester* at low pH

4B.1 Introduction

Natural amides are extremely stable, and their half-life for spontaneous hydrolysis is estimated to be 350 to 600 years at neutral pH and room temperature (for example Figure 4B.1).[1-3] The *N*-acetamide bond of peptides is extremely stable even under the strongest organic acid TFMSA. The carbonyl group of amides are poor electrophiles because of the typical resonance stability. Though amides can readily cleaved/hydrolyzed by enzymes such as proteases, it is difficult to cleave the C-N bond of an amide selectively using synthetic chemistry. The cleavage/hydrolysis of amide bonds usually requires heating under strongly acidic or basic conditions. However, the cyclic amides, lactams, are easily cleavable as compare to the linear amides because of ring strained amides.[4, 5] A large number of the highly strained lactams are synthesized and found as cleavable amide bond under mild conditions. Since the C=O bond of reactive lactams amide is arising to the strong electrophile as "ketonic" carbonyl because of the resonance decoupling through N-C=O torsion. Brown and co-workers have mechanistically confirmed that resonance decoupling induces the remarkable enhancement in hydrolysis rate by direct nucleophilic attack at the carbonyl of strained amide.[6-8] The most twisted amide trimethylazatricyclodecanone is highly strained lactam with large resonance decoupling through N-C bond, and found as readily hydrolysable. This twisted lactam also shows the dual reactivity of amide as nucleophilic character of amine and electrophilic nature of carbonyl. Booker-Milburn and co-workers have reported the solvolysis/substitution reaction of acyclic synthetic amides at room temperature and neutral conditions via the formation of ketene intermediates.[9] In literature, the cleavage of few natural

amide bonds is precisely demonsterated through the formation of thiazolinone derivatives (by Edman degradation), oxazolinium intermediates and chemical modification of the amide backbone.[10-13] Most importantly, Edman degradation method for cleavage of amide bond is the foundation for protein sequencing technology.[10] The presences of Zinc species in active sites of metalloproteases have inspired synthetic chemists to cleave amide bond using Zn^{2+}. Mashima and co-workers have reported the cleavage of amides bearing a β-hydroxyethyl group under mild conditions with Lewis acid $Zn(OTf)_2$.[14] Kostic and co-workers have prepared an artificial peptidase which cleavage of the sequence specific amide bonds as Pro-Met/Pro-His segments of peptide under mild conditions with Pd-catalyst.[15] Hauk, Garg and co-workers have established the conversion of amide functional group into ester group by cleaving C-N bond of amide using Ni-catalyst.[16-18] The conversion of an amide into another amide with different amine group, transamidation, is important chemical transformation reaction for ligation or removal of amino functionalized chemical moieties at carboxylate functionalized molecules.[19-21] Gellman, Stahl and co-workers have explored the amide groups for transamidation reaction, between secondary amines and tertiary amines, using Zr/Hf-catalyst.[19] Other Metal catalyzed and metal free transamidation reactions are also reported.[20-24]



(*J. Am. Chem. Soc.* **1996**, *118*, 6105-6109)

Figure 4B.1. Stability of acetamide bond in glycine dipeptide.

Recently, our lab established the transamidation reaction in tropolonylalkylglycinate peptide derivatives with different amino acid/peptides. Our lab, also found the instability of amide bonds containing *N*-Tropolonylaminoalkylglycinate (*Trag*) in short peptides under mild acidic

conditions. The formation of stable *tropolonium cations* followed by the formation of ketene intermediate is main cause of the cleavage of *Trag*-amide bonds.[25-27]

4B.2 Objective of this work

We have explored the syntheses and supramolecular self-assembly studies of pyrazolone containing amino acid analogues in chapter 4A. One of those analogue is 4-aminopyrazolonyl acetate has been conjugated with natural amino acid/short peptides for understanding the role of pyrazolonyl moiety in conformational changes of peptides.[28] In extension to that work to synthesize homo peptides of aminopyrazolone, we encountered an unusual amide hydrolysis (Figure 4B.2). This is a rare chemical event in peptide bond. Herein we studied the syntheses and chemical instability of amide/acid/ester bond containing aminopyrazolonyl acetyl residue under mild acidic conditions at room temperature by NMR and Mass analyses.

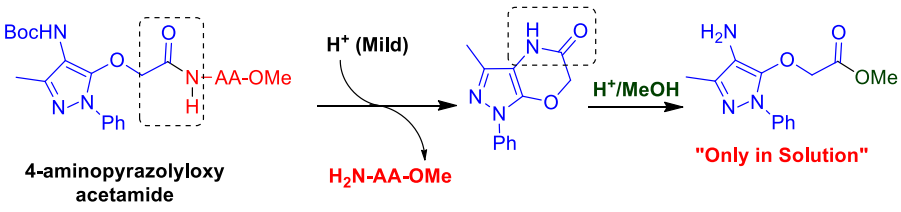


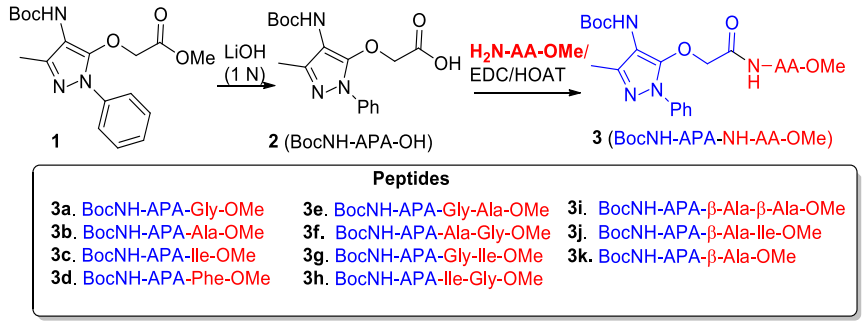
Figure 4B.2. Chemical events observed during the synthesis APA homopeptides.

4B.3 Results and Discussion

We used previously synthesized 4-aminopyrazolonyl amino acid (APA) ester derivative (**1**) for synthesis of conjugated peptides. Ester (**1**) was hydrolyzed into 4-aminopyrazolonyl amino acid, APA acid (**2**) with LiOH (1 N in THF), and then coupled with amino group of various natural amino acids/peptides (H_2N-AA-OMe) under peptide coupling reaction conditions which produced conjugated peptides **3** (APA-AA-OMe). Natural α/β-amino acid ester derivatives of Gly-OMe, Ala-OMe, Ile-OMe, and Phe-OMe produced unnatural respective conjugated dipeptides as

BocNH-APA-Gly-OMe (**3a**), BocNH-APA-Ala-OMe (**3b**), BocNH-APA-Ile-OMe (**3c**), BocNH-APA-Phe-OMe (**3d**), BocNH-APA-β-Ala-OMe (**3k**). Further natural α/β-amino acid dipeptide esters, Gly-Ala-OMe, Ala-Gly-OMe, Gly-Ile-OMe and Ile-Gly-OMe gave respective unnatural conjugated tripeptides as BocNH-APA-Gly-Ala-OMe (**3e**), BocNH-APA-Ala-Gly-OMe (**3f**) BocNH-APA-Gly-Ile-OMe (**3g**), and BocNH-APA-Ile-Gly-OMe (**3h**), BocNH-APA-β-Ala-β-Ala-OMe (**3i**), BocNH-APA-b-Ala-Ile-OMe (**3j**) as shown in Scheme 4B.1. The characterization data (^{1}H-/^{13}C-NMR/Mass) of all synthesized peptides are provided in Appendix.

Scheme 4B.1. Synthesis of *N*-Boc-aminopyrazolonyl amino acid and their peptides.



For further elongation of peptides at *N*-terminal of hybrid peptides **3a-3k**, the removal of *N*-Boc group was necessary. The removal of Boc group of *N*-Boc-aminopyrazolonyl amino acid containing peptides **3a-3k** was attempted with 20% TFA in DCM; unfortunately, we could not achieve desired aminopyrazolonyl amino acid peptides. Surprisingly, we noticed the cleavage of amide bond by ESI-Mass studies. To ensure this cleavage of amide bond, we performed the similar acid treatment reaction with other hybrid di-/tri-peptides, **3b-3k** and then characterized their isolated products by NMR and Mass analyses. Irrespective of α-amino acid sequence, aminopyrazolyl amino acid containing amide bonds in peptides **3a-3k** were cleaved and produced the cyclized product as lactam **4** (Figure 4B.3). In case of hybrid tripeptides **3e-3j**, other natural amide bonds, dipeptides (between two α-amino acids) were stable under acidic conditions (20%

TFA in DCM). Along with lactam **4**, we also observed *N*-trifluoroacylation with 20% TFA in ESI-MS, this side reaction was minimized/avoided by using 10% TFA or 4.0 N HCl in 1,4-dioxane (Appendix, Figure S13-S15). These results strongly suggest the cleavage of specific amide bond containing 4-aminopyrazolonyl amino acid residue at C-terminal of α-amino acid/peptides.

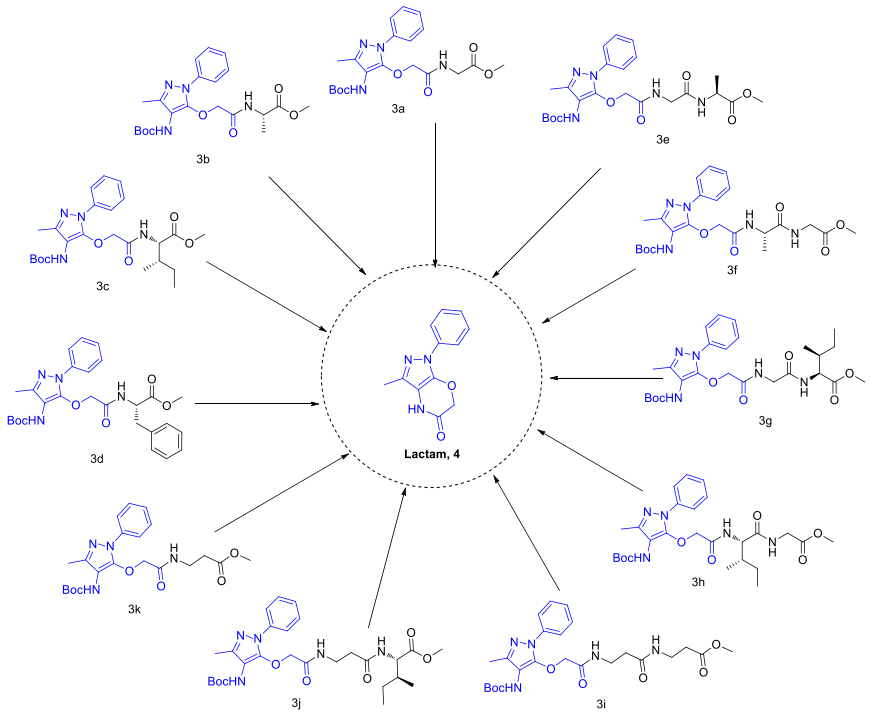
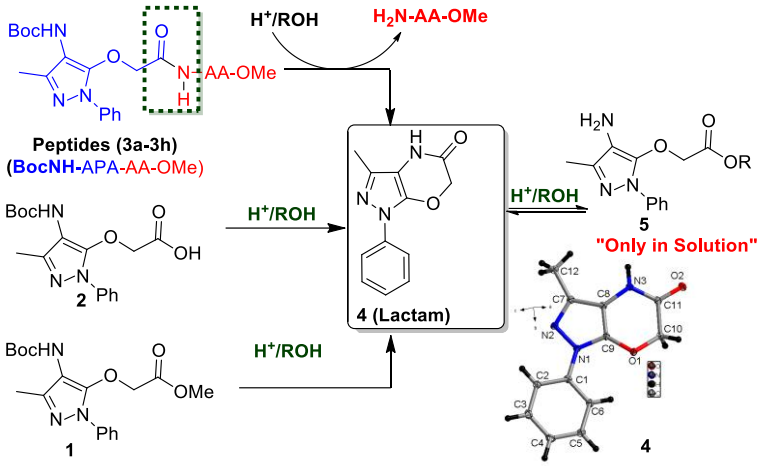


Figure 4B.3. Overview of APA amide bond cleavage and formation of lactam **4** from different APA-peptides (**3a**-**3k**) under mild acidic conditions.

Further, we attempted to remove the Boc group of *N*-Boc-aminopyrazolonyl amino acid ester **1** and its carboxylic acid **2** under similar acidic conditions (10-20% TFA in $CDCl_3$) for conjugation of aminoprazolonyl moiety at C-terminal of above mentioned di-/tri-peptides. Surprisingly, we failed to generate amino functionalized amino ester (**1-NH$_2$**) and acid (**2-NH$_2$**). However, we isolated a new product and characterized by NMR and ESI-Mass analyses. Our NMR and Mass results supported that both ester **1** and its acid **2** gave the common cyclized product as lactam **4**

under mild acidic conditions (Scheme 4A.2). The synthesis of that lactam **4** is extremely unusual from ester **1**/acid **2**/amide **3** under acidic conditions at room temperature which are also unusual in peptide chemistry. The isolated crystalline product, lactam **4** was also studied by single crystal X-ray diffractometer. The X-ray studies confirmed the structure of lactam **4**, the ORTEP diagram is shown in Scheme 4A.2. The cif file of crystal lactam **4** has deposited to CCDC with number CCDC 1890879.

Scheme 4A.2. Reaction of APA derivatives under mild acidic conditions.



The progress of amide cleavage reactions under acidic conditions is studied by ^{1}H-NMR technique. We recorded the time dependent ^{1}H-NMR spectra of peptide **3a** under 10% TFA in deuterated solvent $CDCl_3$. The array of the ^{1}H-NMR spectra with time is provided in Figure 4B.4, a full spectrum is provided in Appendix Figure S25. Most importantly, the gradual disappearance of NMR signal, singlet at δ 5.0 (ppm), is noticed in peptide **3a** with respect to time under that acidic conditions. A transient increase and decrease of signal at δ 4.5 (ppm) is also observed.

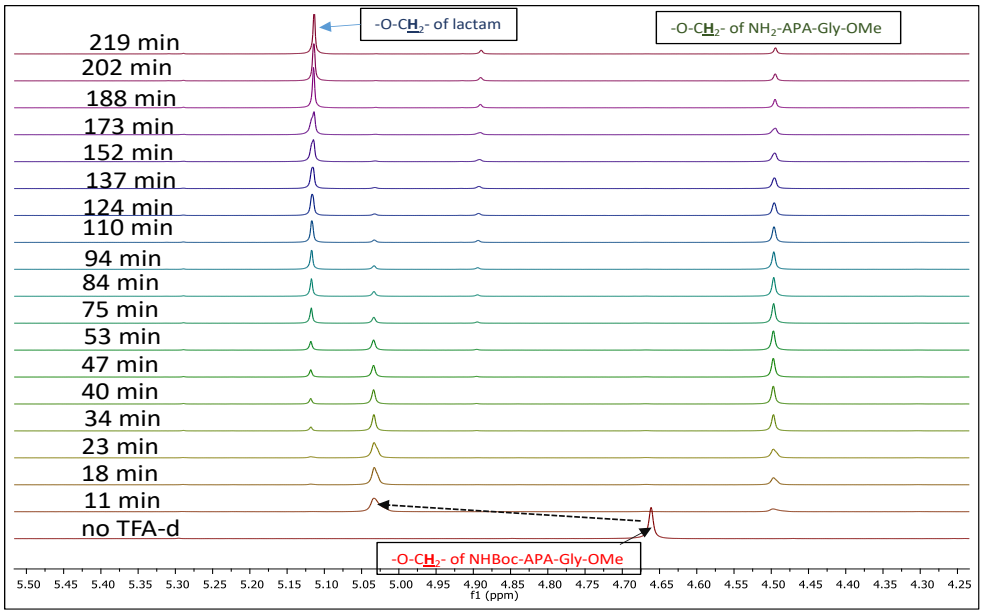


Figure 4B.4. Time dependent ^{1}H-NMR spectra of compound **3a** in 10% TFA/$CDCl_3$

Similar time dependent NMR experiments were performed under acidic conditions with another peptide **3e**. Their NMR data are provided in Appendix (Figure S27). After the NMR study the same samples were subjected to ESI-MS in solvent CH_3CN (neat). Their mass spectra exhibit the mass (m/z) peak at 230, in solvent CH_3CN, for both peptides **3a**/**3e**, which belongs to mass of lactam **4** as $[M+H]^+$ Appendix (Figure S26 & S28). These results (NMR and Mass) strongly support the involvement of α-CH_2 of aminopyrazolonyl residue in the cleavage of peptide **3a**/**3e** under acidic condition (10% TFA in $CDCl_3$) via the formation of stable lactam **4**. Further, we performed the similar NMR and Mass experiments with *N*-Boc-aminopyrazolonyl amino acid ester **1** and its carboxylic acid **2**. Their NMR spectra, NMR profiles and mass spectra are provided in the Appendix. These results suggest that the formation of lactam **4** is common for both aminopyrazolonyl ester **1** and its carboxylic acid **2** under mild acidic conditions (10% TFA in $CDCl_3$). We studied the ^{1}H-NMR kinetics of amide **3a**/acid **2**/ester **1** hydrolysis in $CDCl_3$. The relative concentrations were calculated in mole fractions (*x*) from relative integrations of –*O*-C**H**$_2$-

peak of amide **3a**/acid **2**/ester **1** with gradually increasing lactam **4** with time. The plots of mole fraction (x) with time for APA amide **3a** and APA acid **2** is provided in Figure 4B.5, and APA ester **1** is provided in Appendix (Figure S35A). We have also generated simulated kinetic plots from the sequential first-order reaction model with COPASI software program using experimentally obtained k_1 and k_2 values and then fitted with our experimental kinetic plots (Figure S41-S43).[29, 30]

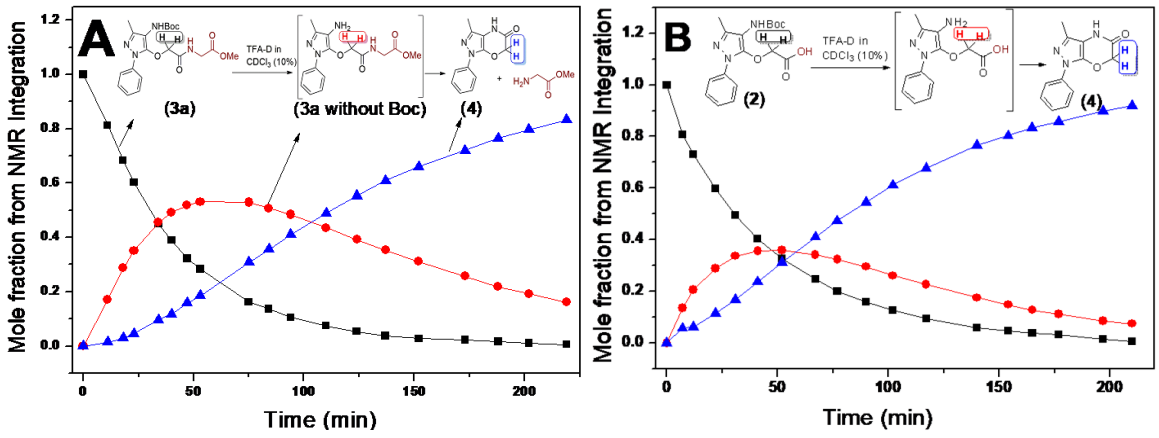


Figure 4B.5. Time dependent ^{1}H-NMR studies of APA peptide **3a** (A) and APA acid **2** (B) in solvent $CDCl_3$ with 20 mM concentrations.

From the plots of concentration versus time, a transient increase and decrease of intermediate is also observed, this suggest that the cyclization of amide **3**/acid **2**/ester **1** is proceeding through an intermediate. After seeing these plots, we propose that compound **1**, **2**, & **3a** are following consecutive first order kinetic pathway and the rate laws are given by equations 1-3 in Appendix. The plots of integral mole fraction versus time, (ln(x) vs time) are provided in Appendix (Figure S33-S35) and gave a straight line with negative slope which are consistent with literature reports for consecutive reaction pathway.[31-33] The rate constants k_1, k_2 and their respective half-lives ($t_{1/2}$) are calculated from these integral plots and provided in Table 4B.1. The observed rate constants

and half-lives from Table 4B.1 suggest that the acid **2** is hydrolysed relatively faster than amide **3a** followed by ester **1**.

Table 4B.1. Kinetic data of 4-aminopyrazolyloxy derivatives from ^{1}H-NMR spectrum.

Compound	k_1/min^{-1}	k_2/min^{-1}	k_2/k_1	$t_{1/2}$ /min (for k_1)	$t_{1/2}$ /min (for k_2)
3a (Amide)	2.28×10^{-2}	0.83×10^{-2}	0.36	30.34	83.00
2 (Acid)	1.97×10^{-2}	1.09×10^{-2}	0.55	35.09	63.35
1 (Ester)	1.95×10^{-2}	0.78×10^{-2}	0.4	35.43	88.74

Interestingly, during the hydrolysis of amide bond in peptides **3a-3h**, we observed methanolysis of amide bond in lactam and exist in methyl ester form in presence of MeOH (Appendix, Figure S13-S18). To further understanding the reactivity of this amide bond, the lactam **4** is monitored in several aspects using mass spectroscopy (Appendix, Figure S37), from these studies it is apparent that, acid and MeOH both are needed for the amide bond cleavage in lactam **4**. Moreover, there is no complete disappearance of lactam in any of the cleavage conditions implemented during these studies. This cleavage is also monitored by time dependent ^{1}H-NMR in CD_3CN, CD_3OD and PTSA (~ 3.0 equivalents) at 293 K, from the array of spectra, the reaction proceeds with the formation of hydrolysed methyl ester and remains constant after ~500 min without complete conversion of lactam **4**. The relative concentrations were calculated in mole fractions (x) from relative integrations of –*O*-C**H**$_2$- peak of lactam with gradually increasing methyl ester **4-OCD$_3$** with time, the resultant plots indicate the existence of equilibrium (Figure 4B.6) and the rate law for this equilibrium is given by equation 4 in Appendix. The equilibrium constant (K) from the plot of mole ratio vs time has a value of 2.72 (Appendix Figure S36B).

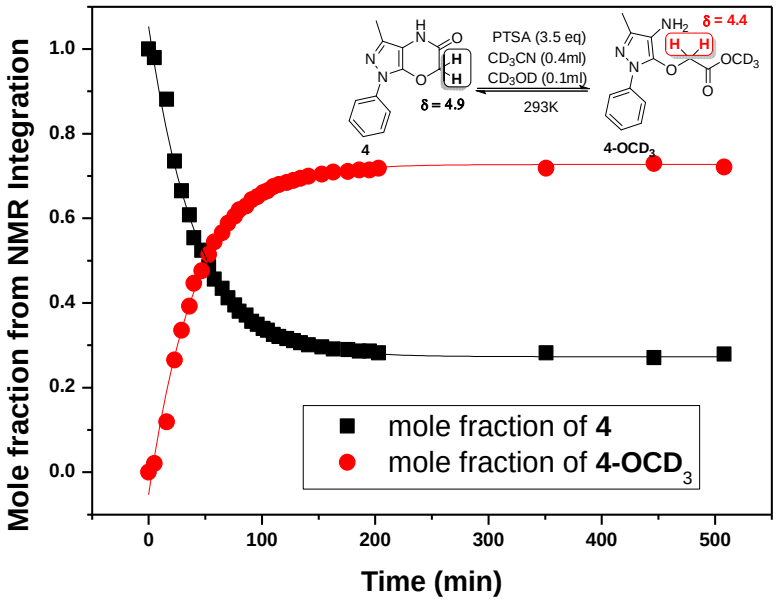


Figure 4B.6. Time dependent ^{1}H-NMR studies of lactam **4** in CD_3CN, CD_3OD and PTSA.

After NMR study, the same sample is subjected to ESI-MS, we observed the mass of both deuterated methyl ester (m/z 265) and lactam (Figure 4B.7). Importantly, we were not able to isolate this methyl ester. From these studies, we assume that the lactam **4** and its methyl ester **4-OCH$_3$** forms are in equilibrium in solution state in presence of acid and MeOH.

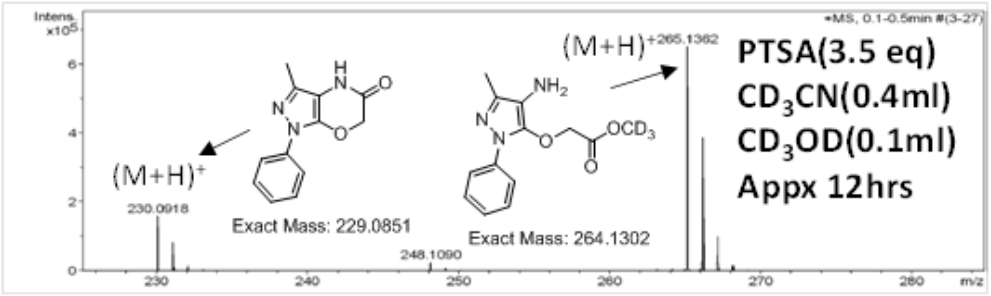


Figure 4B.7. Methanolysis of lactam **4**. Even after 12 h the lactam is not completely disappeared.

Since aminopyrazolone is aromatic compound and UV active in nature. Thus we attempted to monitor the cleavage of lactam **4** into its amino ester derivative **4-OCH$_3$** under acidic condition by UV studies. We found that the UV spectrum of pure lactam **4** exhibited two characteristic peaks

at wavelength (λ) 240 nm and 280 nm under neutral pH conditions (Appendix, Figure S40). Then we recorded UV spectra of lactam **4** at different time intervals of times under acidic conditions (0.1 N HCl in MeOH) which exhibit the significant depletion of absorption intensity of lactam **4** at $\lambda_{240\ nm}$ and $\lambda_{280\ nm}$ (Figure 4B.8). Such changes in UV spectra strongly support the formation of new species under that mild acidic condition.

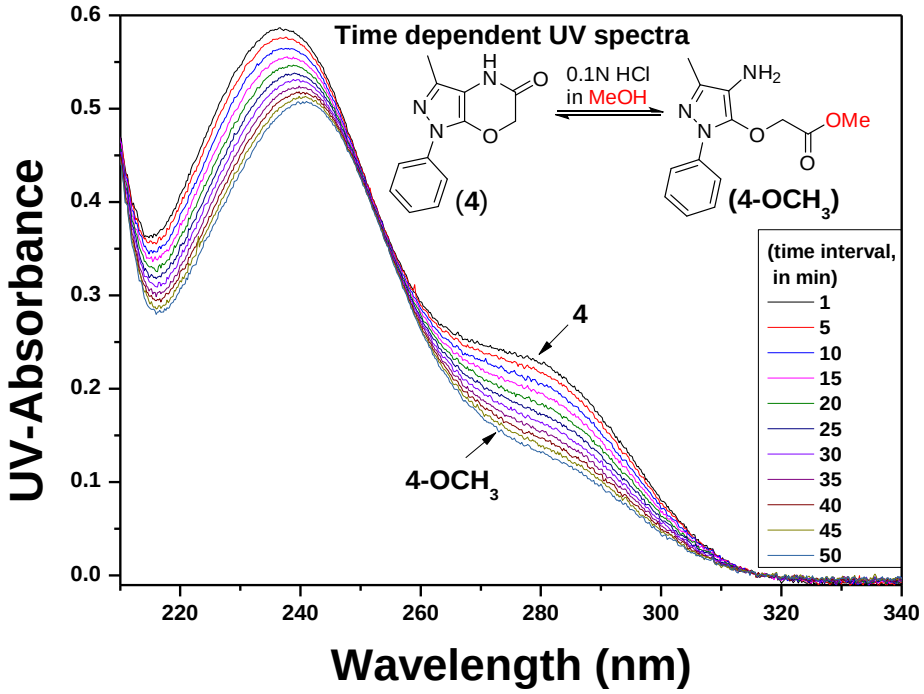


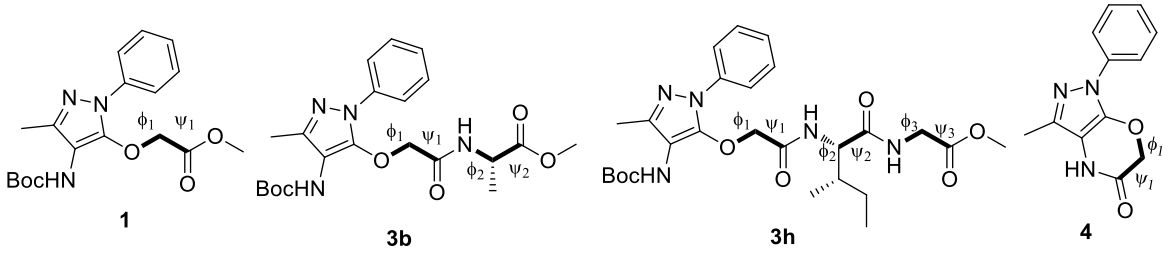
Figure 4B.8. Time dependent UV-Spectra of lactam **4** under acidic conditions, time intervals in min. (right panel).

In chapter 4A, we have confirmed an unusual *syn* conformation between -***O***-***C***H_2–***C***O-***N***H- in crystal structures of peptides **3b**/**3h** and a weak intramolecular hydrogen bonding between *H*-atom of amide (-CONH-) and *O*-atom (–***O***-*C*H_2-).[28] Similar *syn* conformation is also noticed in crystal structure of APA ester **1**. Crystal structural analyses of APA peptides **3b**/**3h** reveal that the APA amide bond is twisted with tau (τ) ~ 5.5^o (Table 4B.2).

Table 4B.2. Structural parameters of APA-peptides (**3b**,**3h**) lactam (**4**) and ester(**1**).

Compound	phi (ϕ)	psi (ψ)	tau (τ)	Sum of bond angle at "N"	Sum of bond angle at "C=O"	C-N Bond length (in Å)	C-O Bond length (in Å)
3h	ϕ_1 (144.345) ϕ_2 (-135.735) ϕ_3 (-142.304)	ψ_1 (26.561) ψ_2 (139.39) ψ_3 (-169.368)	5.538 0.468	360 360	359.99 360	1.305 1.360	1.196 1.182
3b[a]	ϕ_1 (133.536) ϕ_2 (-139.577)	ψ_1 (-12.46) ψ_2 (58.43)	5.5389(A) & 3.847 (B)	359.093 (A) & 359.663(B)	359.992 (A) & 359.99 (B)	1.333 1.322	1.231 1.228
1	ϕ_1 (-165.135)	ψ_1 (3.889)					
4	ϕ_1 (21.52)	ψ_1 (-12.303)	3.985	360	359.98	1.334	1.233

[a]unit cell has two molecules



We also noticed intramolecular hydrogen bonding between *H*-atom (-COOH-) and *O*-atom (–***O***-*CH*$_2$-) in APA acid **2** from NMR studies (Appendix, Figure S1-S4). The strong synergetic effects, such as electronegative –***O***-*CH*$_2$-, twisted amide bond (τ ~ 5.5^o) and unfavourable *syn* conformation (ψ) weaken the stability of that amide bond. As a resultant, the favourable intramolecular attack of APA amine (6-exo-trig) led to the cleavage of that amide bond under mild acidic conditions (Figure 4B.9).

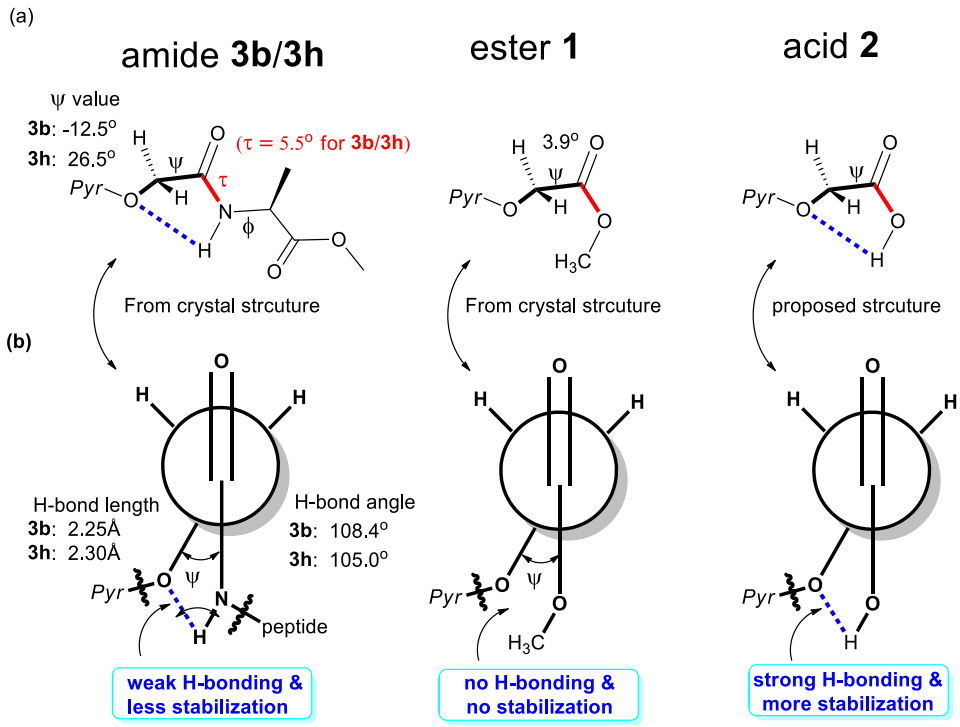


Figure 4b.9. (a) Structures of amide **3b/3h**, ester **1**, acid **2**; (b) Their Newman Projection.

The Plausible mechanism as could be as: removal of Boc group of aminopyrazolonyl derivatives **1/2/3** occurs under acidic conditions (10-20% TFA in DCM/CH_3CN) and generates the protonated amine intermediates (**a**) which may protonate the intramolecular carbonyl group of ester/acid/amide derivatives as intermediates (**b**). Subsequently the amine nucleophile possibly attacks at the protonated carbonyl electrophile as typical intramolecular nucleophilic addition-elimination reaction to facilitate the formation of cyclic amide (lactam) derivative **4** (Figure 4B.10). Most importantly, this lactam is stable in solvent free conditions and its structure is confirmed by single crystal X-ray analyses. The stability of this lactam **4** is further investigated under protic nucleophilic solvent (MeOH) under acidic conditions. Importantly, this lactam **4** is unstable in nucleophilic solvents (H_2O/MeOH) under acidic conditions and proceeds to amide

hydrolysis possibly due to protonation of amide bond *N*-atom under acidic conditions. As resultant, lactam **4** hydrolyzes into acid/ester derivatives (**f**) with respective solvents H_2O/MeOH.

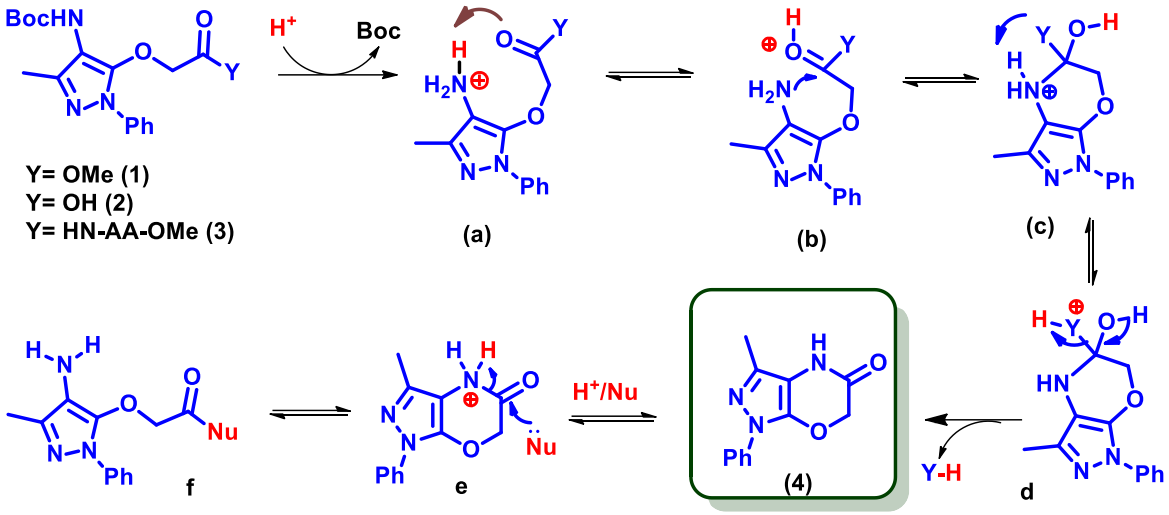


Figure 4B.10. Plausible mechanism of amide hydrolysis.

4B.4 Conclusion

In summary, 4-aminopyrazolyloxy acetate was successfully incorporated at *N*-terminal of natural di-/tri-peptides aminopyrazolyloxy acetamide derivatives which are hydrolysable under mild acidic conditions such amide hydrolyses are rare. This unique character of amide bond is elaborated by NMR/UV/Mass/X-ray studies which helped to demonstrate the mechanism of reactions. Since aminopyrazoly rings and its derivatives are chromophores. Thus 4-aminopyrazolyloxy acetate could be employed in protection of free amine via acid sensitive chromophoric amide bond unlike carbamate group of other protecting groups.

4B.5 Experimental section

General methods**.** All required materials were obtained from commercial suppliers and used without any further purification. Dimethylformamide (DMF) was distilled over calcium hydride. Reactions were monitored by thin layer chromatography, visualized by UV and ninhydrin. Column chromatography was performed in 230-400 mesh silica. Mass spectra and HRMS were obtained from Bruker micrOTOF-Q II Spectrometer. ^{1}H NMR, ^{13}C NMR, were recorded on Bruker AV-400 or 700 MHz at 298 K. ^{1}H and ^{13}C NMR chemical shifts were recorded in ppm downfield from tetramethylsilane or residual solvent peak. Splitting patterns are abbreviated as: s, Singlet; d, doublet; dd, doublet of doublet; t, triplet; q, quartet; dq, doublet of quartet; m, multiplet.

Experimental procedure. Experimental procedure and their characterization data for compounds **1**, **3b-3d**, **3g-3h** are previously reported from our lab.[28]

N-phenyl, 4-aminopyrazolyloxy acetic acid, BocNH-APA-OH (**2**). Compound **1** (1.26 g, 3.52 mmol) is dissolved in THF (25 mL) and cooled to 0 ^{o}C, then added aqueous LiOH (1 M, 25mL) and at the same temperature stirred for 20 min. The solvents are evaporated under vacuum to half of its volume and adjusted the pH to 6-7 with 1 M HCl. The resulting aqueous layer extracted with EtOAc (3*30 mL) and dried over anhydrous Na_2SO_4 and concentrated under vacuum. The product is precipitated with chloroform and hexane, dried the solid to produce 1.0 g (90%) of title compound. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.90 (OH, s, 1H), 7.63 (d, *J* = 7.8 Hz, 2H), 7.40 (t, *J* = 7.8 Hz, 2H), 7.27 (dd, *J* = 6.5, 4.7 Hz, 1H), 6.02 (NH, s, 1H), 4.67 (s, 2H), 2.15 (s, 3H), 1.49 (s, 9H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 172.6, 158.3, 147.4, 146.9, 138.2, 128.9, 126.7, 122.5, 101.2, 82.9, 67.8, 28.3. HRMS (ESI-TOF) m/z: $[M + H]^+$ Calcd for $C_{17}H_{22}N_3O_5$ 348.1554; Found 348.1567.

General experimental peptide coupling procedure for compounds, **3a-3h**:

Compound **2**, corresponding amine (TFA or HCl salt form) (1.5 equivalent), HOAT (1.5 equivalent) was dissolved in Dry DMF (1.5 M). After stirring for 10 min, N-methyl morpholine (3 equivalent) was added drop wise and cooled the temperature to 0 ^{o}C and added EDC.HCl (1.5 equivalent). After 20 min allowed to warm to rt followed by heating at 60 ^{o}C for 8 h. The crude reaction mixture was evaporated under reduced pressure. The resultant crude was purified by column chromatography with MeOH in CH_2Cl_2 (1-3%).

methyl 2-(2-((4-((tert-butoxycarbonyl)amino)-3-methyl-1-phenyl-1H-pyrazol-5-yl)oxy)acetamido)acetate, *BocNH-APA-Gly-OMe* (**3a**): R_f 0.25 (3% MeOH in CH_2Cl_2); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.58 (d, *J* = 7.8 Hz, 2H), 7.44 (t, *J* = 7.9 Hz, 2H), 7.30 (t, *J* = 7.4 Hz, 1H), 6.94 (t, 1H), 6.15 (s, 1H), 4.65 (s, 2H), 4.07 (d, *J* = 5.7 Hz, 2H), 3.77 (s, 3H), 2.19 (s, 3H), 1.48 (s, 9H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.0, 167.8, 155.1, 147.1 (two peaks merged), 138.1, 129.2, 126.9, 122.6, 102.6, 80.6, 70.7, 52.5, 40.5, 28.2, 11.8. HRMS (ESI-TOF) m/z: $[M + H]^+$ Calcd for $C_{20}H_{27}N_4O_6$ 419.1925; Found 419.1914.

(*S*)-methyl-2-(2-(2-((4-((tert-butoxycarbonyl)amino)-3-methyl-1-phenyl-1H-pyrazol-5-yl)oxy)acetamido)acetamido)propanoate *BocNH-APA-Gly-Ala-OMe* (**3e**): R_f 0.18 (3% MeOH in CH_2Cl_2); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.58 (d, *J* = 7.7 Hz, 2H), 7.42 (t, *J* = 7.6 Hz, 2H), 7.30 (d, *J* = 7.3 Hz, 1H), 7.16 (s, 1H), 6.69 (s, 1H), 6.48 (s, 1H), 4.63 (s, 2H), 4.62 – 4.53 (m, 1H), 3.98 (d, *J* = 5.3 Hz, 2H), 3.74 (s, 3H), 2.18 (s, 3H), 1.79 (d, *J* = 7.0 Hz, 2H), 1.46 (s, 9H), 1.41 (d, *J* = 7.1 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 173.3, 168.3 (two peaks merged together), 155.3, 147.2 (two peaks merged together), 138.1, 129.1, 126.9, 122.5, 102.8, 80.4, 70.7, 52.5, 48.2, 42.0, 28.2, 17.9, 11.9. HRMS (ESI-TOF) m/z: $[M + H]^+$ Calcd for $C_{23}H_{30}N_5O_{7s}Na$ 512.2116; Found 512.2121.

(*S*)-methyl 2-(2-(2-((4-((tert-butoxycarbonyl)amino)-3-methyl-1-phenyl-1H-pyrazol-5-yl)oxy)acetamido)propanamido)acetate *BocNH-APA-Ala-Gly-OMe* (**3f**): R_f 0.32 (3% MeOH in CH_2Cl_2); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.58 (d, *J* = 7.8 Hz, 2H), 7.44 (t, *J* = 7.8 Hz, 2H), 7.30 (t, *J* = 7.4 Hz, 1H), 6.91 (d, *J* = 6.8 Hz, 1H), 6.78 (s, 1H), 6.32 (s, 1H), 4.63 (s, 2H), 4.59 – 4.49 (m, 1H), 4.09 – 3.92 (m, 2H), 3.74 (s, 3H), 2.18 (s, 3H), 1.47 (s, 9H), 1.34 (d, *J* = 7.0 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 172.3, 170.3, 167.7, 155.3, 147.3, 147.2, 138.1, 129.3, 127.1, 122.7, 102.8, 80.6, 70.8, 52.4, 48.2, 41.2, 28.3, 17.9, 11.9. HRMS (ESI-TOF) m/z: $[M + H]^+$ Calcd for $C_{23}H_{31}N_5O_7$ 490.2296; Found 490.2295.

3-methyl-1-phenyl-4,6-dihydropyrazolo[3,4-b][1,4]oxazin-5(1H)-one, (**4**): 20% TFA in CH_2Cl_2 (5 mL) was added to **1** (0.50 g, 1.38 mmol) at room temperature and stirred for 1 h, then refluxed for an additional 1.5 h at 45 ^{o}C. The solvents were removed under vacuum and resulting pale orange residue was partitioned between EtOAc (25 mL) and saturated $NaHCO_3$ (10 mL). The aqueous layer was extracted with EtOAc (2 * 25 mL). The combined organic extracts were washed with saturated aqueous NaCl (10 mL), dried over anhydrous Na_2SO_4 and concentrated under vacuum. The residual viscous oil was purified via chromatography with 3% MeOH in CH_2Cl_2 to produce 0.293 g (92%) of **4** as a crystalline solid. *R_f 0.2* (0.3:9.7 MeOH/CH_2Cl_2); ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.51 (s, 1H), 7.72 – 7.58 (m, 2H), 7.45 (dd, *J* = 10.8, 5.2 Hz, 2H), 7.25 (t, *J* = 7.4 Hz, 1H), 4.84 (s, 2H), 2.15 (s, 3H). ^{13}C NMR (101 MHz, DMSO-d_6) δ 162.8, 140.2, 138.5, 136.0, 129.8, 126.3, 119.5, 106.5, 70.0, 12.2. HRMS (ESI-TOF) m/z: $[M + H]^+$ Calcd for $C_{12}H_{12}N_3O_2$ 230.0924; Found 230.0933. mp: 216-218 ^{o}C.

4B.6 Reference

1. A. J. Barrett, N. D. R., J. F. Woessner, Enzymes can of course efficiently orchestrate amide cleavage under physiological conditions. *Handbook of Proteolytic Enzymes, 2nd ed., Academic Press, New York* **2004**.
2. Radzicka, A.; Wolfenden, R., Rates of uncatalyzed peptide bond hydrolysis in neutral solution and the transition state affinities of proteases. *J. Am. Chem. Soc.* **1996,** *118*, 6105-6109.
3. Smith, R. M.; Hansen, D. E., The pH-rate profile for the hydrolysis of a peptide bond. *J. Am. Chem. Soc.* **1998,** *120*, 8910-8913.
4. Clayden, J.; Moran, W. J., The Twisted Amide 2-Quinuclidone: 60 Years in the Making. *Angew. Chem. Int. Ed. Engl.* **2006,** *45*, 7118-7120.
5. Szostak, M.; Aubé, J., Medium-bridged lactams: a new class of non-planar amides. *Org. Biomol. chem.* **2011,** *9*, 27-35.
6. Somayaji, V.; Brown, R., Distorted amides as models for activated peptide NC: O units produced during enzyme-catalyzed acyl transfer reactions. 1. The mechanism of hydrolysis of 3, 4-dihydro-2-oxo-1, 4-ethanoquinoline and 2, 3, 4, 5-tetrahydro-2-oxo-1, 5-ethanobenzazepine. *J. Org. Chem* **1986,** *51*, 2676-2686.
7. Slebocka-Tilk, H.; Brown, R., Effect of distortion on the hydrolytic reactivity of amides. 2. N-Pyramidalization: decomposition of N-benzoylaziridines in aqueous media. *J. Org. Chem.* **1987,** *52*, 805-808.
8. Bennet, A.; Somayaji, V.; Brown, R.; Santarsiero, B., The influence of altered amidic resonance on the infrared and carbon-13 and nitrogen-15 NMR spectroscopic

characteristics and barriers to rotation about the NC (O) bond in some anilides and toluamides. *J. Am. Chem. Soc.* **1991,** *113*, 7563-7571.

9. Hutchby, M.; Houlden, C. E.; Haddow, M. F.; Tyler, S. N.; Lloyd-Jones, G. C.; Booker-Milburn, K. I., Switching pathways: room-temperature neutral solvolysis and substitution of amides. *Angew. Chem. Int. Ed. Engl.* **2012,** *51*, 548-51.

10. Edman, P., Mechanism of the phenyl isothiocyanate degradation of peptides. *Nature* **1956,** *177*, 667.

11. Samaritoni, J. G.; Copes, A. T.; Crews, D. K.; Glos, C.; Thompson, A. L.; Wilson, C.; O'Donnell, M. J.; Scott, W. L., Unexpected hydrolytic instability of N-acylated amino acid amides and peptides. *The Journal of Organic Chemistry* **2014,** *79*, 3140-3151.

12. Nalbone, J. M.; Lahankar, N.; Buissereth, L.; Raj, M., Glutamic acid selective chemical cleavage of peptide bonds. *Organic letters* **2016,** *18*, 1186-1189.

13. Mahesh, S.; Tang, K.-C.; Raj, M., Amide bond activation of biological molecules. *Molecules* **2018,** *23*, 2615.

14. Kita, Y.; Nishii, Y.; Onoue, A.; Mashima, K., Combined catalytic system of scandium triflate and boronic ester for amide bond cleavage. *Adv. Synth. Catal.* **2013,** *355*, 3391-3395.

15. Parac, T. N.; Kostić, N. M., Effects of linkage isomerism and of acid− base equilibria on reactivity and catalytic turnover in hydrolytic cleavage of histidyl peptides coordinated to Palladium (II). identification of the active complex between Palladium (II) and the histidyl residue. *J. Am. Chem. Soc.* **1996,** *118*, 5946-5951.

16. Tobisu, M.; Nakamura, K.; Chatani, N., Nickel-catalyzed reductive and borylative cleavage of aromatic carbon–nitrogen bonds in N-aryl amides and carbamates. *J. Am. Chem. Soc.* **2014,** *136*, 5587-5590.

17. Hie, L.; Nathel, N. F. F.; Shah, T. K.; Baker, E. L.; Hong, X.; Yang, Y.-F.; Liu, P.; Houk, K.; Garg, N. K., Conversion of amides to esters by the nickel-catalysed activation of amide C–N bonds. *Nature* **2015,** *524*, 79.

18. Simmons, B. J.; Weires, N. A.; Dander, J. E.; Garg, N. K., Nickel-catalyzed alkylation of amide derivatives. *ACS Catal.* **2016,** *6*, 3176-3179.

19. Stephenson, N. A.; Zhu, J.; Gellman, S. H.; Stahl, S. S., Catalytic transamidation reactions compatible with tertiary amide metathesis under ambient conditions. *J. Am. Chem. Soc.* **2009,** *131*, 10003-10008.

20. Tamura, M.; Tonomura, T.; Shimizu, K.-i.; Satsuma, A., Transamidation of amides with amines under solvent-free conditions using a CeO 2 catalyst. *Green Chem.* **2012,** *14*, 717-724.

21. Baker, E. L.; Yamano, M. M.; Zhou, Y.; Anthony, S. M.; Garg, N. K., A two-step approach to achieve secondary amide transamidation enabled by nickel catalysis. *Nat. commun.* **2016,** *7*, 11554.

22. Acosta-Guzmán, P.; Mateus-Gómez, A.; Gamba-Sánchez, D., Direct Transamidation Reactions: Mechanism and Recent Advances. *Molecules* **2018,** *23*, 2382.

23. Vanjari, R.; Allam, B. K.; Singh, K. N., Hypervalent iodine catalyzed transamidation of carboxamides with amines. *RSC Adv.* **2013,** *3*, 1691-1694.

24. Zhang, F.; Li, L.; Ma, J.; Gong, H., MoS 2-Catalyzed transamidation reaction. *Sci. Rep.* **2019,** *9*, 2536.

25. Balachandra, C.; Sharma, N. K., Synthesis and conformational analysis of new troponyl aromatic amino acid. *Tetrahedron* **2014,** *70*, 7464-7469.

26. Balachandra, C.; Sharma, N. K., Instability of Amide Bond Comprising the 2-Aminotropone Moiety: Cleavable under Mild Acidic Conditions. *Org. Lett.* **2015,** *17*, 3948-3951.

27. Balachandra, C.; Sharma, N. K., Direct/Reversible Amidation of Troponyl Alkylglycinates via Cationic Troponyl Lactones and Mechanistic Insights. *ACS Omega* **2018,** *3*, 997-1013.

28. Bollu, A.; Sharma, N. K., Syntheses and Conformational Analysis of Aminopyrazolonyl Amino Acid (APA)/Peptides. *Eur. J. Org. Chem.* **2019,** *6*, 1286-1292.

29. Hoops, S.; Sahle, S.; Gauges, R.; Lee, C.; Pahle, J.; Simus, N.; Singhal, M.; Xu, L.; Mendes, P.; Kummer, U., COPASI—a complex pathway simulator. *Bioinformatics* **2006,** *22*, 3067-3074.

30. Shi, Q.; Mower, M. P.; Blackmond, D. G.; Rebek, J., Water-soluble cavitands promote hydrolyses of long-chain diesters. *Proc. Natl. Acad. Sci. U.S.A.* **2016,** *113*, 9199-9203.

31. House, J. E., Principles of chemical kinetics Academic Press. **2007**, 47– 53.

32. Koudriavtsev, A. B.; Linert, W., Keto-Enol Equilibrium from NMR Data: A Closer Look at the Laboratory Experiment. *J Chem Educ* **2009,** *86*, 1234-1237.

33. Nichols, M. A.; Waner, M. J., Kinetic and Mechanistic Studies of the Deuterium Exchange in Classical Keto-Enol Tautomeric Equilibrium Reactions. *J Chem Educ* **2010,** *87*, 952-955.

4B.7 Appendix

1. ^{1}H-/ ^{13}C-/ESI-MS/HRMS spectra of compound **2**

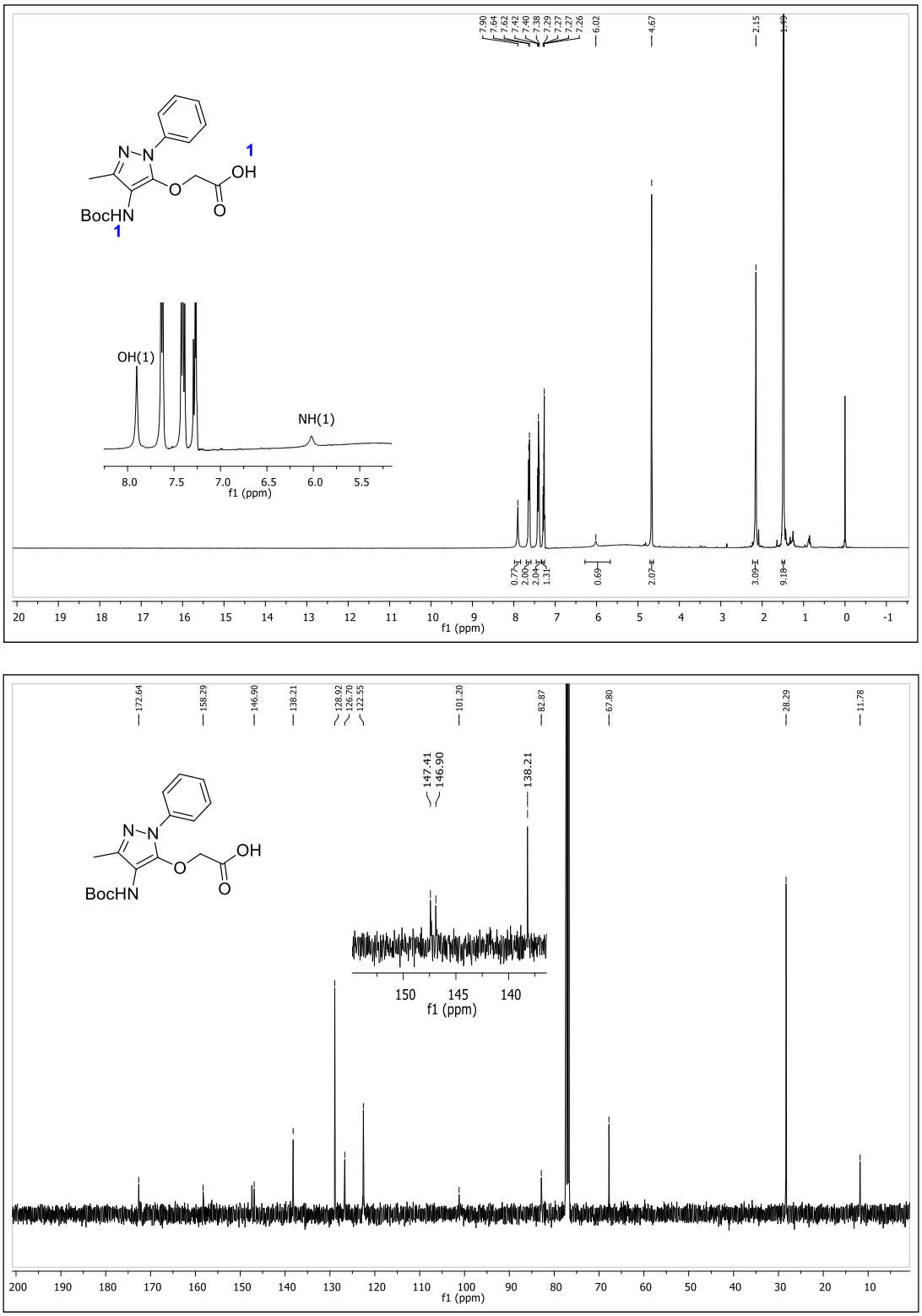


Figure S1. ^{1}H and ^{13}C NMR spectra of compound **2** in $CDCl_3$

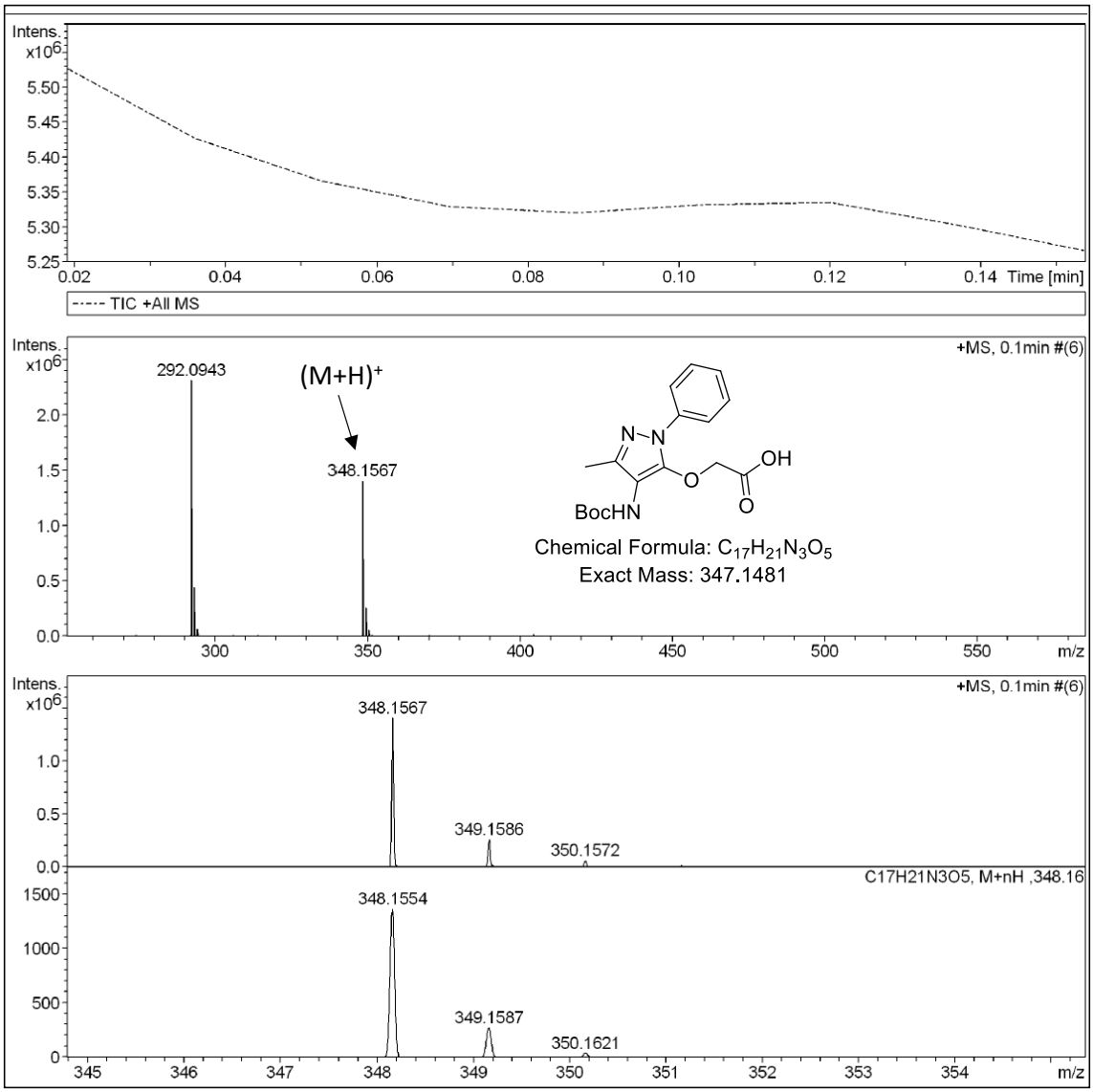
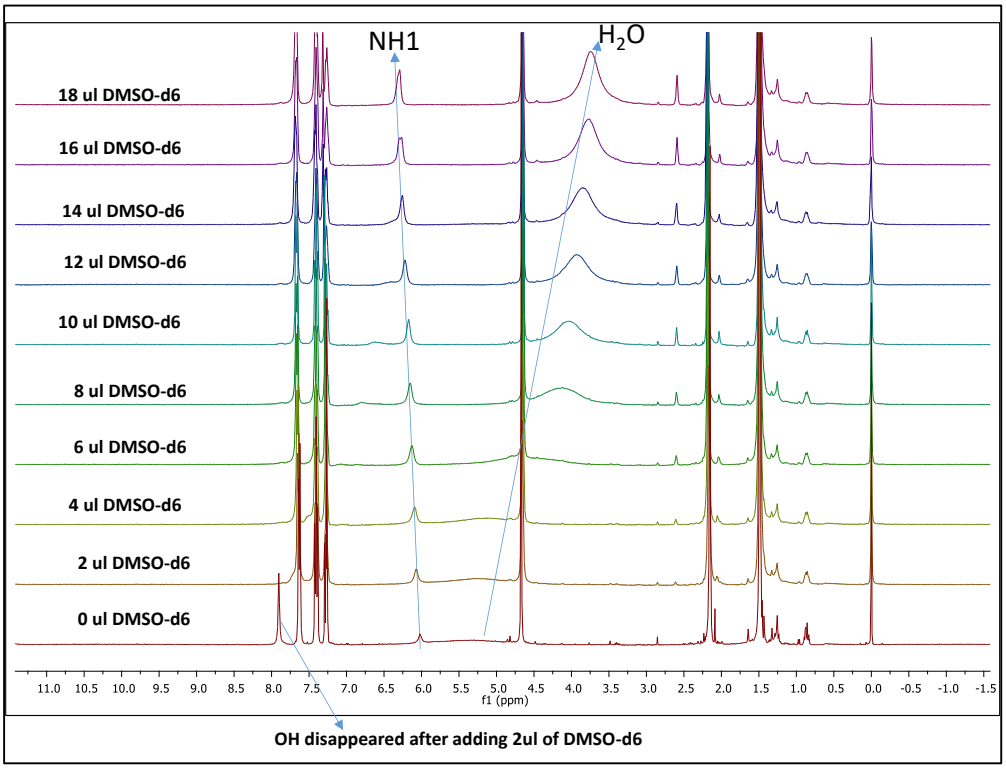


Figure S2. ESI-MS/HRMS spectra of compound **2**

2.DMSO-d_6 titration ^{1}H-NMR spectra for compound **2**



	NH(1)			**OH(1)**		
Compound	0 µl DMSO-d_6	20 µl DMSO-d_6	dδ (ppm)	0 µl DMSO-d_6	20 µl DMSO-d_6	dδ (ppm)
2	6.015	6.288	0.273	7.903	Disappeared after 2 µl DMSO-d_6	-

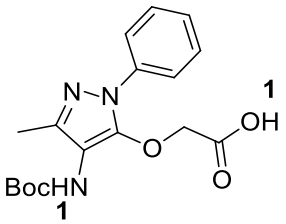


Figure S3. DMSO-d6 titration profile and difference in chemical shift, dδ (ppm) of compund **2** in $CDCl_3$

3.^{1}H-NOESY spectra of compound **2**

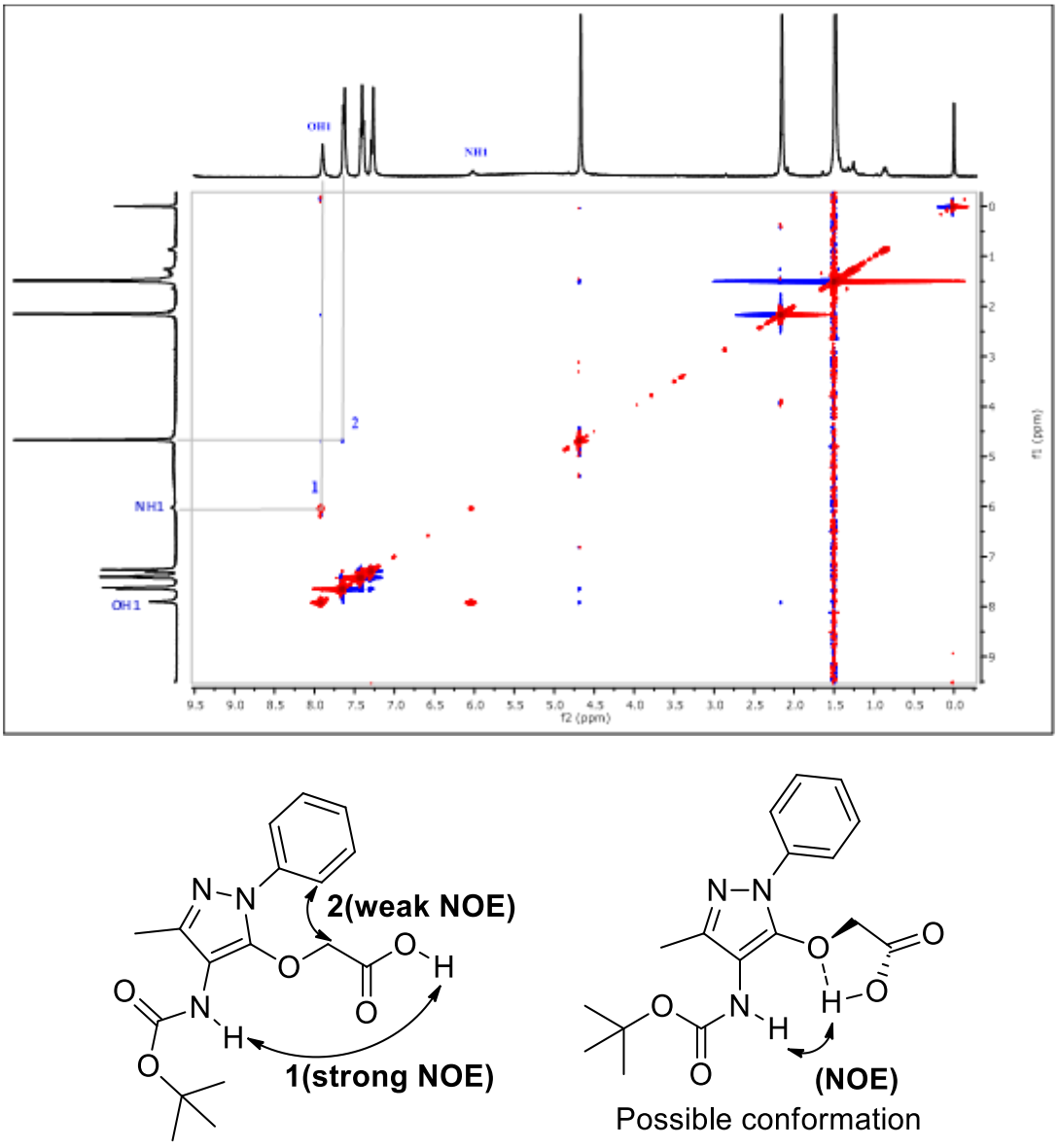


Figure S4. ^{1}H-NOESY spectra, NOE assignments and possible structure in $CDCl_3$ of compound **2**

4.^{1}H-/ ^{13}C-/ESI-MS/HRMS spectra of compound **3a**

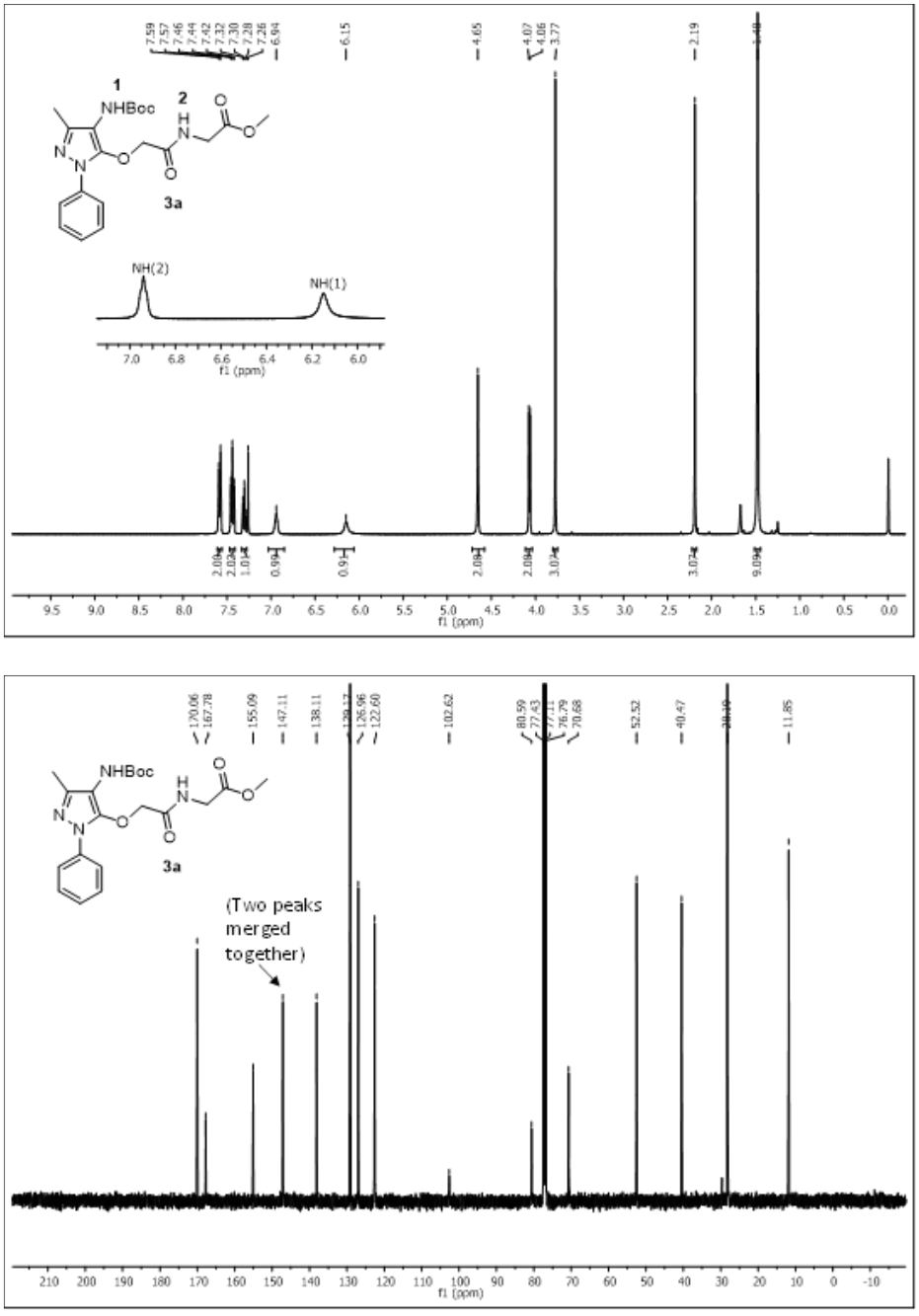


Figure S5. ^{1}H and ^{13}C NMR spectra of compound **3a** in $CDCl_3$

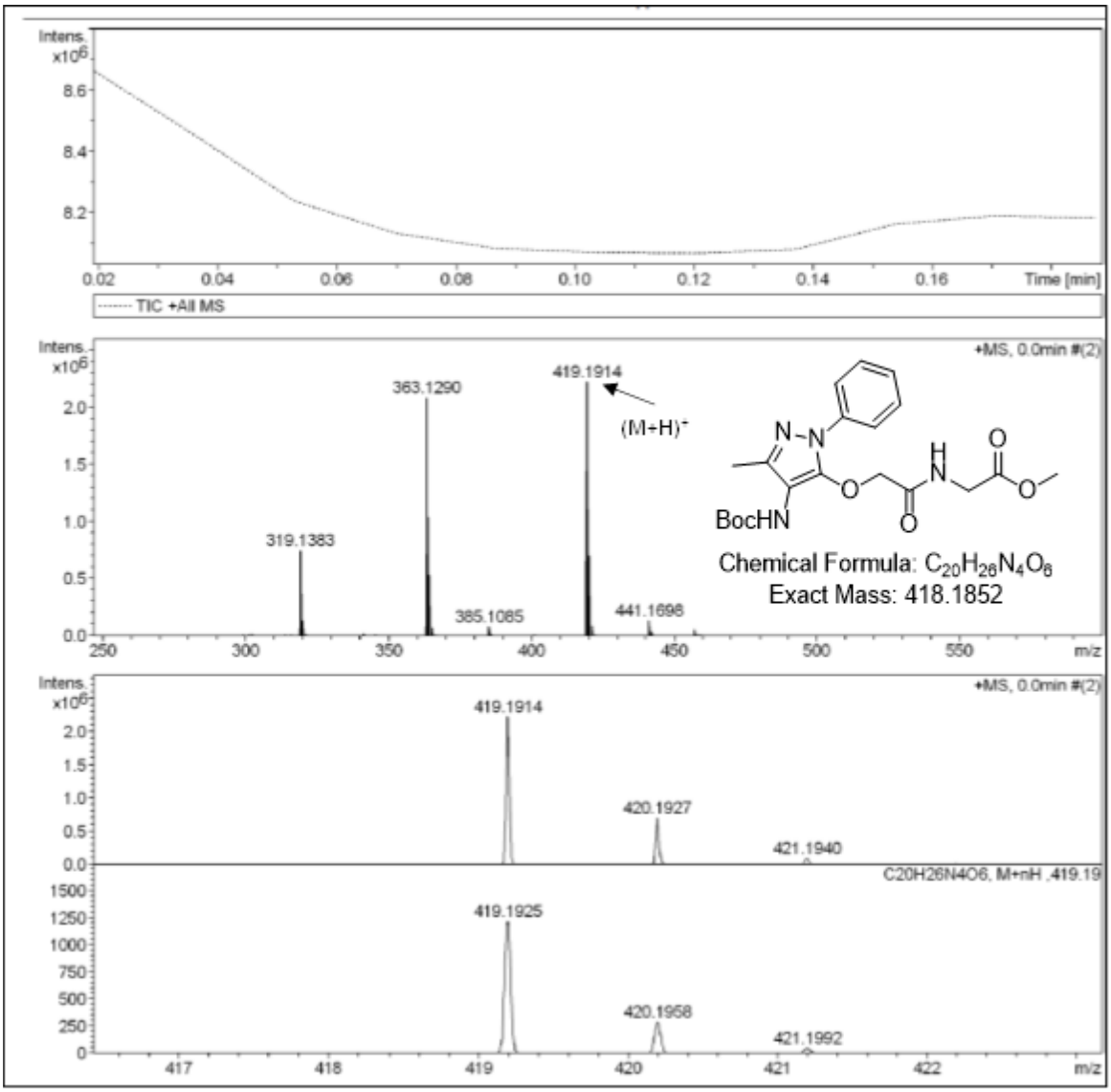


Figure S6. ESI-MS/HRMS spectra of compound **3a**

5.^{1}H-/ ^{13}C-/ESI-MS/HRMSspectra of compound **3e**

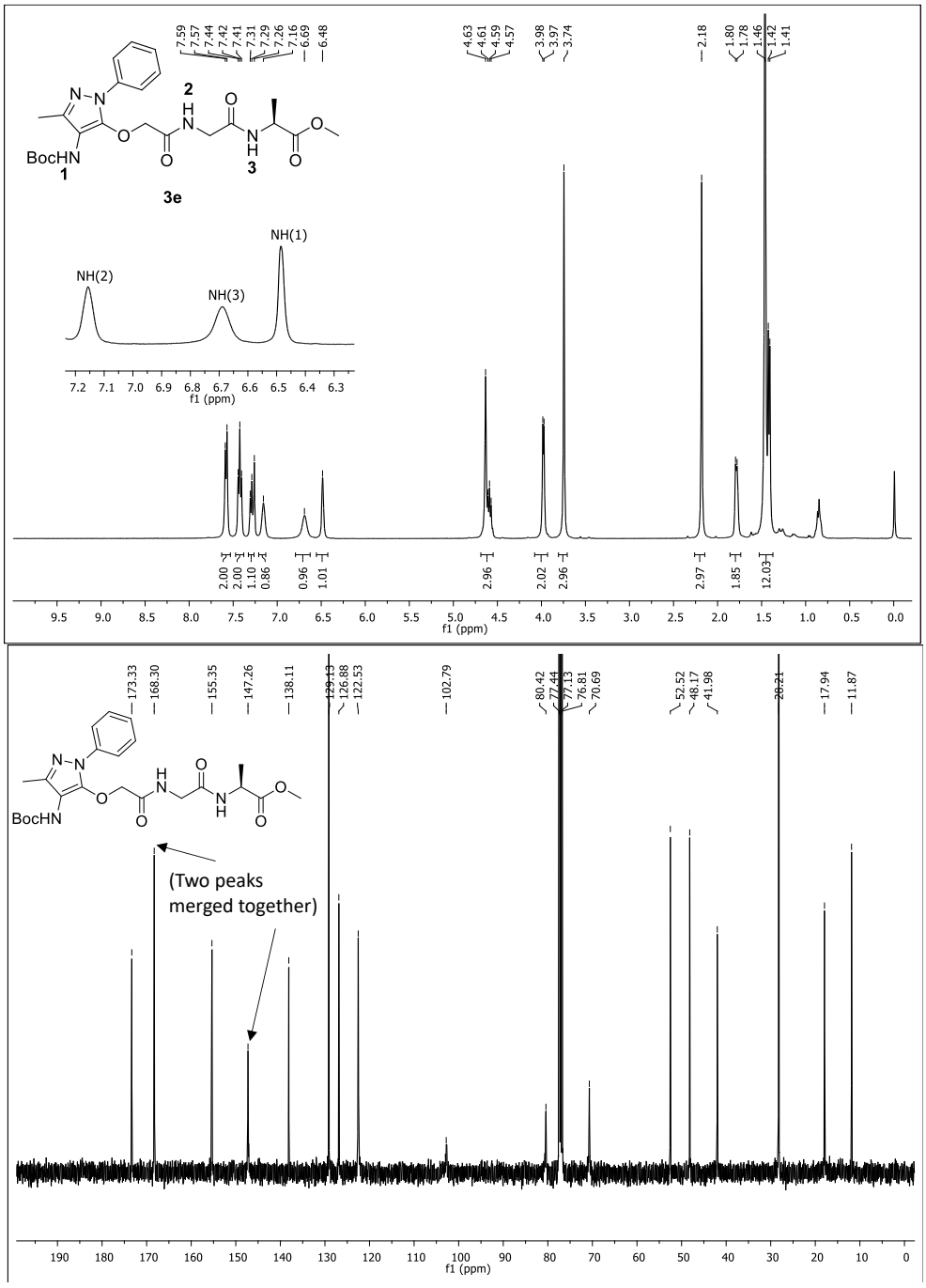


Figure S7.^{1}H and ^{13}C NMR spectra of compound **3e** in $CDCl_3$

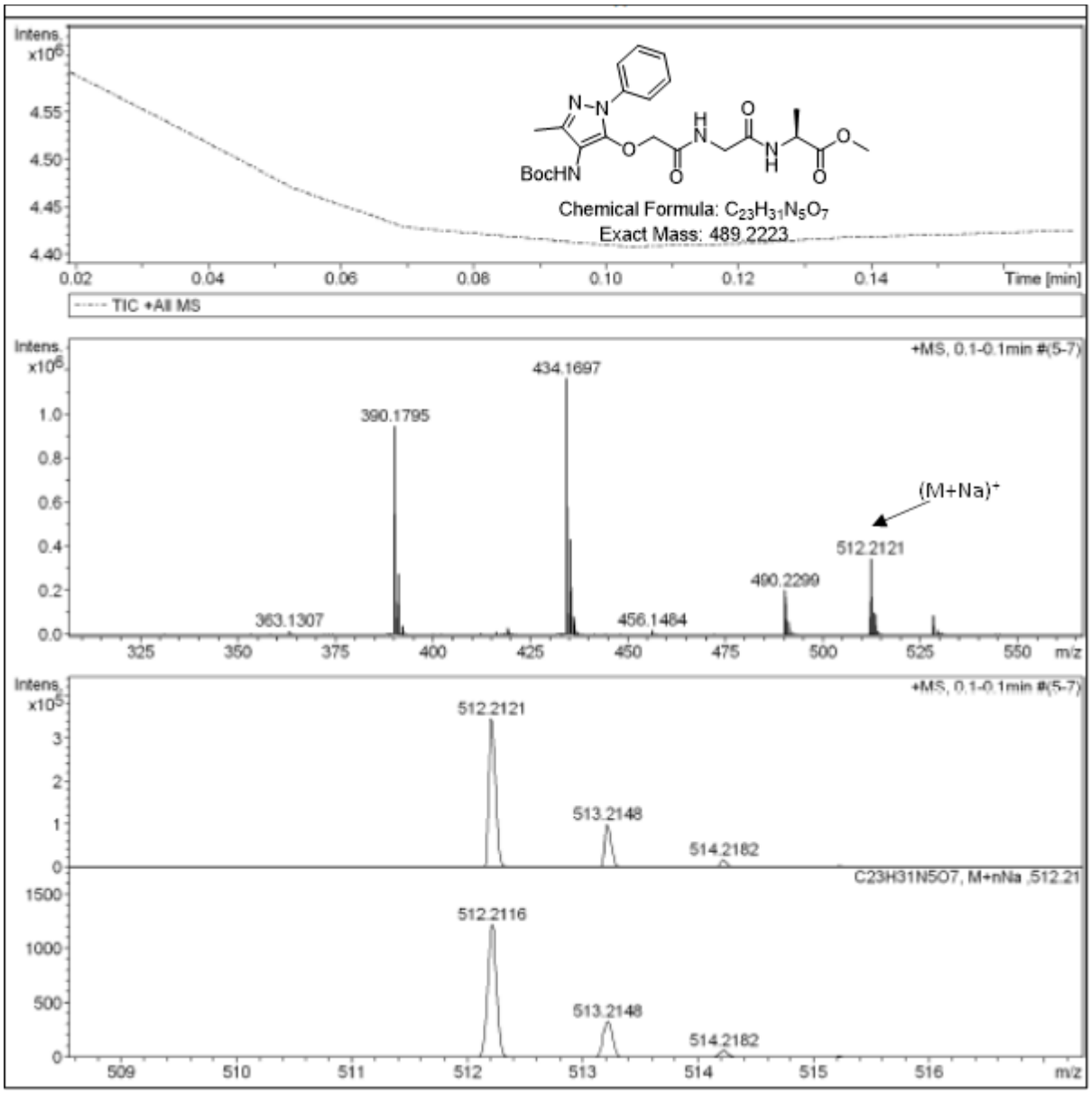


Figure S8. ESI-MS/HRMS spectra of compound **3e**

6.^{1}H-/ ^{13}C-/ESI-MS/HRMS spectra of compound **3f**

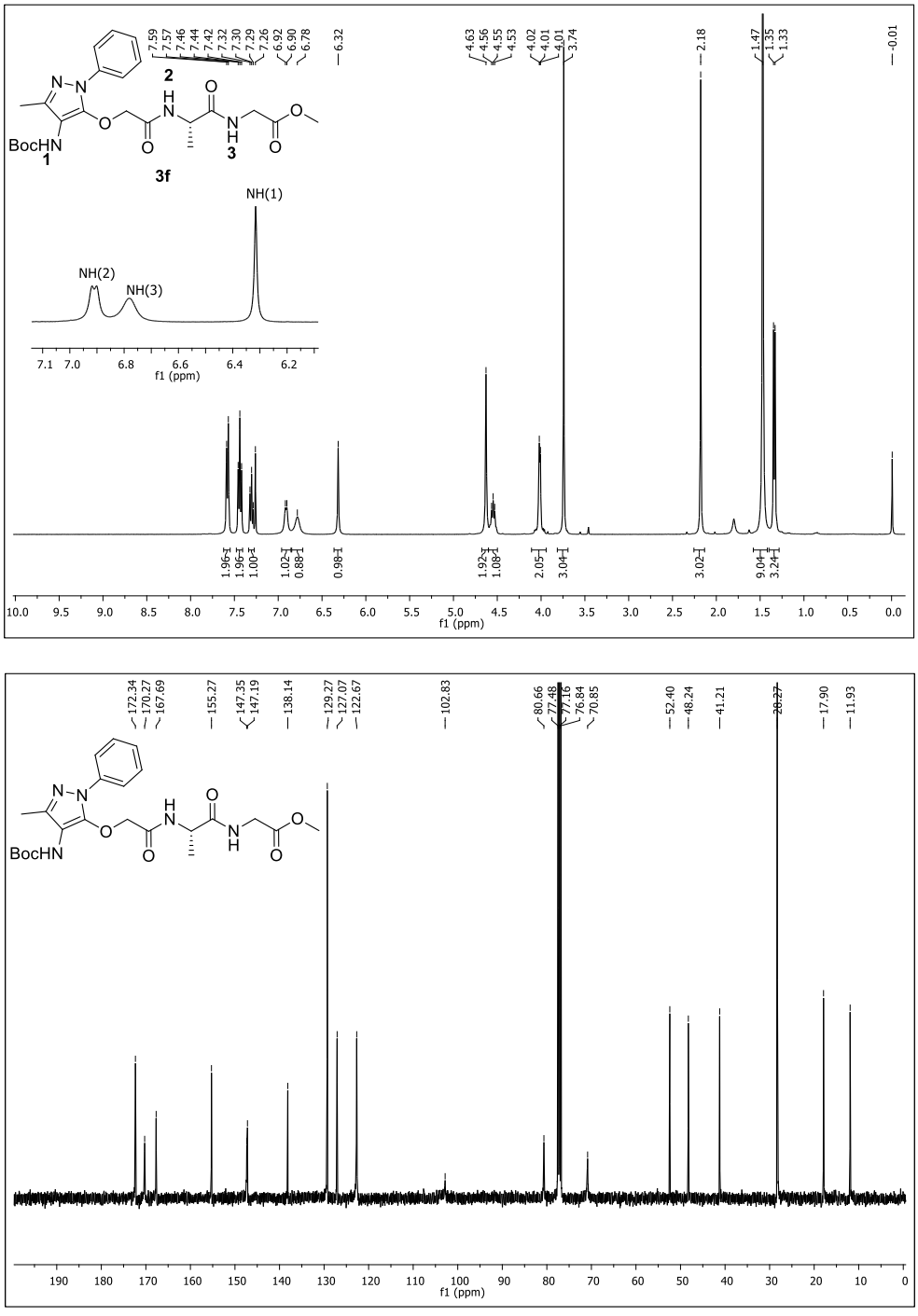


Figure S9.^{1}H and ^{13}C NMR spectra of compound **3f** in $CDCl_3$

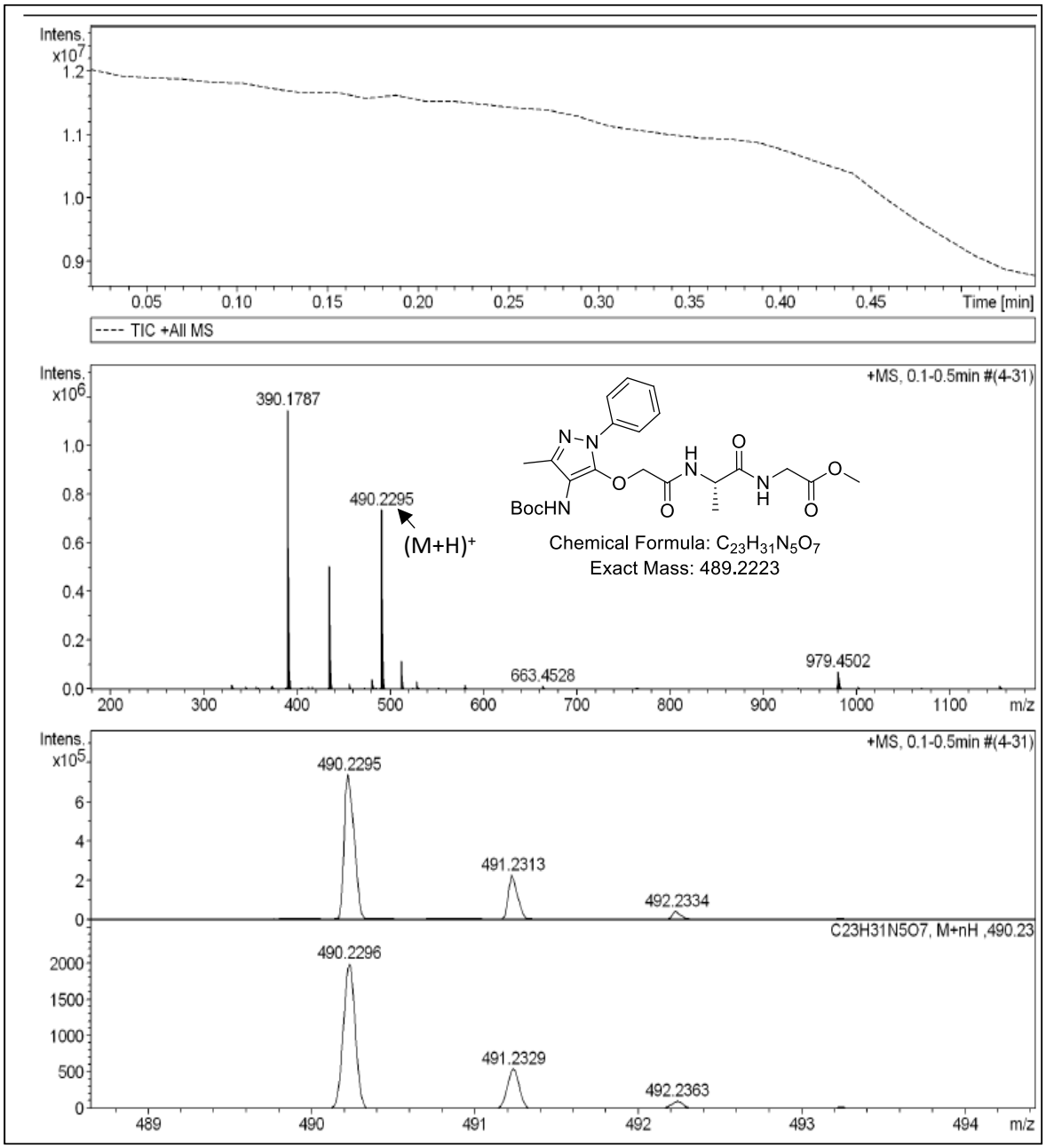


Figure S10. ESI-MS/HRMS spectra of compound **3f**

7.^{1}H-/ ^{13}C-/ESI-MS/HRMS spectra of lactam **4**

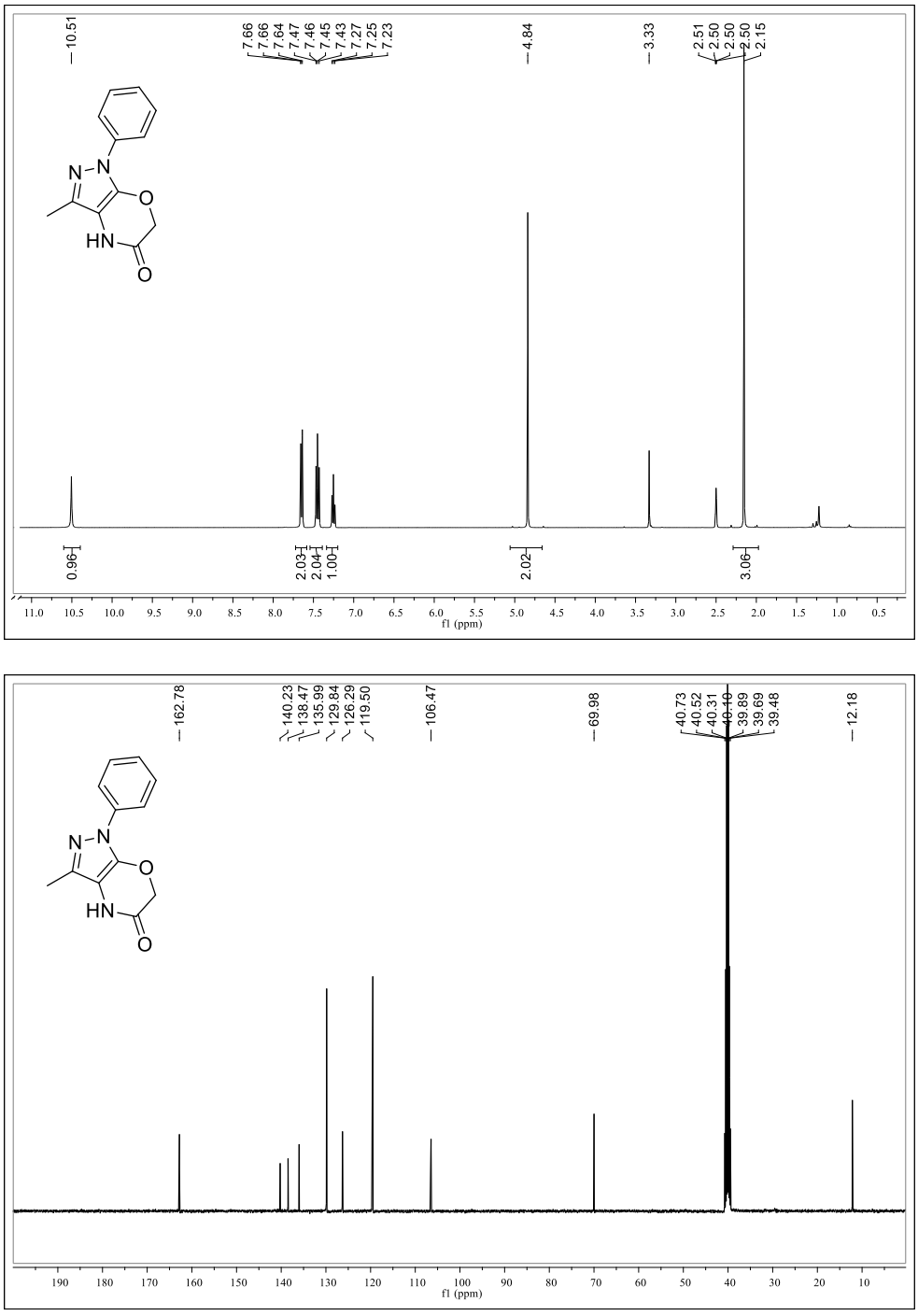


Figure S11. ^{1}H and ^{13}C NMR spectra of lactam **4** in $CDCl_3$

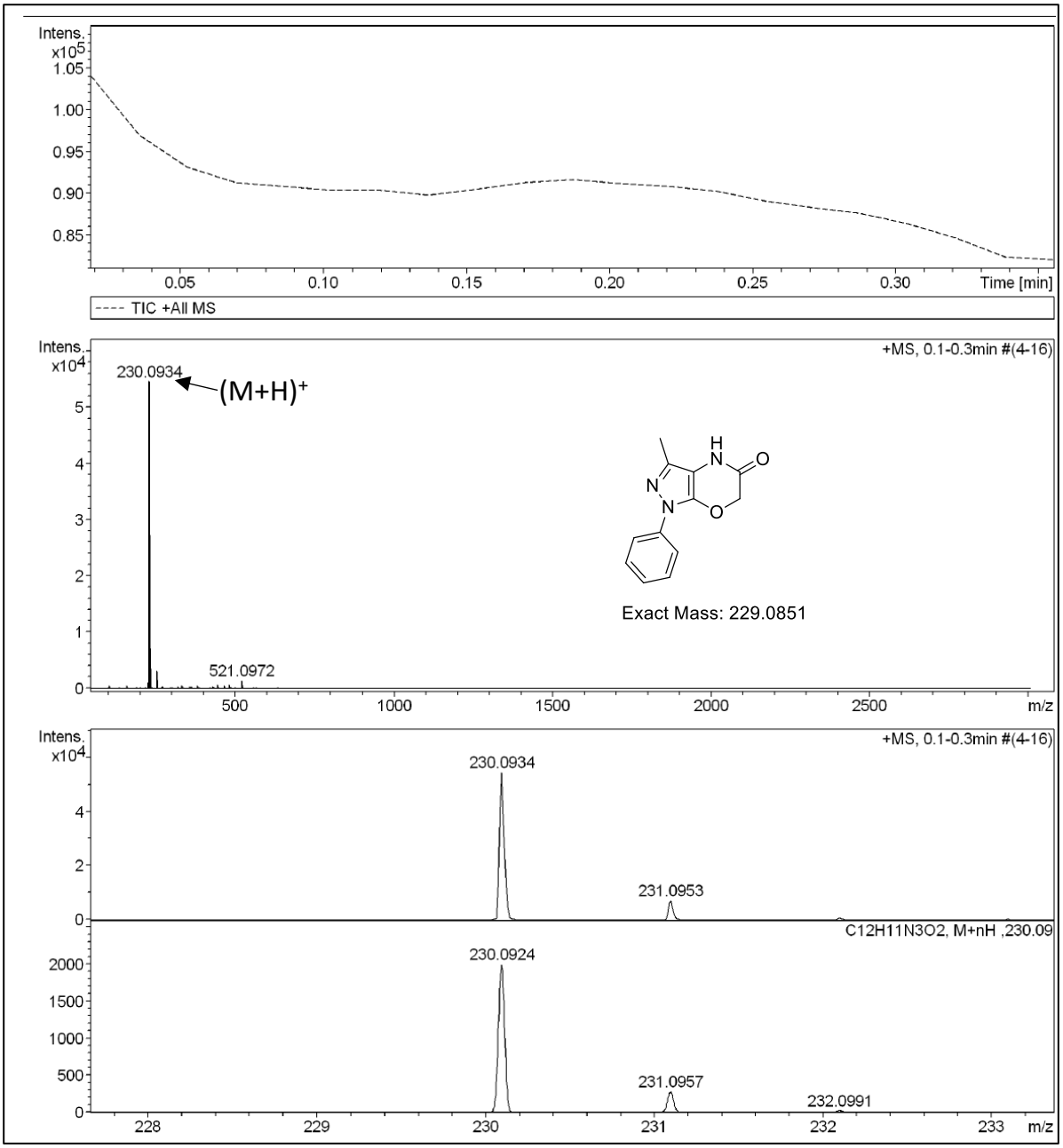


Figure S12. ESI-MS/HRMS spectra of lactam **4**

8.Time dependent ESI-MS spectra of compound **3a**

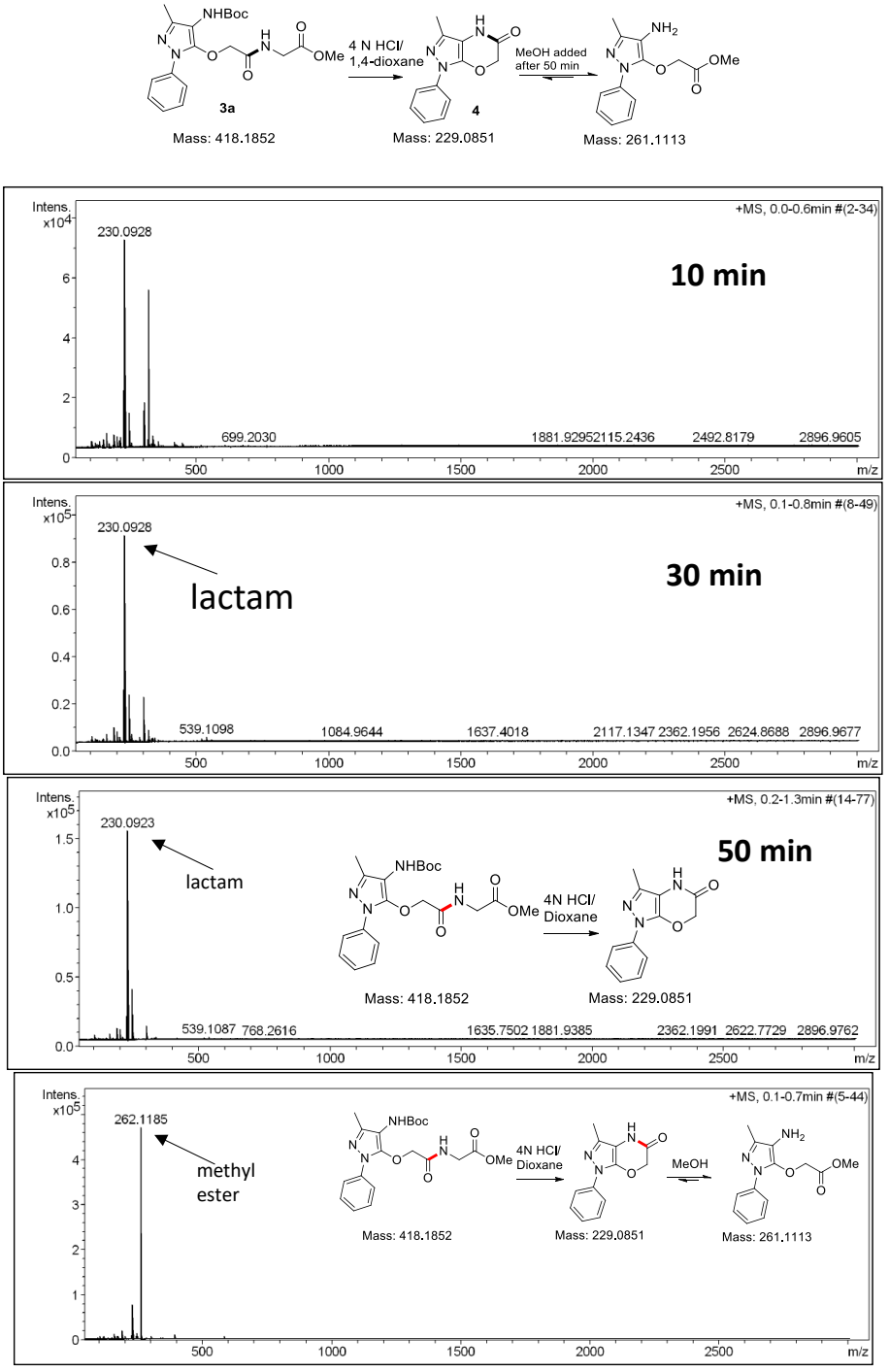
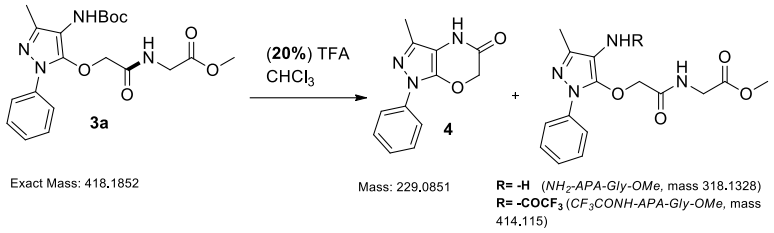


Figure S13. ESI-MS spectra of compound **3a** at different time intervals, MeOH was added after 50 min.

9. Time dependent ESI-MS spectra of compound **3a** in 20% TFA/$CHCl_3$



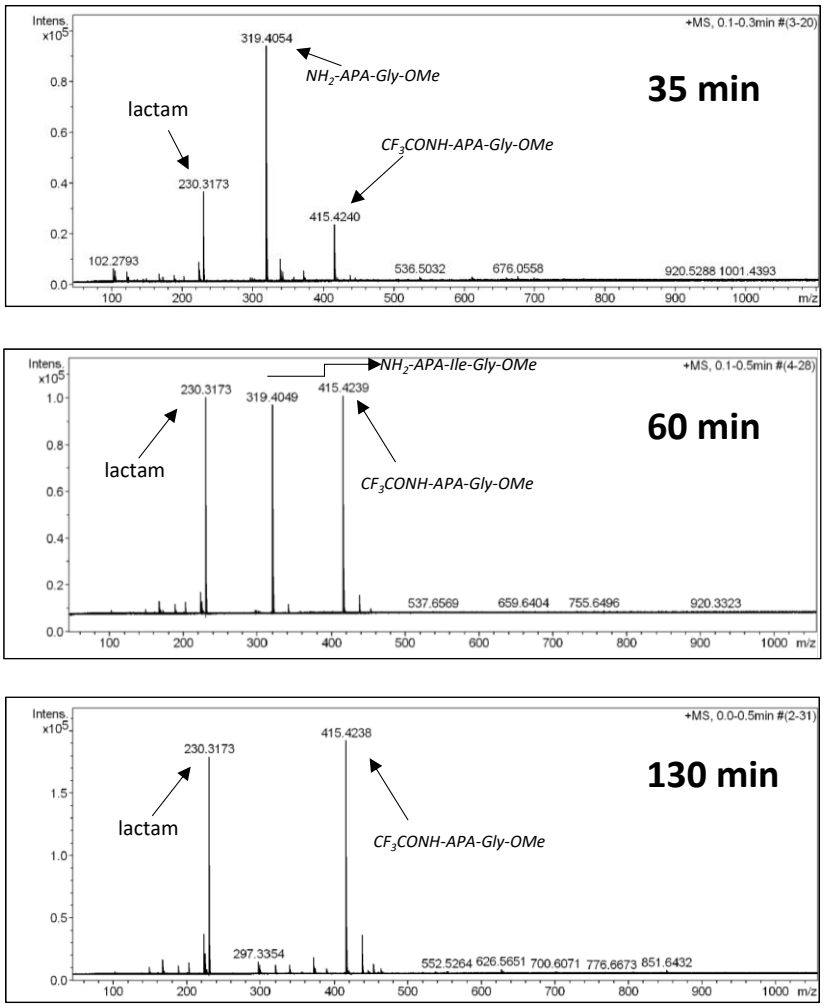
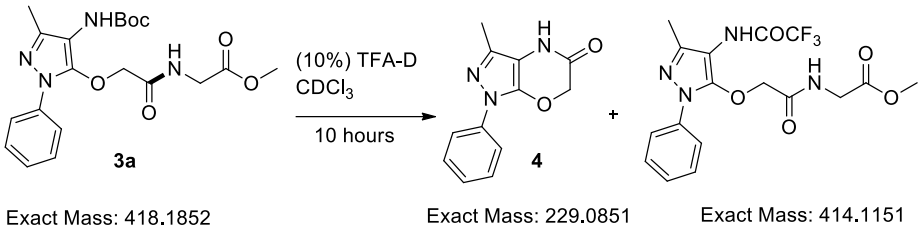


Figure S14. ESI-MS spectra of compound **3a** at different time intervals in 20% TFA in $CHCl_3$

10.ESI-MS spectra of compound **3a** in 10% TFA/$CHCl_3$



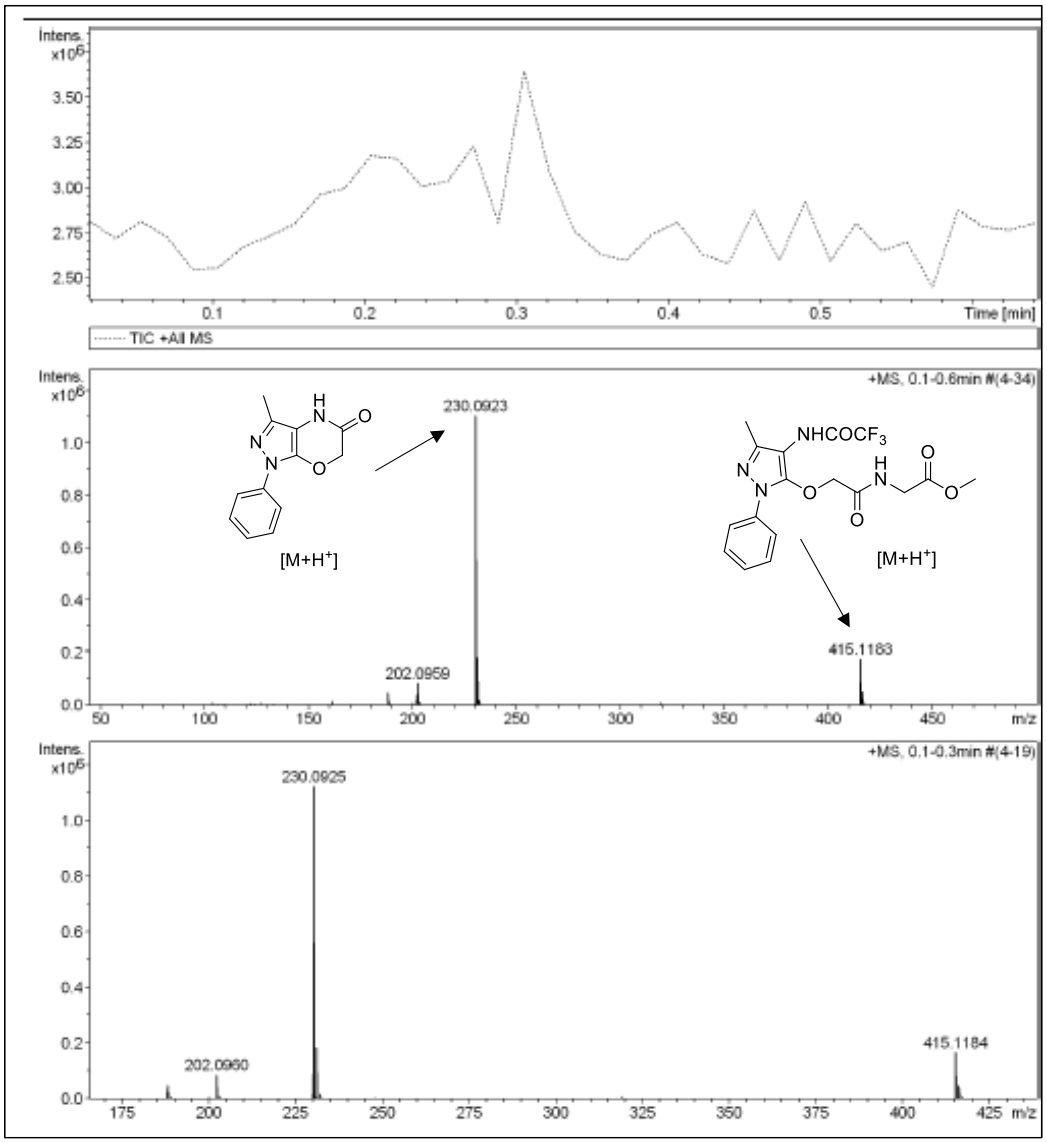
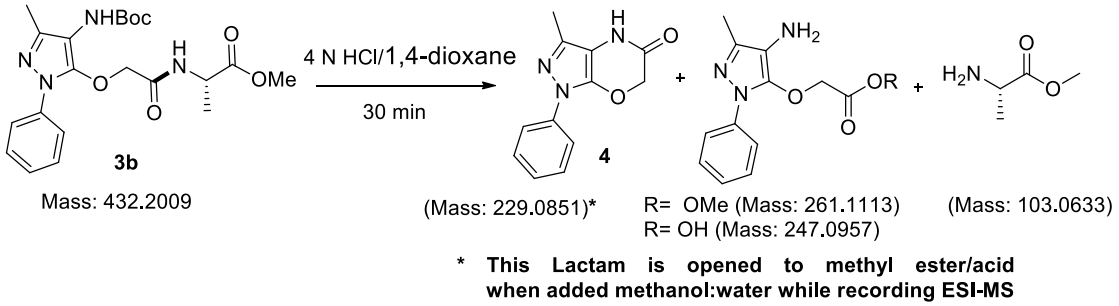


Figure S15. ESI-MS spectra of peptide compound **3a** in 10% TFA in $CHCl_3$

11.ESI-MS spectra of amide hydrolysis in compound **3b**



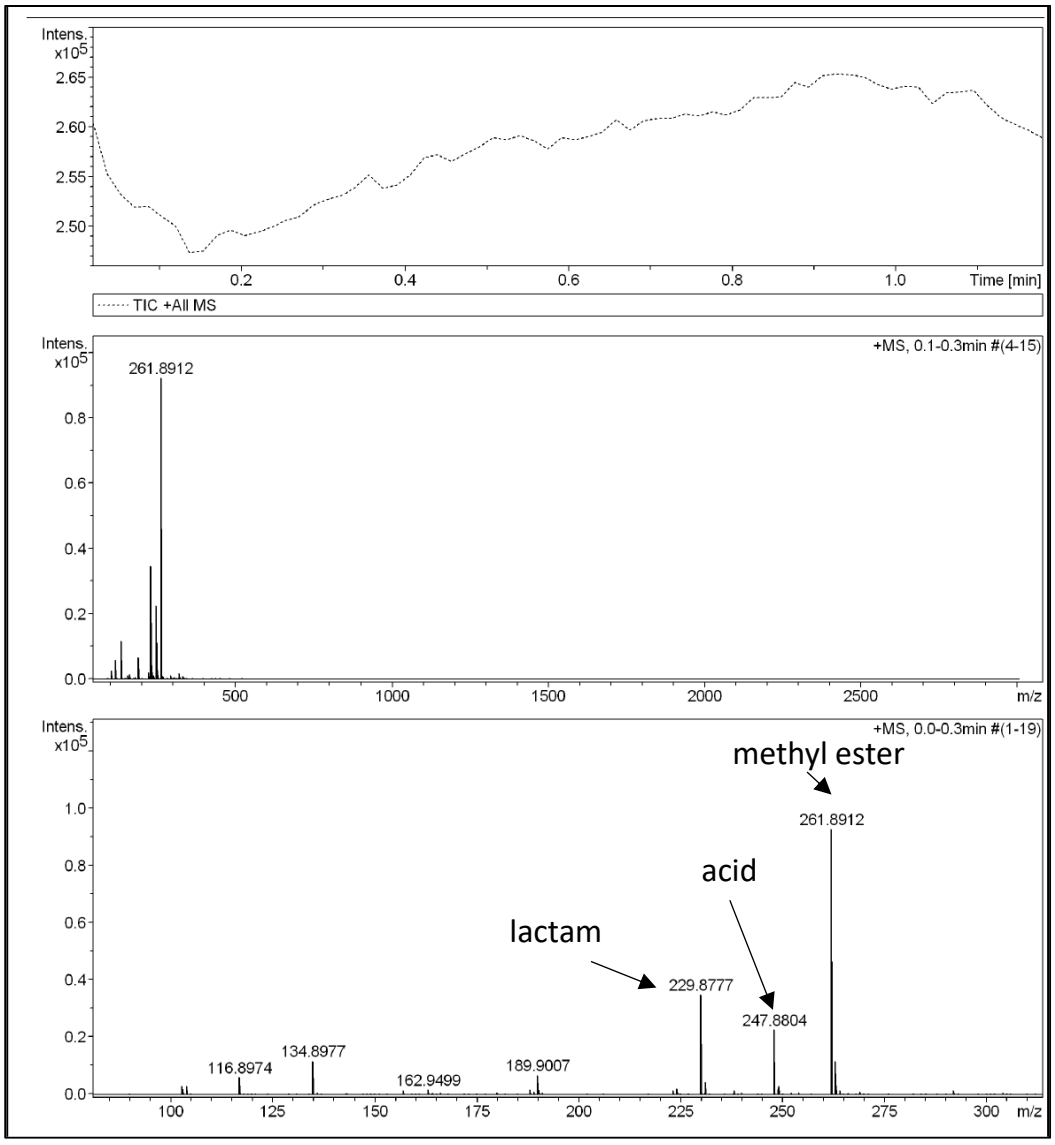
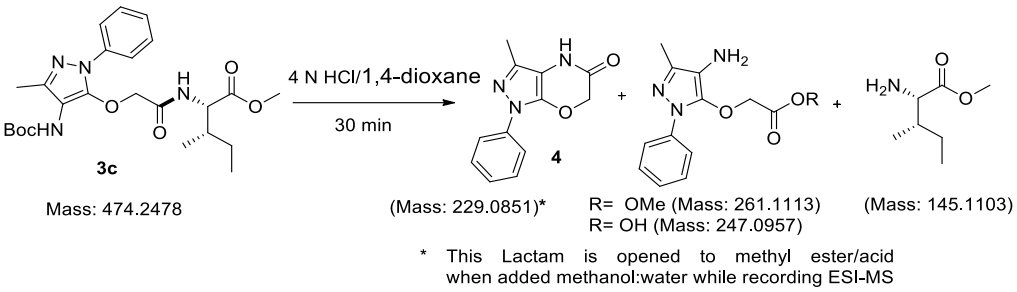


Figure S16. ESI-MS spectra of compound **3b** in 4 N HCl/1,4-dioxane, after 30 min

12.ESI-MS spectra of amide hydrolysis in compound **3c**



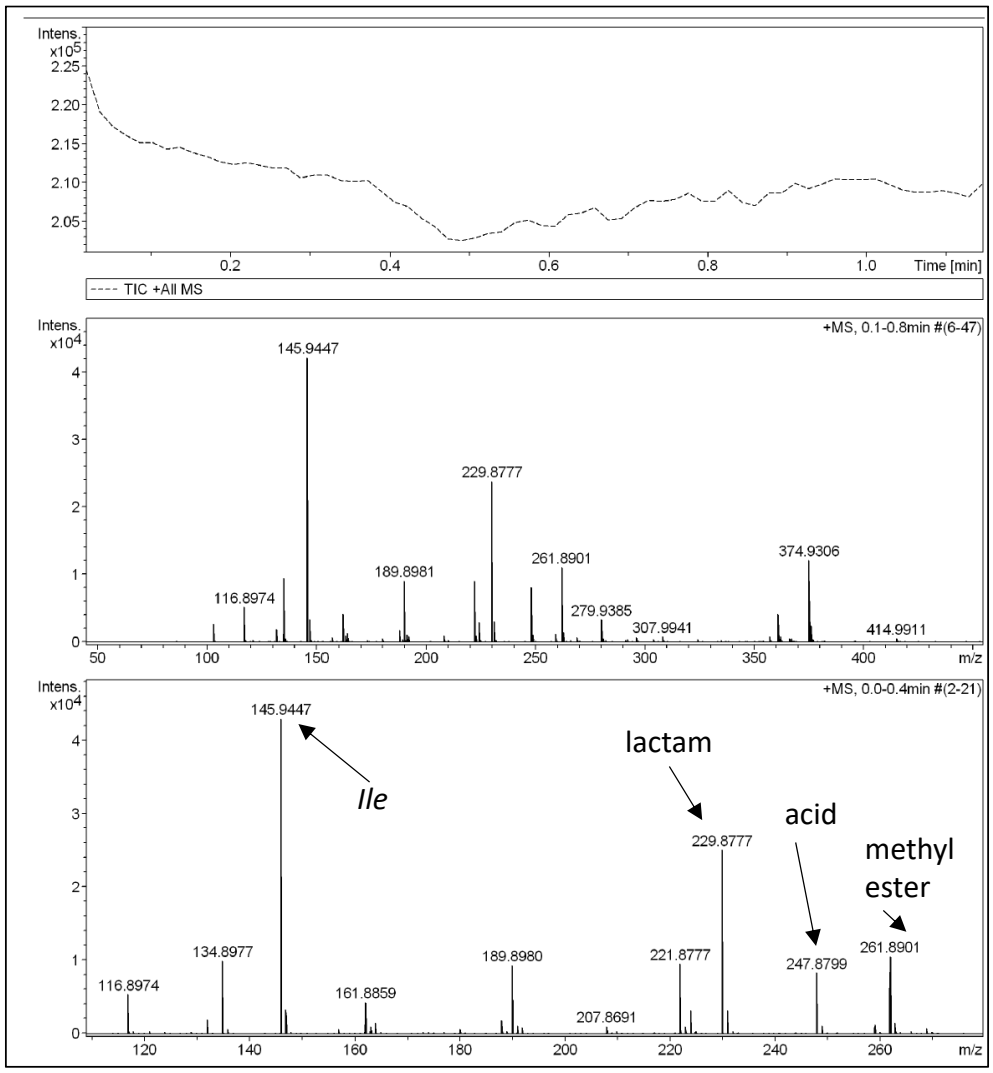
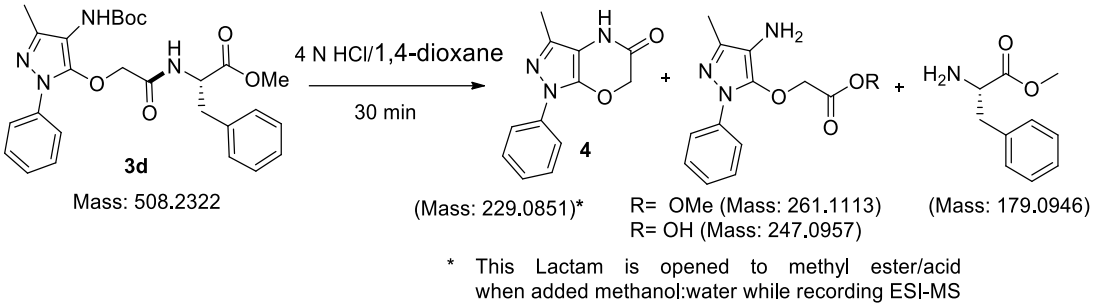


Figure S17. ESI-MS spectra of compound **3c** in 4 N HCl/1,4-dioxane, after 30 min

13.ESI-MS spectra of amide hydrolysis in compound **3d**



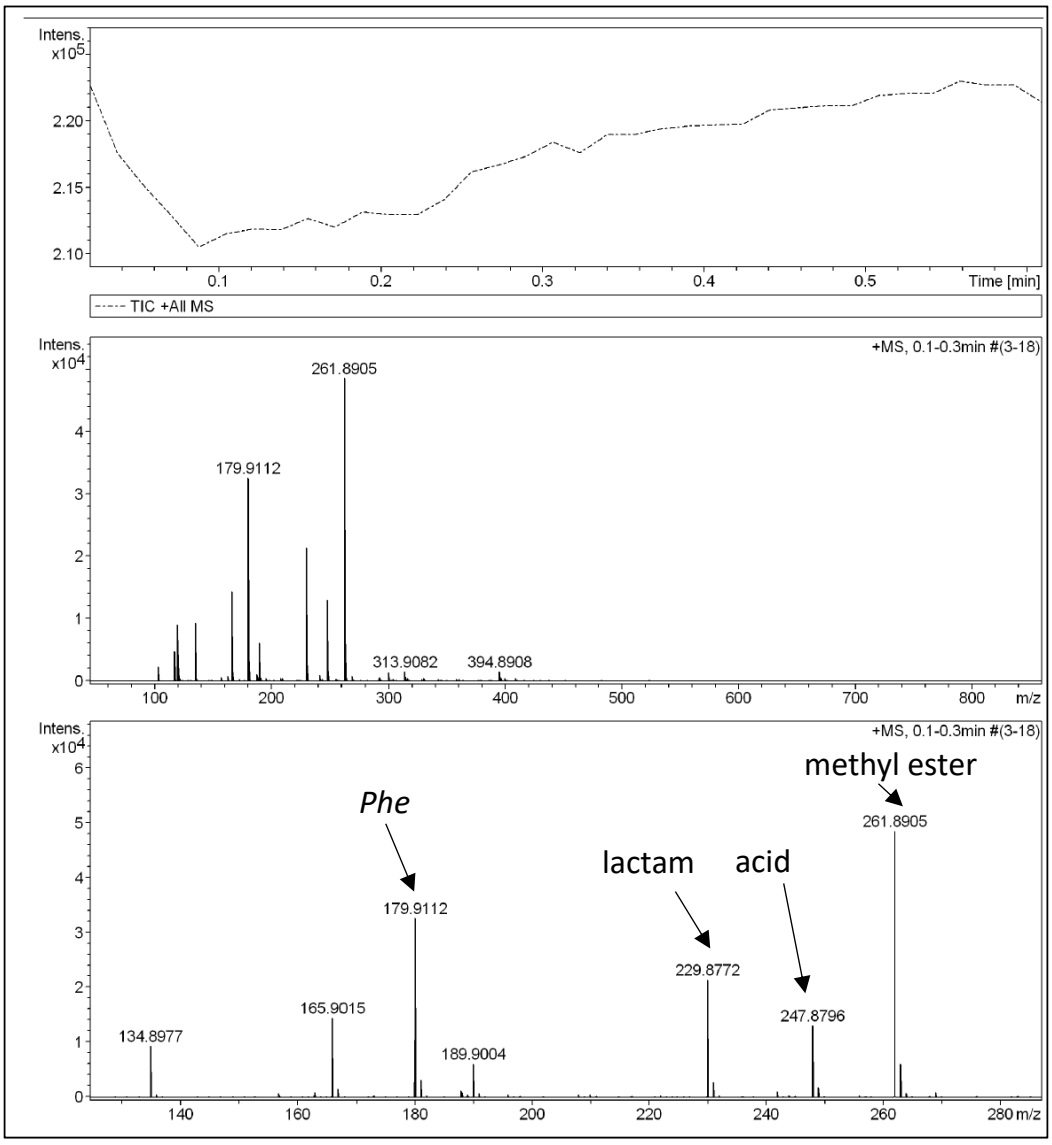
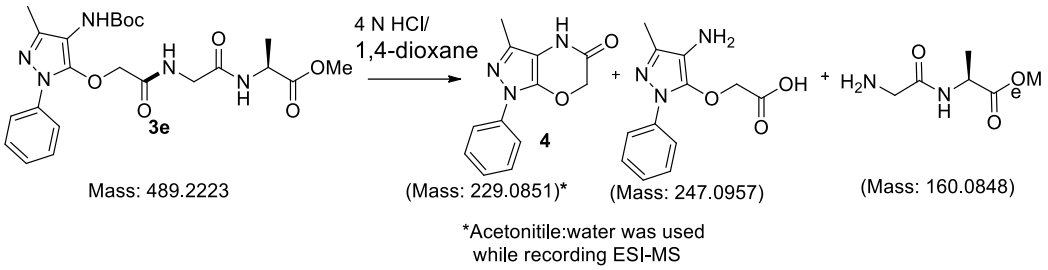


Figure S18. ESI-MS spectra of compound **3d** in 4 N HCl/1,4-dioxane, after 30 min

14.ESI-MS spectra of amide hydrolysis in compound **3e**



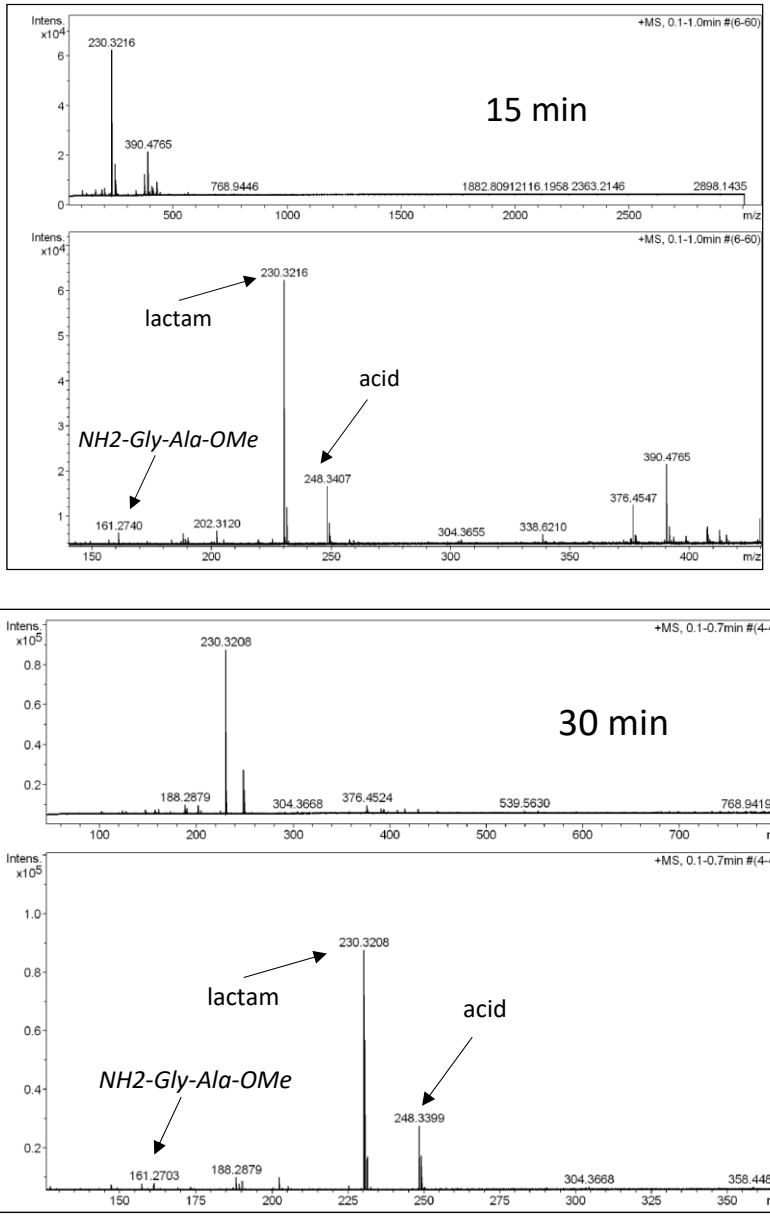
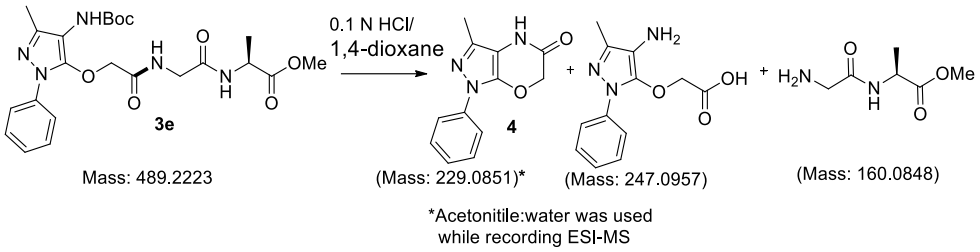


Figure S19. ESI-MS spectra of compound **3e** at 15 min & 30 min in 4 N HCl/1,4-dioxane

15.ESI-MS spectra of amide hydrolysis in compound **3e**



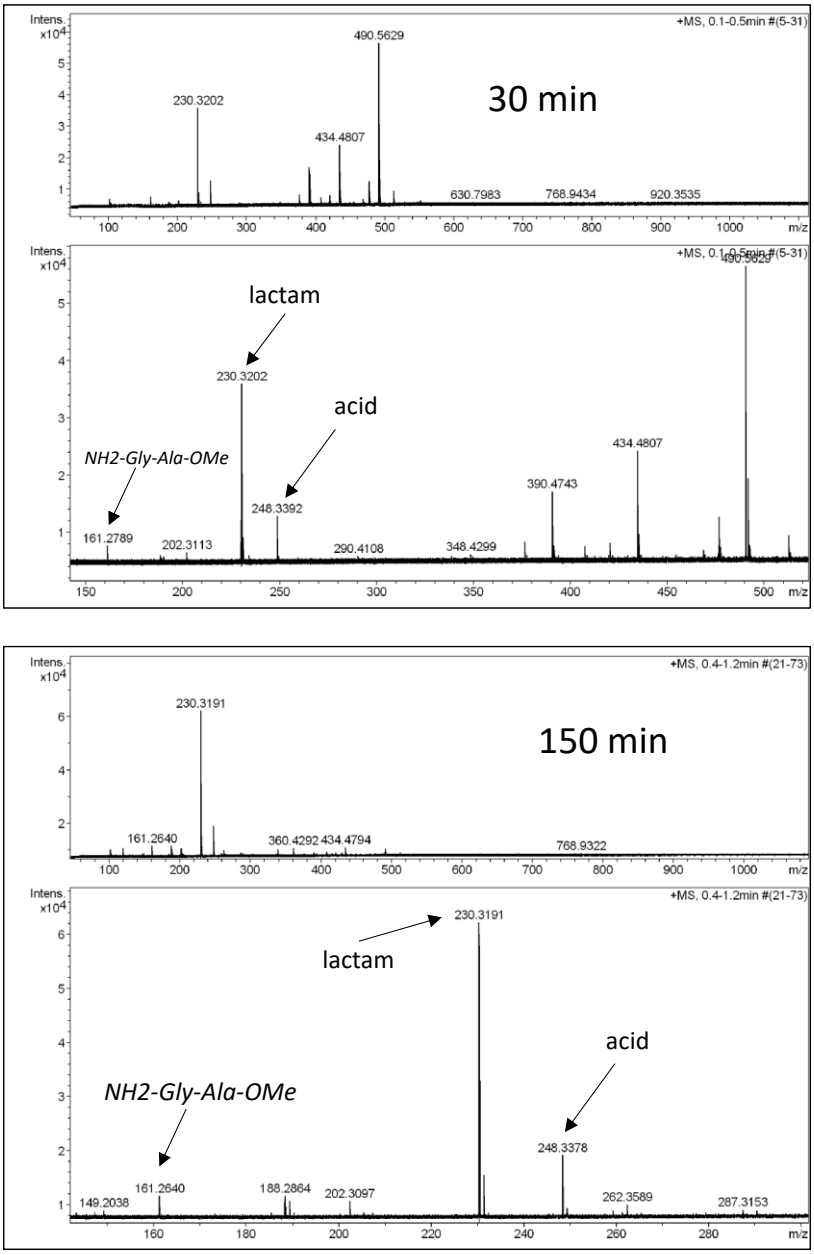
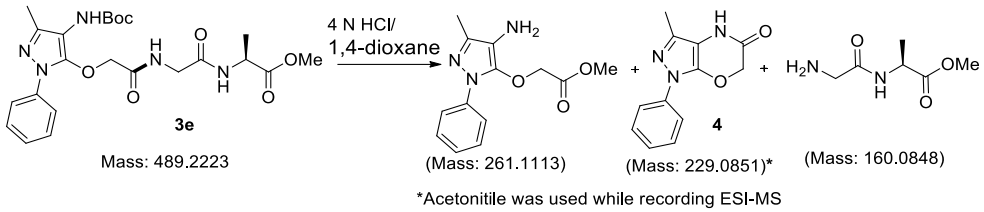
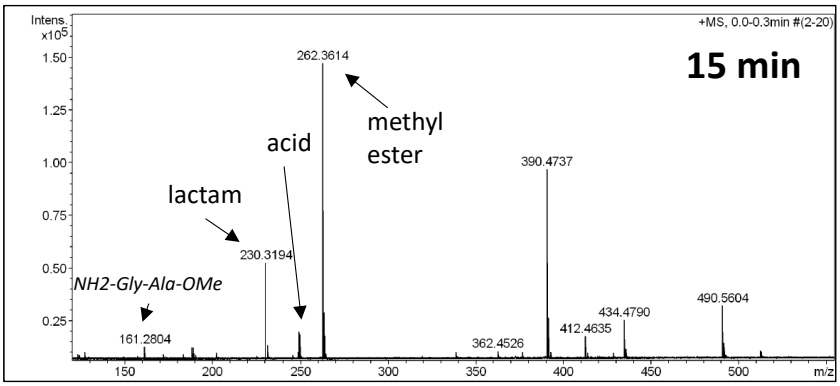
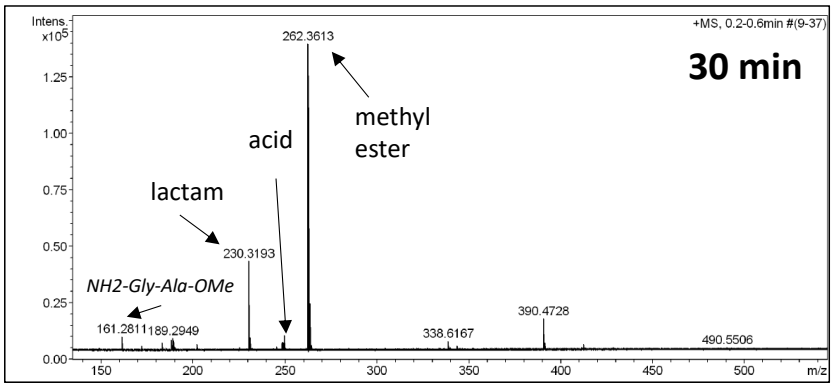


Figure S20. ESI-MS spectra of compound **3e** at 30 & 150 min. in 0.1 N HCl/1,4-dioxane

16.ESI-MS spectra of amide hydrolysis in compound **3e**







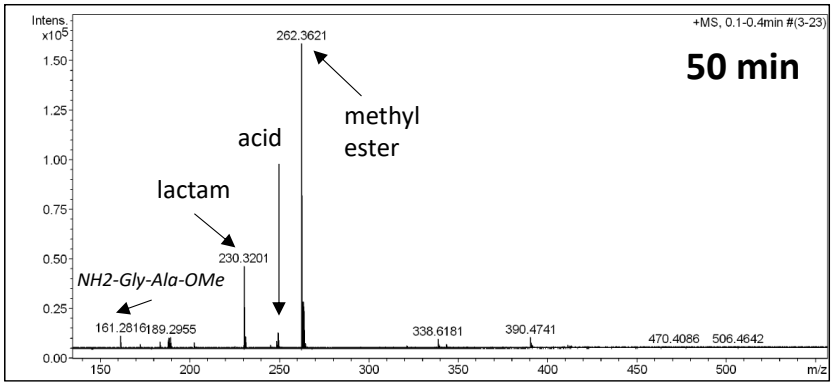
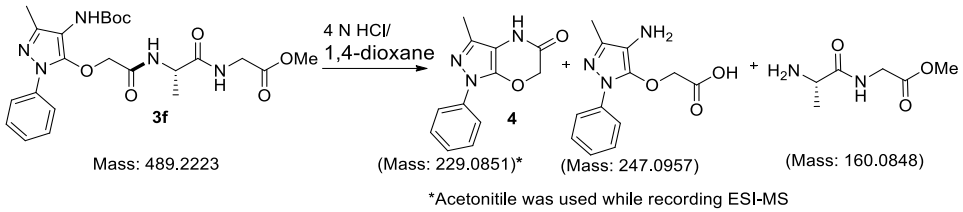


Figure S21. ESI-MS spectra of compound **3e** 15, 30 & 50 min in 4 N HCl/1,4-dioxane

17.ESI-MS spectra of amide hydrolysis in compound **3f**



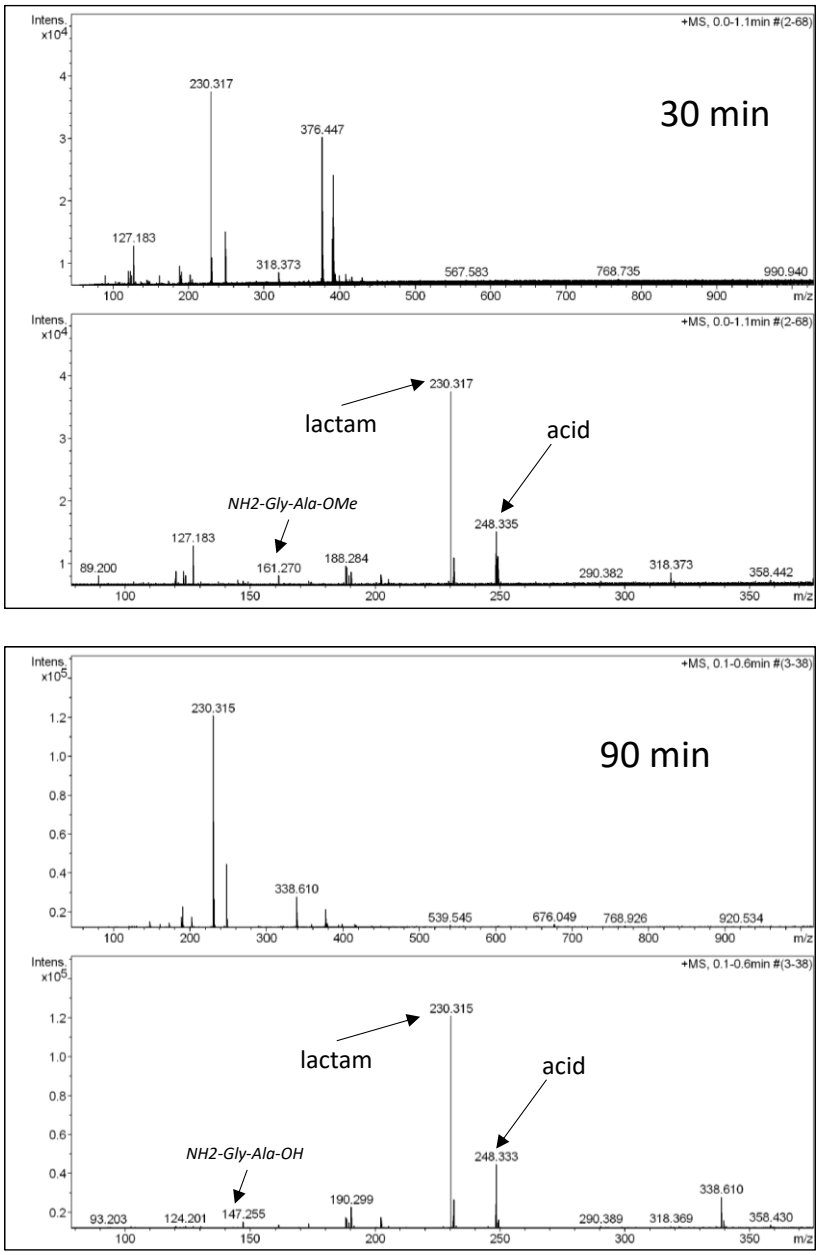
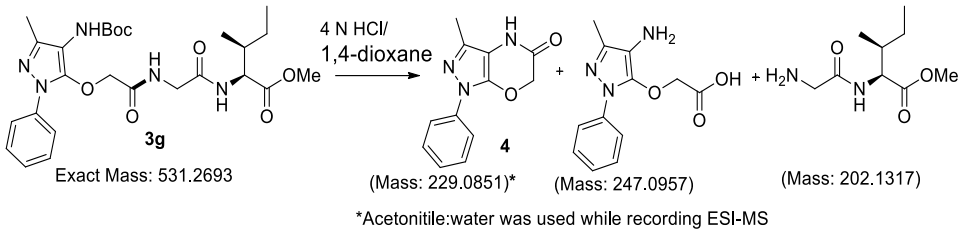


Figure S22. ESI-MS spectra of compound **3f** at 30 min & 90 min in 4 N HCl/1,4-dioxane

18.ESI-MS spectra of amide hydrolysis in compound **3g**



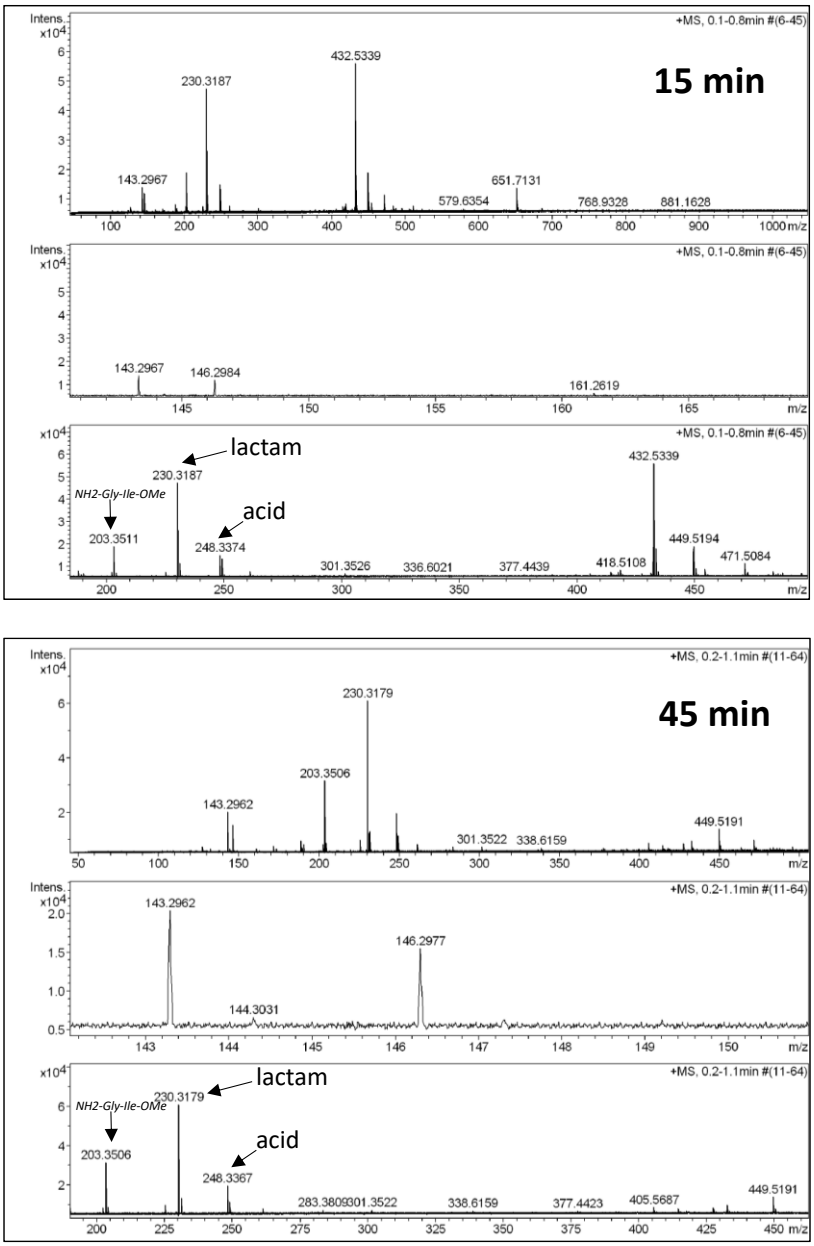
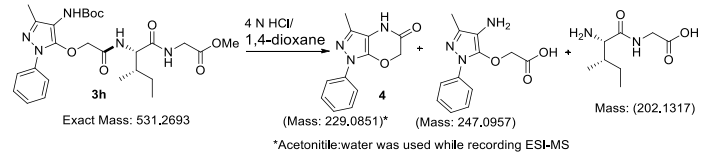
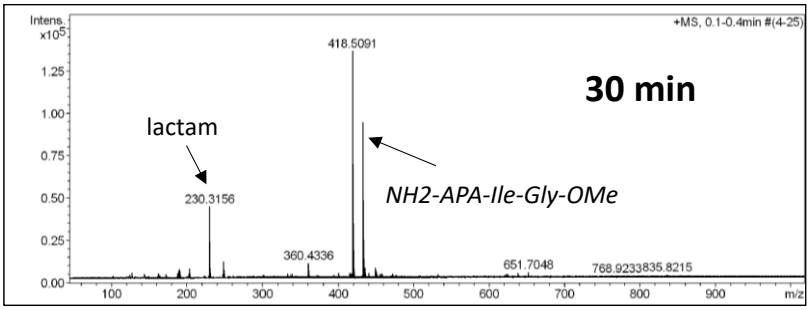
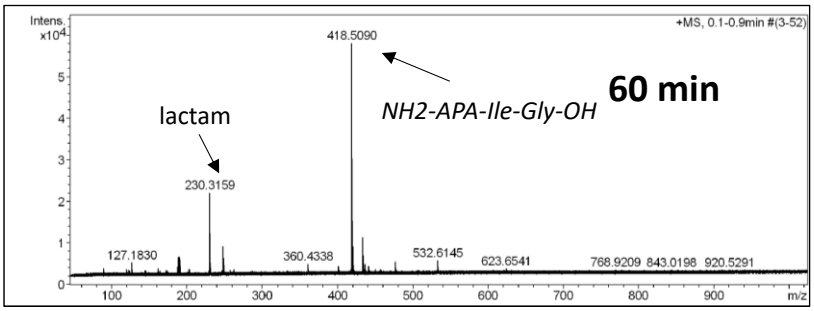


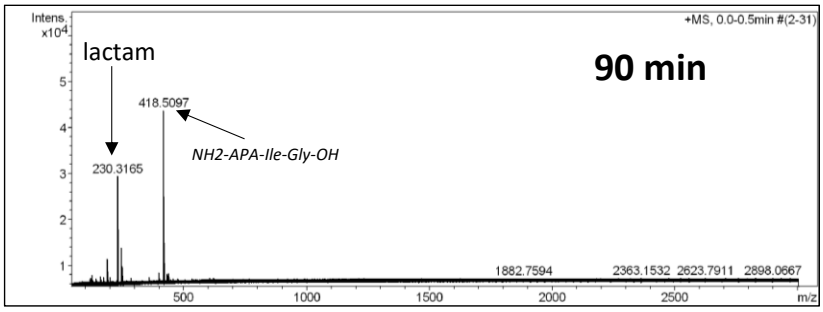
Figure S23. ESI-MS spectra of compound **3g** at 15 min & 45 min in 4 N HCl/1,4-dioxane

19.ESI-MS spectra of amide hydrolysis in compound **3h**









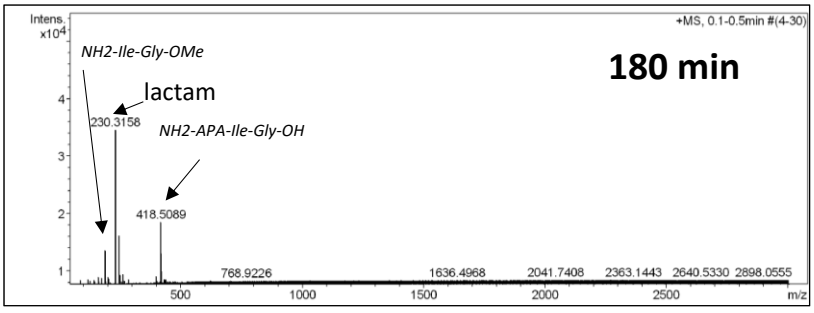
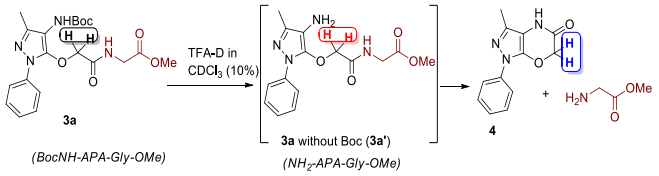
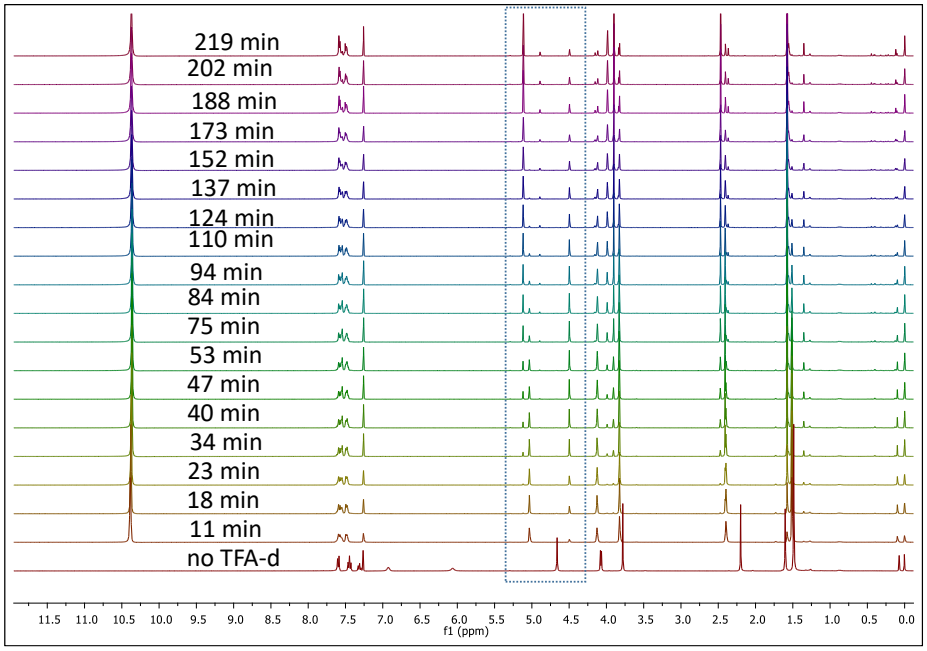


Figure S24. ESI-MS spectra of compound **3h** at 30, 60, 90 & 180 min in 4 N HCl/1,4-dioxane

20.Time dependent ^{1}H-NMR spectra of amide hydrolysis in compound **3a**





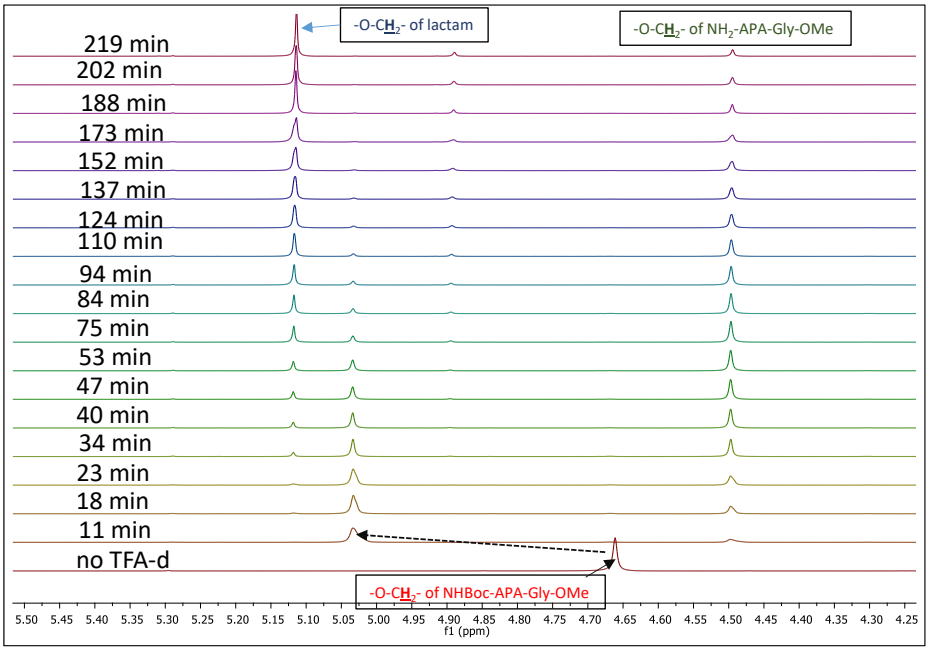


Figure S25. Time dependent ^{1}H-NMR spectra of compound **3a** in 10% TFA/$CDCl_3$

21.ESI-MS spectra of compound **3a** after time dependent ^{1}H-NMR spectra

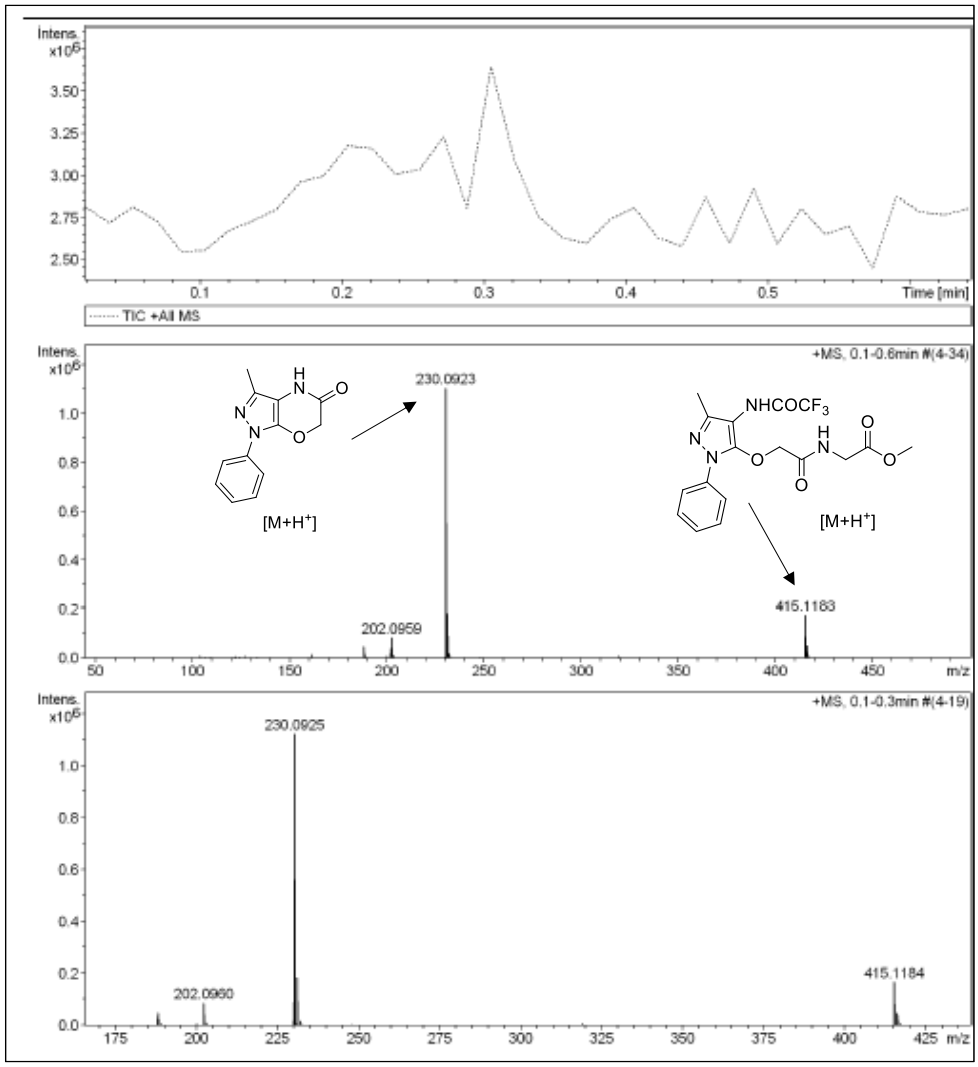
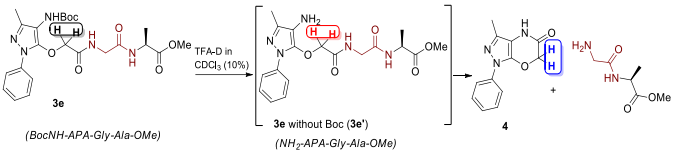
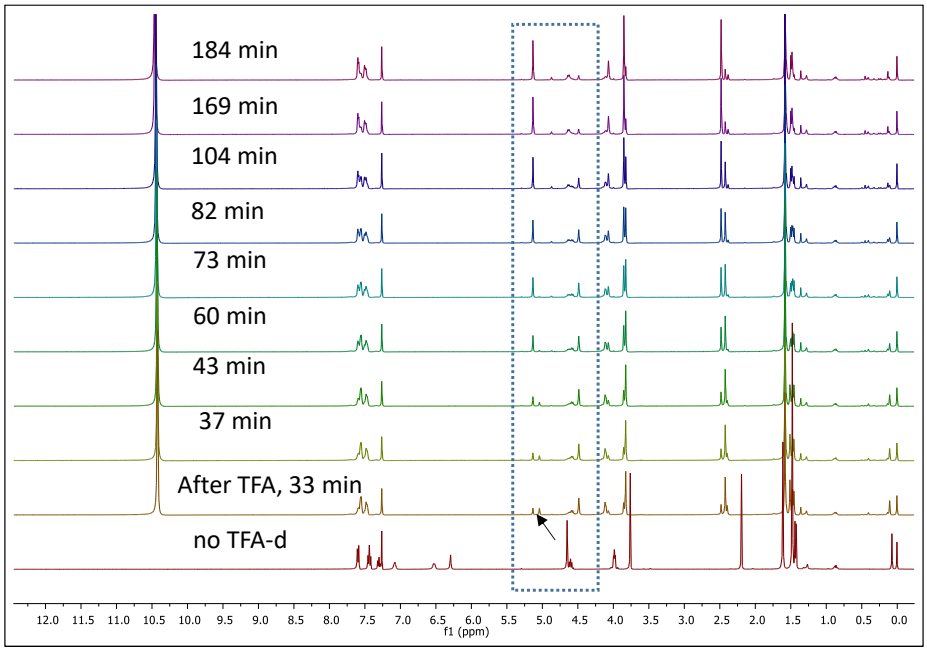


Figure S26. ESI-MS spectra of compound **3a** after time dependent ^{1}H-NMR spectra in 10% TFA/$CDCl_3$

22.Time dependent ^{1}H-NMR spectra of amide hydrolysis in compound **3e**





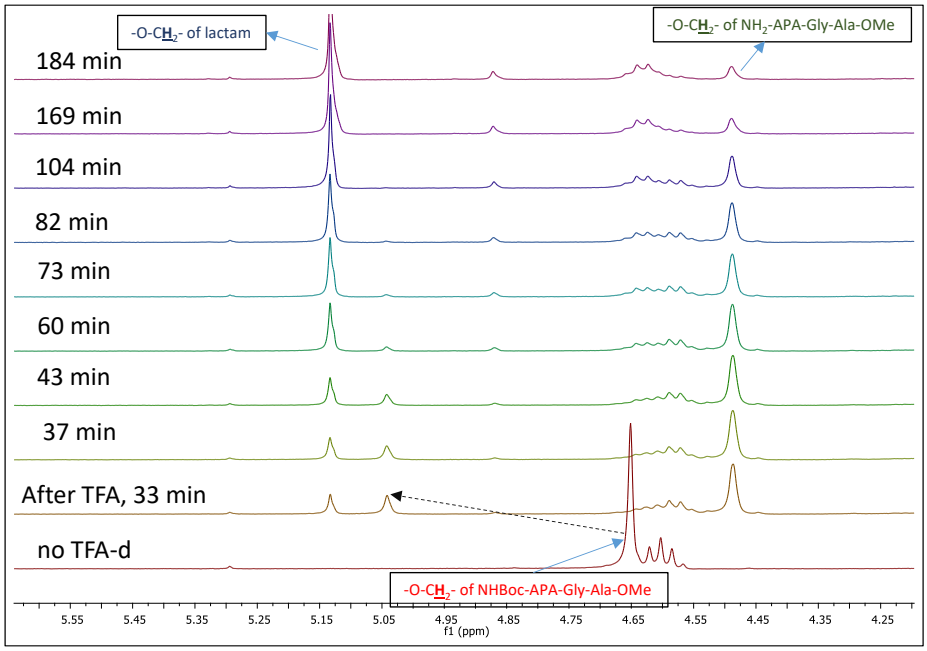


Figure S27. Time dependent ^{1}H-NMR spectra of compound **3e** in 10% TFA/$CDCl_3$

23.ESI-MS spectra of compound **3e** after time dependent ^{1}H-NMR spectra

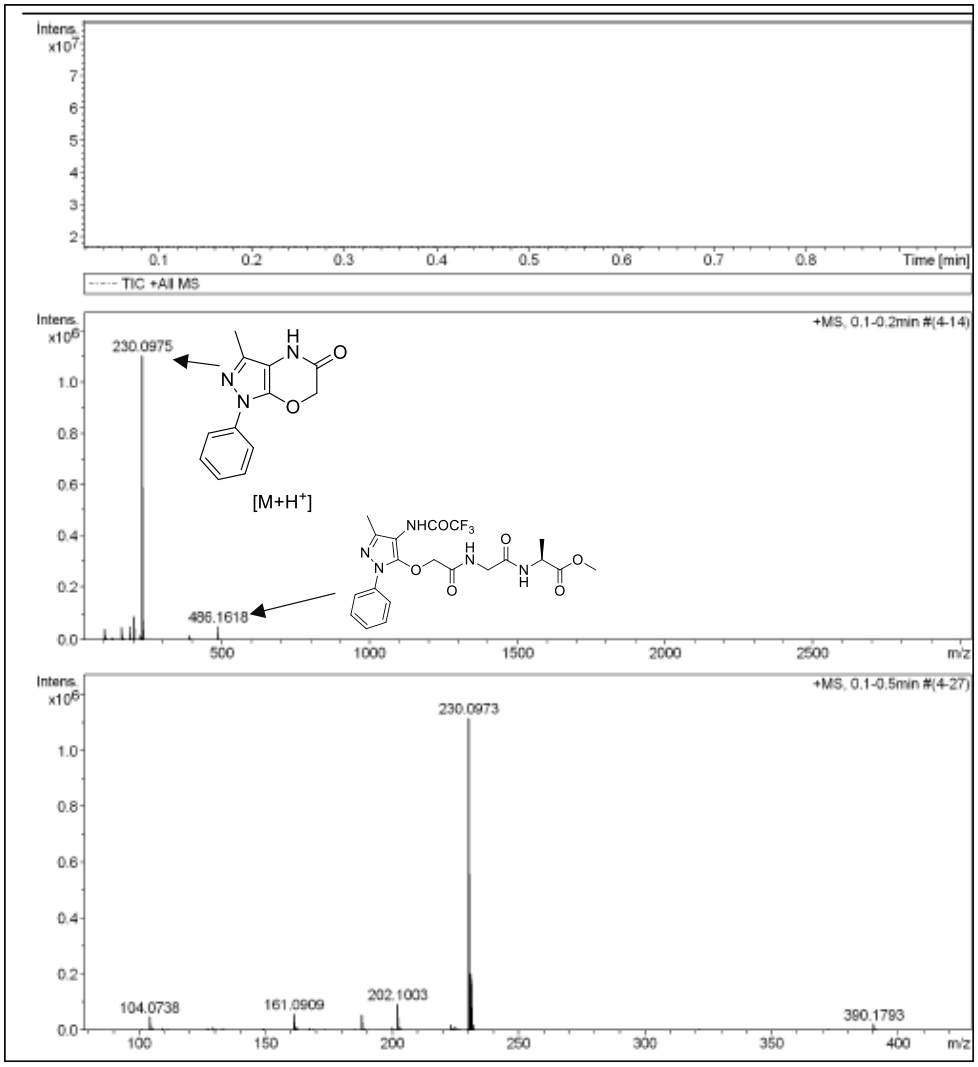
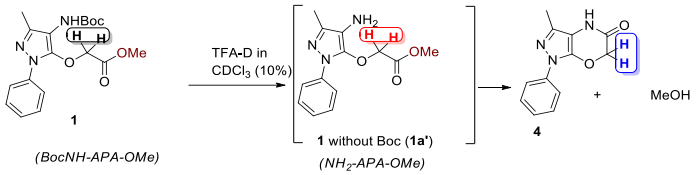


Figure S28. ESI-MS spectra of compound **3e** after time dependent ^{1}H-NMR in 10% TFA/$CDCl_3$

24.Time dependent ^{1}H-NMR spectra of ester cyclization in compound **1**



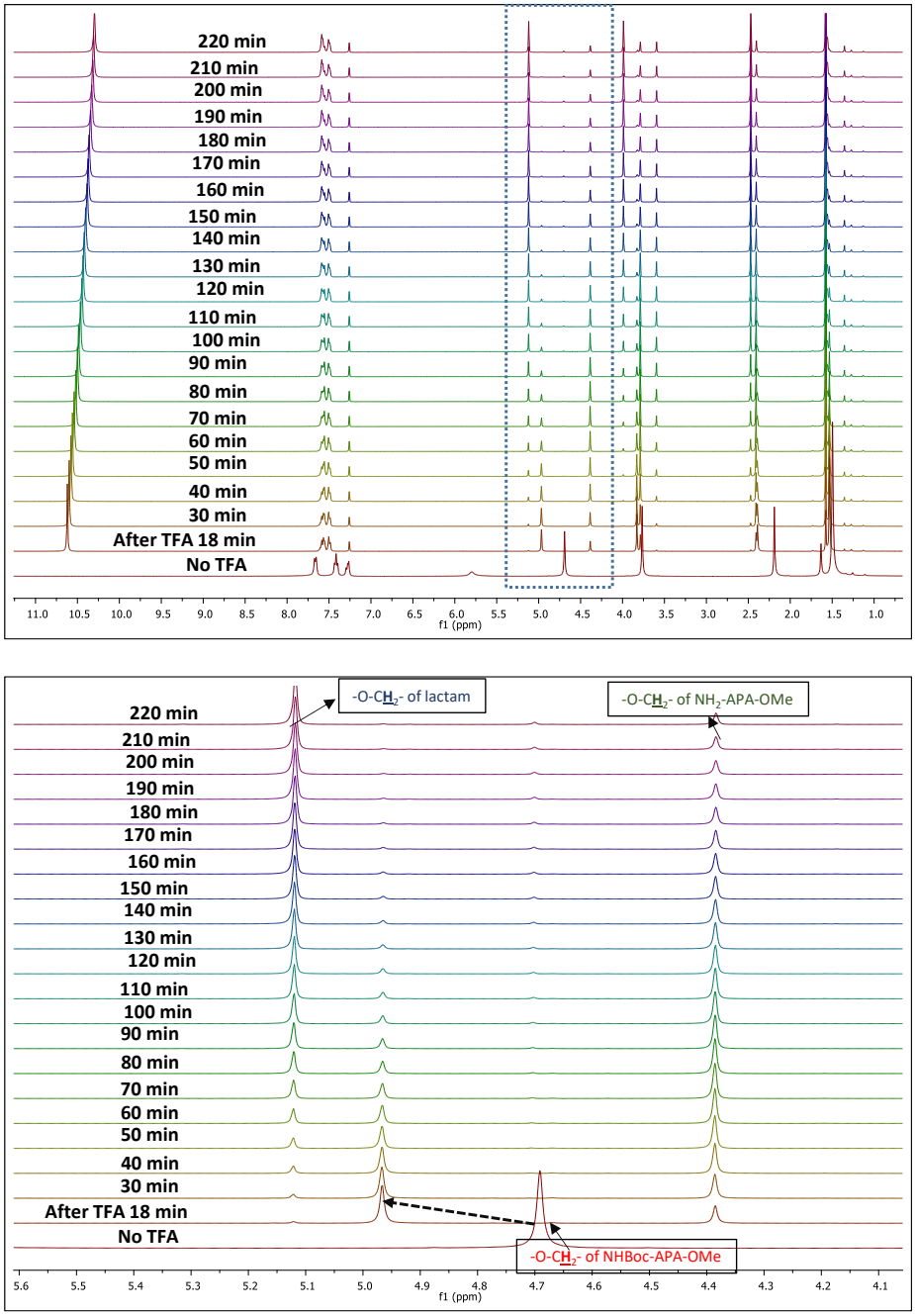


Figure S29. Time dependent ^{1}H-NMR spectra of compound **1** in 10% TFA/$CDCl_3$

25.Time dependent ^{1}H-NMR spectra of cyclization of compound **2**

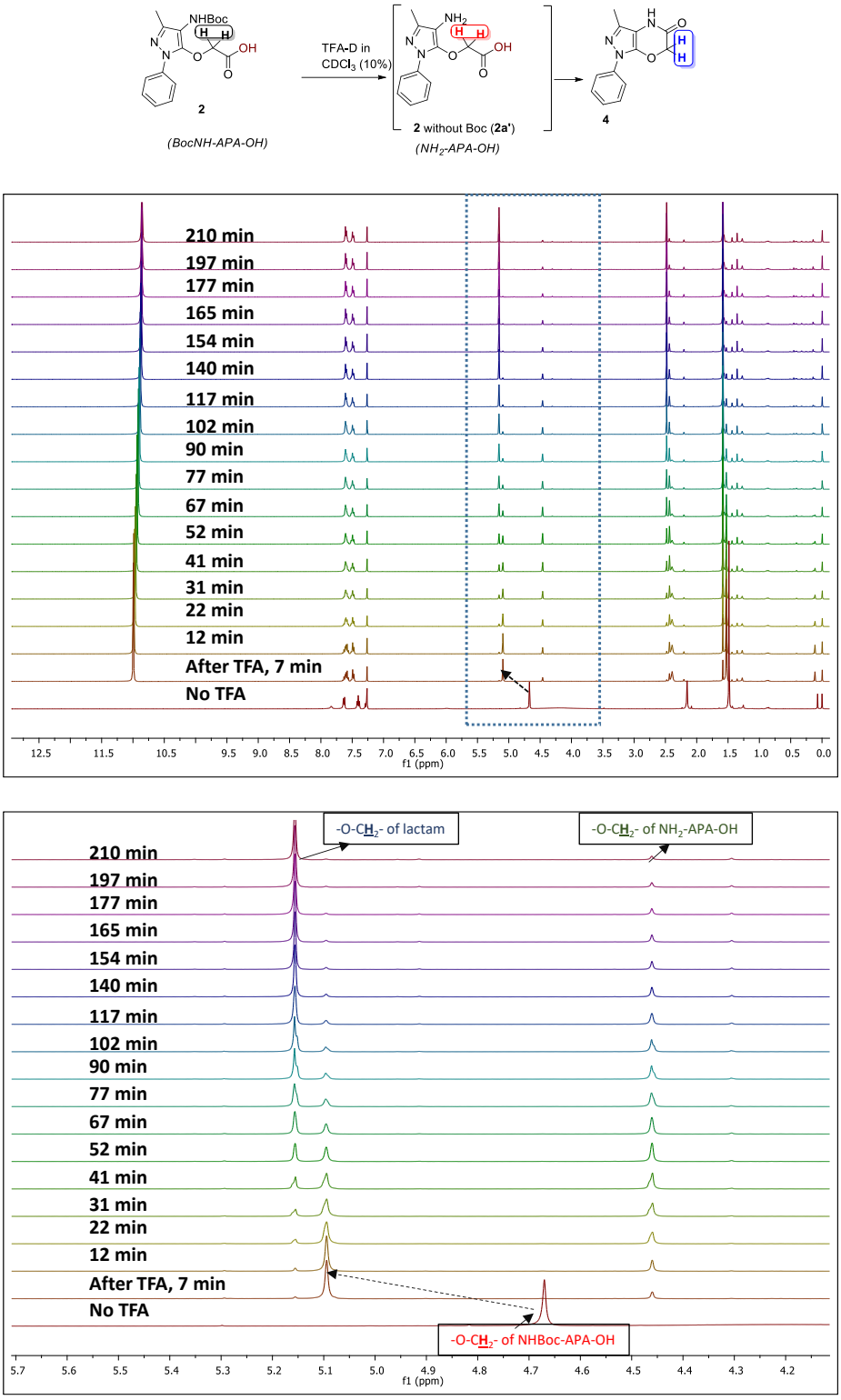


Figure S30. Time dependent ^{1}H-NMR spectra of compound **2** in 10% TFA/$CDCl_3$

26.ESI-MS spectra of compound **2** after time dependent NMR

After recording NMR added CD_3OD and recorded ESI-MS

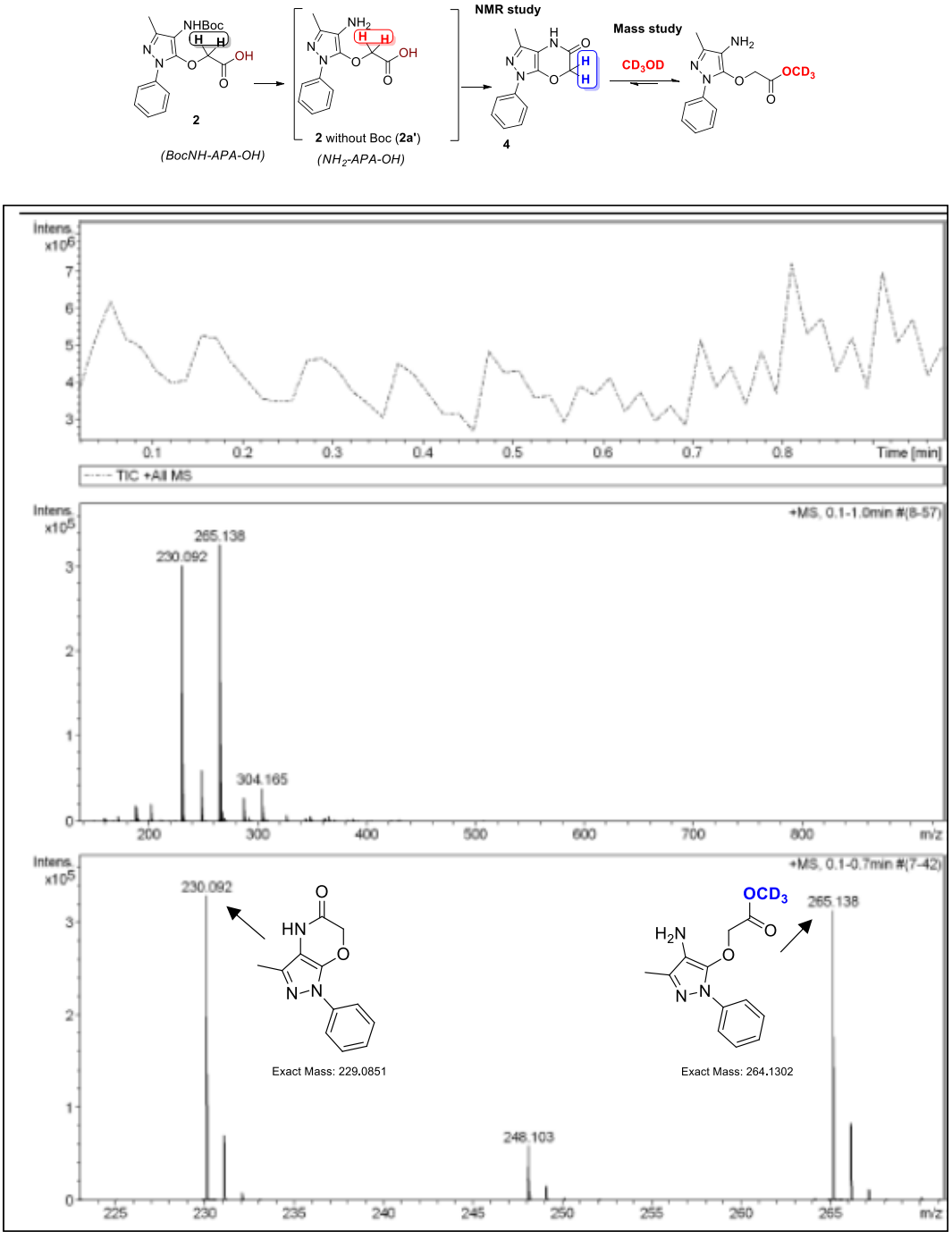


Figure S31. ESI-MS spectra of compound **2** after time dependent ^{1}H-NMR in 10% TFA/$CDCl_3$

(After recording NMR added CD_3OD and recorded ESI-MS)

27.Time dependent ^{1}H-NMR spectra of lactam **4**

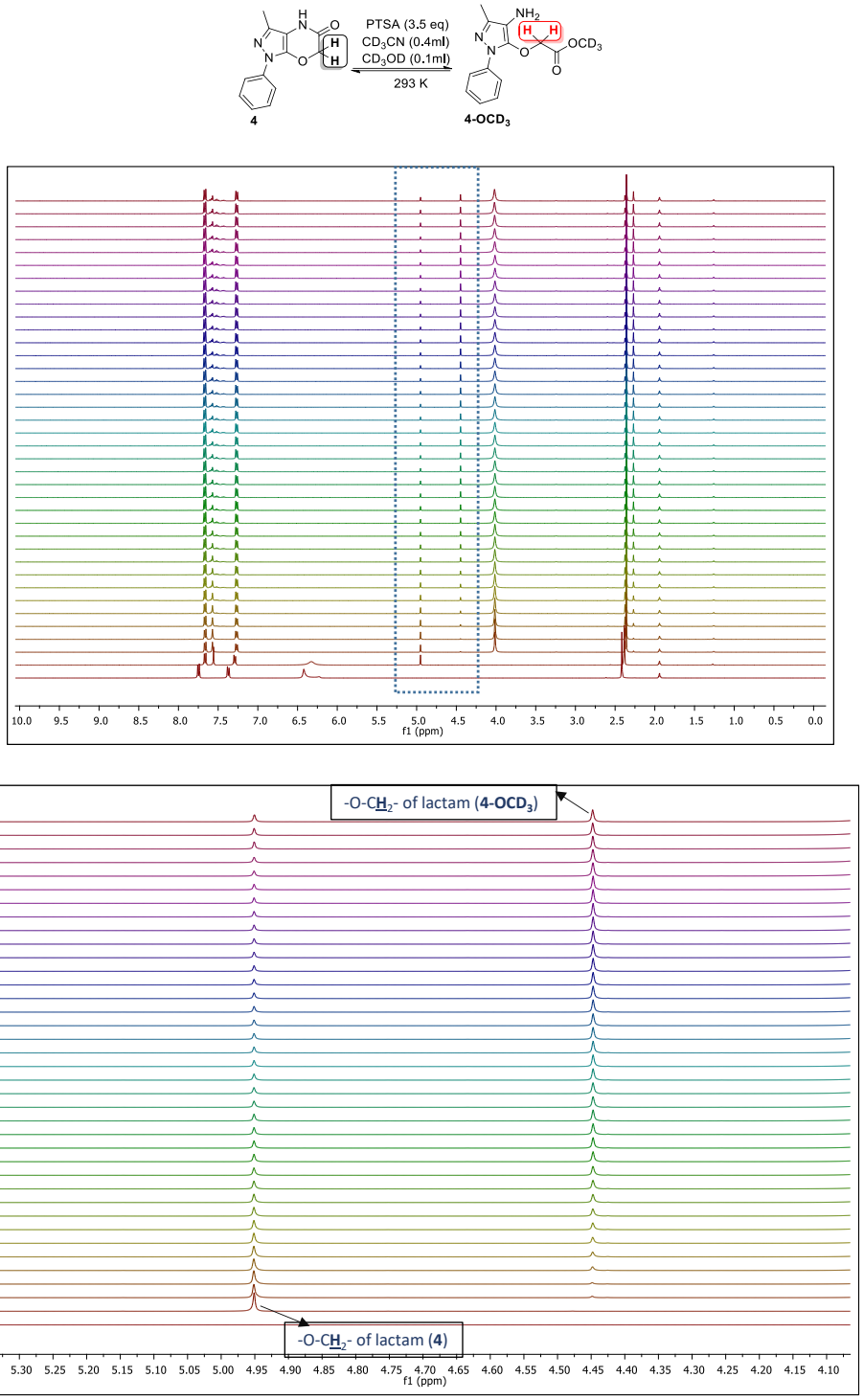
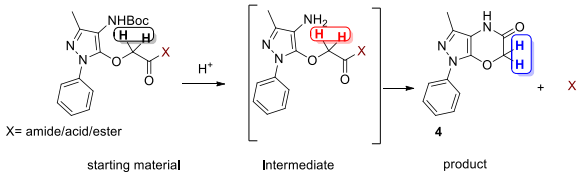


Figure S32. Time dependent ^{1}H-NMR spectra of lactam **4** in PTSA/$CDCl_3$/CD_3OD at 293K for every 5-8 min.

28.Determination of rate constants (k_1), (k_2) and half-life ($t_{1/2}$) of APA derivatives

Procedure for rate constant determination for compounds **1**, **2**, & **3a**:

5 mg of APA peptides/ester/acid was taken in 450 µL of $CDCl_3$ solution and studied with ^{1}H-NMR (Figure S27, S29, and S30) before and after adding TFA-D (10%) with time, the spectra were acquired for every ~ 7-10 min. All the spectra were baseline corrected and adjusted the solvent residual peak, TMS to zero. For precise integration all the spectra were overlapped and integrated using MestReNova NMR software. The mole fractions (x) were calculated from their desired peaks from area under the anticipated peak. From integrated NMR spectra, the mole fraction (x) of –O-C**H**$_2$- in APA peptide and their respective intermediates and products vs time is plotted. The integral plots of mole fraction, (x) vs Time were used to calculate the rate constants. After seeing these plots, we propose that compounds **1**, **2**, & **3a** are following consecutive first order kinetic pathway.[2-4] and their rate laws are shown as:



$$A \xrightarrow{k_1} B \xrightarrow{k_2} C$$

$$[A] = [A]_0\, e^{-k_1 t} \qquad (1)$$

$$[B] = \frac{k_1[A]_0}{k_2 - k_1}\,(e^{-k_1 t} - e^{-k_2 t}) \qquad (2)$$

$$[C] = [A]_0\left(1 - \frac{k_2}{k_2 - k_1}\, e^{-k_1 t} + \frac{k_1}{k_2 - k_1} e^{-k_2 t}\right) \qquad (3)$$

A_0 = Initial of concentration of starting material (A) at time zero, (initial integral value)

T = Time; k_1 and k_2 = rate constants of the reaction;

Half-life of reaction ($t_{1/2}$) = ln (2)/k

Rate constant and half-life of compound **3a**:

Herein we calculated the rate constant for hydrolysis of APA peptide **3a** with 10% TFA in $CDCl_3$ by time dependent ^{1}H-NMR studies. We have extracted integration of –O-C**H**$_2$- of APA peptide/intermediate/product at different interval of time by considering the CH_3 residue on pyrazole ring of APA peptide **3a** as reference with integration of 3.0.

Table S1. Integration of the –O-C**H**$_2$- at different time points from their ^{1}H-NMR spectra.

S.No	Time (min)	sm	intd	prdt	Total integral value	*x* (sm)	*x* (intd)	*x* (prdt)	ln *x* of sm	ln *x* of intd	ln *x* of prdt
1	0	2	0	0	2	1	0	0	0		
2	11	1.61	0.34	0.03	1.98	0.81	0.17	0.02	-0.21	-1.76	-4.19
3	18	1.36	0.57	0.06	1.99	0.68	0.29	0.03	-0.38	-1.25	-3.5
4	23	1.2	0.7	0.09	1.99	0.6	0.35	0.05	-0.51	-1.04	-3.1
5	34	0.89	0.9	0.19	1.98	0.45	0.45	0.1	-0.8	-0.79	-2.34
6	40	0.76	0.96	0.23	1.95	0.39	0.49	0.12	-0.94	-0.71	-2.14
7	47	0.63	1.01	0.31	1.95	0.32	0.52	0.16	-1.13	-0.66	-1.84
8	53	0.55	1.03	0.36	1.94	0.28	0.53	0.19	-1.26	-0.63	-1.68
9	75	0.31	1.01	0.59	1.91	0.16	0.53	0.31	-1.82	-0.64	-1.17
10	84	0.26	0.97	0.68	1.91	0.14	0.51	0.36	-1.99	-0.68	-1.03
11	94	0.2	0.92	0.78	1.9	0.11	0.48	0.41	-2.25	-0.73	-0.89
12	110	0.14	0.82	0.92	1.88	0.07	0.44	0.49	-2.6	-0.83	-0.71
13	124	0.1	0.73	1.03	1.86	0.05	0.39	0.55	-2.92	-0.94	-0.59
14	137	0.07	0.65	1.12	1.84	0.04	0.35	0.61	-3.27	-1.04	-0.5
15	152	0.05	0.57	1.21	1.83	0.03	0.31	0.66	-3.6	-1.17	-0.41
16	173	0.04	0.47	1.32	1.83	0.02	0.26	0.72	-3.82	-1.36	-0.33
17	188	0.03	0.4	1.4	1.83	0.02	0.22	0.77	-4.11	-1.52	-0.27
18	202	0.02	0.35	1.45	1.82	0.01	0.19	0.8	-4.51	-1.65	-0.23
19	219	0.01	0.29	1.5	1.8	0.01	0.16	0.83	-5.19	-1.83	-0.18

sm = starting material, intd = intermediate, prdt = product, *x* = mole fraction

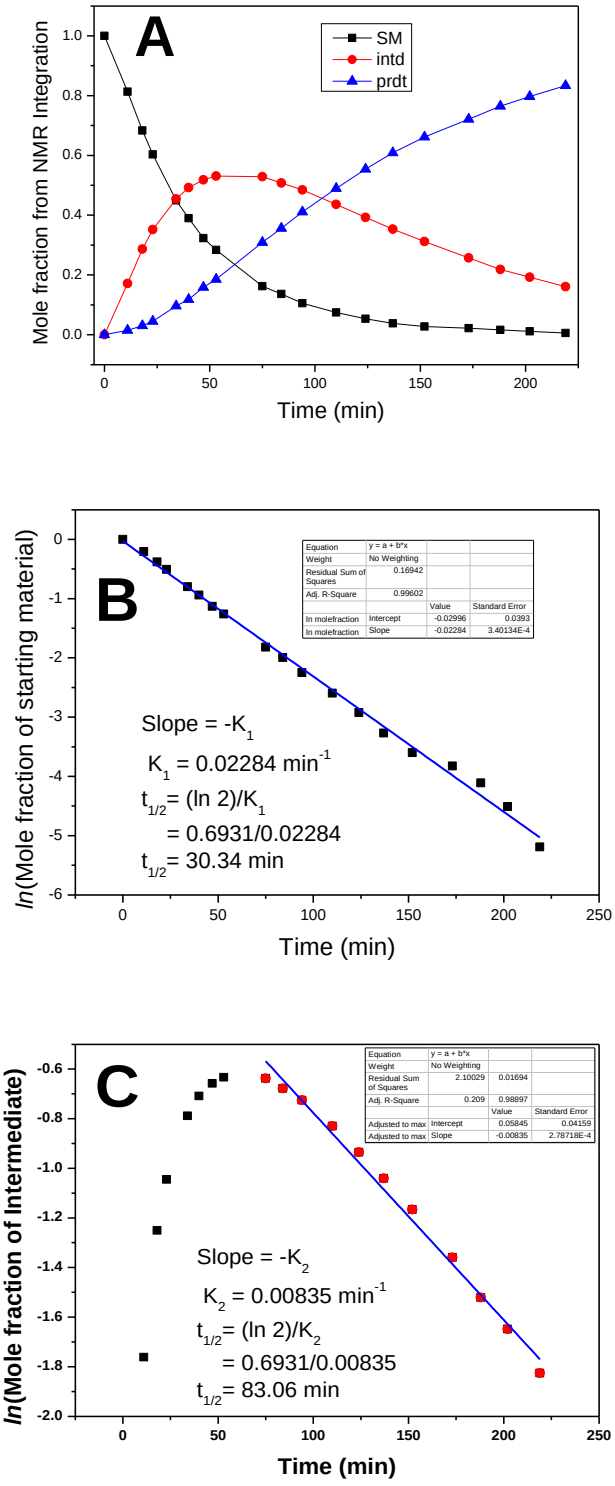


Figure S33. Rate constants and half-life of amide hydrolysis (formation of lactam **4**) in **3a**

*Rate constant and half-life of compound **2**:*

Herein we calculated the rate constant of hydrolysis of compound **2** with 10% TFA in $CDCl_3$ by time dependent ^{1}H-NMR studies. We have extracted integration of the –O-C**H**$_2$- at different interval of time by considering the CH_3 residue on pyrazole ring as reference with integration 3.0

Table S2. Integration of the –O-C**H**$_2$- at different time points from their ^{1}H-NMR spectra.

S.No	Time (min)	sm	intd	prdt	Total integral value	x (sm)	x (intd)	x (prdt)	ln x of sm	ln x of intd	ln x of prdt
1	0	2	0	0	2	1.00	0.00	0.00	0.00		
2	7	1.55	0.26	0.11	1.92	0.81	0.14	0.06	-0.21	-2.00	-2.86
3	12	1.42	0.4	0.12	1.94	0.73	0.21	0.06	-0.31	-1.58	-2.78
4	22	1.14	0.55	0.22	1.91	0.60	0.29	0.12	-0.52	-1.24	-2.16
5	31	0.94	0.64	0.32	1.9	0.49	0.34	0.17	-0.70	-1.09	-1.78
6	41	0.77	0.68	0.45	1.9	0.41	0.36	0.24	-0.90	-1.03	-1.44
7	52	0.62	0.68	0.59	1.89	0.33	0.36	0.31	-1.11	-1.02	-1.16
8	67	0.47	0.65	0.78	1.9	0.25	0.34	0.41	-1.40	-1.07	-0.89
9	77	0.38	0.61	0.89	1.88	0.20	0.32	0.47	-1.60	-1.13	-0.75
10	90	0.3	0.56	1.03	1.89	0.16	0.30	0.54	-1.84	-1.22	-0.61
11	102	0.24	0.49	1.15	1.88	0.13	0.26	0.61	-2.06	-1.34	-0.49
12	117	0.18	0.43	1.28	1.89	0.10	0.23	0.68	-2.35	-1.48	-0.39
13	140	0.11	0.33	1.44	1.88	0.06	0.18	0.77	-2.84	-1.74	-0.27
14	154	0.09	0.28	1.51	1.88	0.05	0.15	0.80	-3.04	-1.90	-0.22
15	165	0.07	0.24	1.56	1.87	0.04	0.13	0.83	-3.29	-2.05	-0.18
16	177	0.06	0.21	1.61	1.88	0.03	0.11	0.86	-3.44	-2.19	-0.16
17	197	0.03	0.16	1.68	1.87	0.02	0.09	0.90	-4.13	-2.46	-0.11
18	210	0.01	0.14	1.71	1.86	0.01	0.08	0.92	-5.23	-2.59	-0.08

sm = starting material, intd = intermediate, prdt = product, *x* = mole fraction

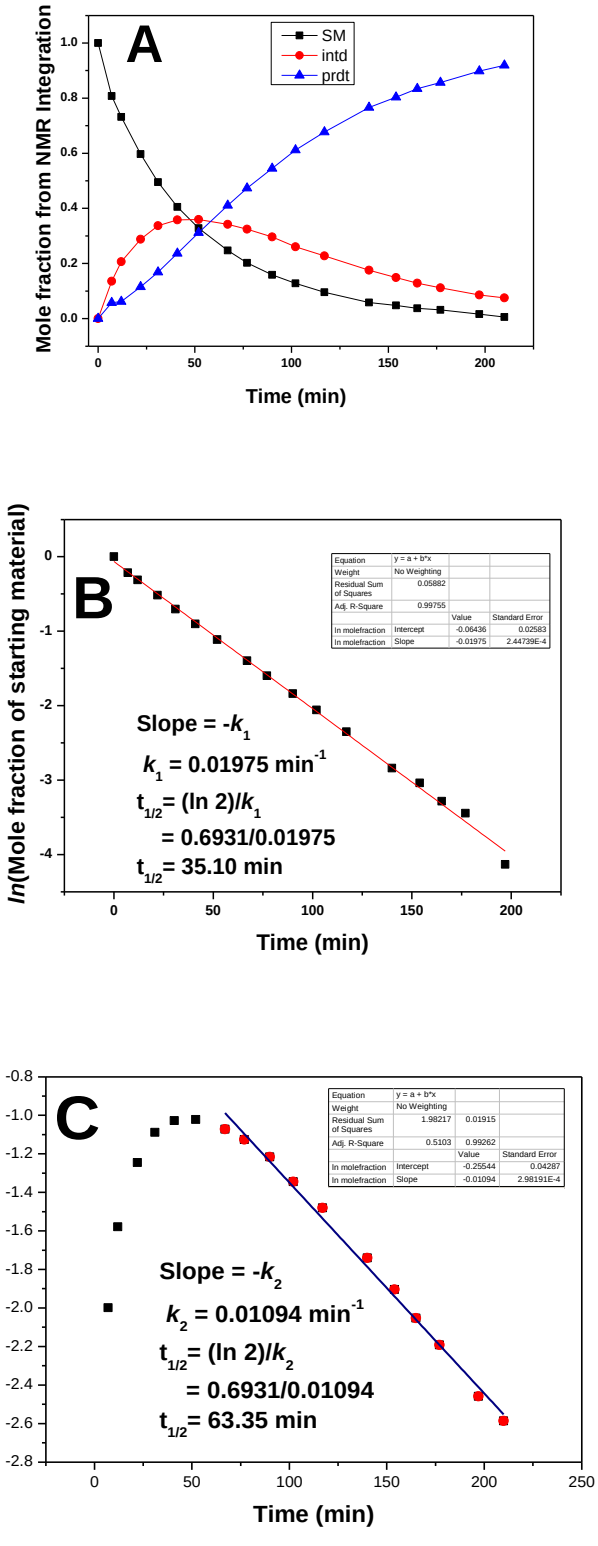


Figure S34. Rate constant and half-life of acid hydrolysis (formation of lactam **4**) in compound **2**

*Rate constant and half-life of compound **1**:*

Herein we calculated the rate constant of hydrolysis of compound (**1**) with 10% TFA in $CDCl_3$ by Time dependent ^{1}H-NMR studies. We have extracted integration of the APA peptide –O-C**H**$_2$- at different interval of time by considering the CH_3 proton of methyl group on pyrazole ring as reference with integration 3.0.

Table S3. Integration of the –O-C**H**$_2$- at different time points from their ^{1}H-NMR spectra

S.No	Time (min)	sm	intd	prdt	Total integral value	*x* of sm	x of intd	x of prdt	ln x of sm	ln x of intd	ln x of prdt
1	0	2	0	0	2.00	1.00	0.00	0.00	0.00		
2	18	1.43	0.62	0.08	2.13	0.67	0.29	0.04	-0.40	-1.23	-3.28
3	23	1.13	0.82	0.14	2.09	0.54	0.39	0.07	-0.61	-0.94	-2.70
4	30	1.04	0.86	0.16	2.06	0.50	0.42	0.08	-0.68	-0.87	-2.56
5	35	0.89	0.94	0.22	2.05	0.43	0.46	0.11	-0.83	-0.78	-2.23
6	40	0.81	0.95	0.26	2.02	0.40	0.47	0.13	-0.91	-0.75	-2.05
7	45	0.72	0.99	0.31	2.02	0.36	0.49	0.15	-1.03	-0.71	-1.87
8	50	0.65	1.01	0.36	2.02	0.32	0.50	0.18	-1.13	-0.69	-1.72
9	55	0.58	1.03	0.42	2.03	0.29	0.51	0.21	-1.25	-0.68	-1.58
10	60	0.53	1.04	0.46	2.03	0.26	0.51	0.23	-1.34	-0.67	-1.48
11	65	0.47	1.03	0.51	2.01	0.23	0.51	0.25	-1.45	-0.67	-1.37
12	70	0.43	1.03	0.56	2.02	0.21	0.51	0.28	-1.55	-0.67	-1.28
13	75	0.38	1.01	0.61	2.00	0.19	0.51	0.31	-1.66	-0.68	-1.19
14	80	0.35	1.01	0.67	2.03	0.17	0.50	0.33	-1.76	-0.70	-1.11
15	85	0.31	0.98	0.72	2.01	0.15	0.49	0.36	-1.87	-0.72	-1.03
16	90	0.28	0.96	0.77	2.01	0.14	0.48	0.38	-1.97	-0.74	-0.96
17	95	0.25	0.94	0.82	2.01	0.12	0.47	0.41	-2.08	-0.76	-0.90
18	100	0.23	0.92	0.86	2.01	0.11	0.46	0.43	-2.17	-0.78	-0.85
19	105	0.21	0.89	0.9	2.00	0.11	0.45	0.45	-2.25	-0.81	-0.80
20	110	0.19	0.86	0.95	2.00	0.10	0.43	0.48	-2.35	-0.84	-0.74
21	115	0.17	0.83	1	2.00	0.09	0.42	0.50	-2.47	-0.88	-0.69
22	120	0.15	0.8	1.04	1.99	0.08	0.40	0.52	-2.59	-0.91	-0.65
23	125	0.13	0.78	1.08	1.99	0.07	0.39	0.54	-2.73	-0.94	-0.61
24	130	0.13	0.75	1.11	1.99	0.07	0.38	0.56	-2.73	-0.98	-0.58
25	135	0.11	0.73	1.15	1.99	0.06	0.37	0.58	-2.90	-1.00	-0.55
26	140	0.1	0.7	1.19	1.99	0.05	0.35	0.60	-2.99	-1.04	-0.51
27	145	0.09	0.67	1.23	1.99	0.05	0.34	0.62	-3.10	-1.09	-0.48
28	150	0.08	0.65	1.26	1.99	0.04	0.33	0.63	-3.21	-1.12	-0.46
29	155	0.07	0.62	1.29	1.98	0.04	0.31	0.65	-3.34	-1.16	-0.43
30	160	0.07	0.59	1.32	1.98	0.04	0.30	0.67	-3.34	-1.21	-0.41
31	165	0.07	0.57	1.34	1.98	0.04	0.29	0.68	-3.34	-1.25	-0.39
32	170	0.06	0.55	1.38	1.99	0.03	0.28	0.69	-3.50	-1.29	-0.37
33	175	0.06	0.52	1.41	1.99	0.03	0.26	0.71	-3.50	-1.34	-0.34
34	180	0.05	0.5	1.44	1.99	0.03	0.25	0.72	-3.68	-1.38	-0.32
35	185	0.04	0.48	1.47	1.99	0.02	0.24	0.74	-3.91	-1.42	-0.30
36	190	0.03	0.46	1.48	1.97	0.02	0.23	0.75	-4.18	-1.45	-0.29
37	195	0.04	0.43	1.51	1.98	0.02	0.22	0.76	-3.90	-1.53	-0.27
38	200	0.03	0.42	1.53	1.98	0.02	0.21	0.77	-4.19	-1.55	-0.26
39	205	0.04	0.4	1.55	1.99	0.02	0.20	0.78	-3.91	-1.60	-0.25
40	210	0.03	0.38	1.58	1.99	0.02	0.19	0.79	-4.19	-1.66	-0.23
41	215	0.03	0.36	1.59	1.98	0.02	0.18	0.80	-4.19	-1.70	-0.22
42	220	0.03	0.34	1.6	1.97	0.02	0.17	0.81	-4.18	-1.76	-0.21

sm = starting material, intd = intermediate, prdt = product, *x* = mole fraction

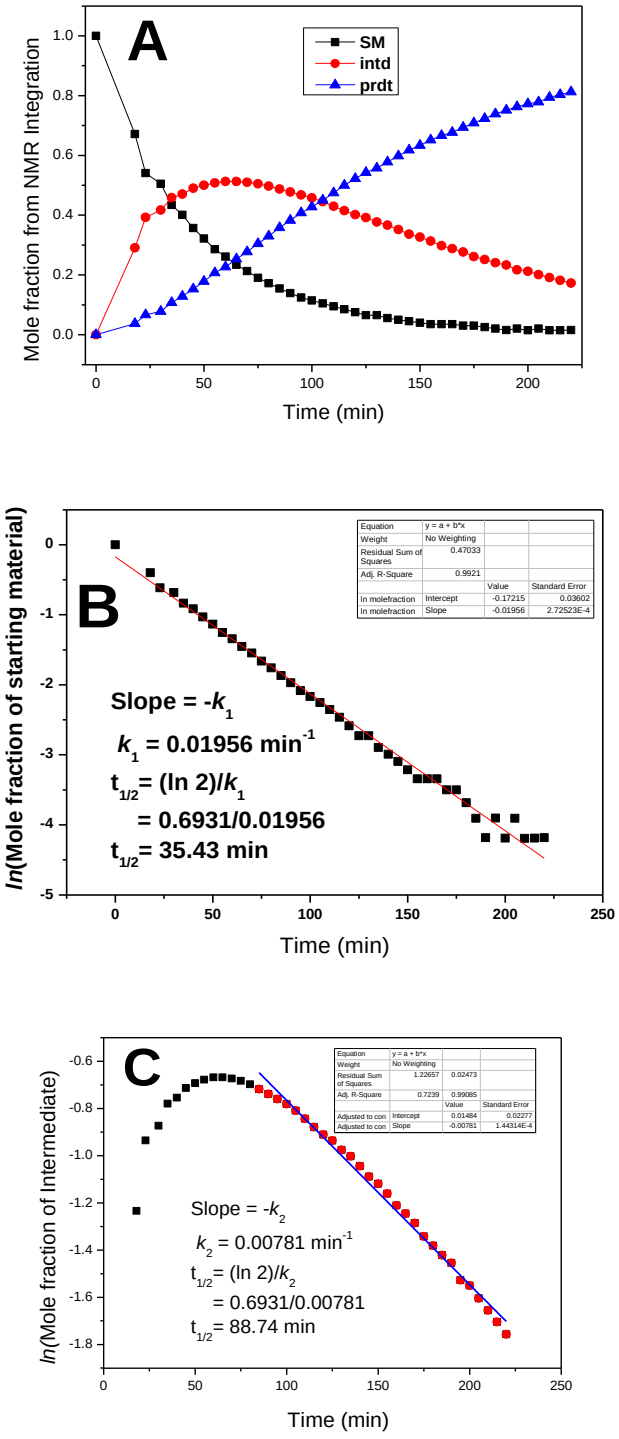
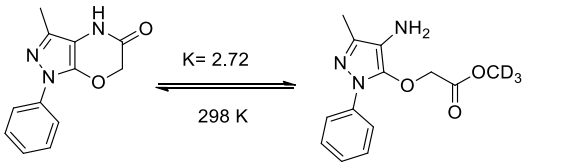


Figure S35. Rate constant and half-life of ester hydrolysis (formation of lactam **4**) in compound **1**

*Procedure for determination of equilibrium rate constant for compound **4**:*

5 mg of compound **4** was taken in 400µl of CD_3CN solution and studied with ^{1}H-NMR (Fig. S27, S29, and S30) before and after adding CD_3OD (100µl) and PTSA (3.5 equivalents) at different time intervals at 298 K, the spectra were acquired for every ~ 7-10 min. All the spectra were baseline corrected and adjusted the solvent residual peak, TMS to zero. For precise integration all the spectra were overlapped and integrated using MestReNova NMR software. The mole fractions (x) were calculated from their desired peaks from area under the anticipated peak. From integrated NMR spectra, the mole fraction (x) of –O-C**H**$_2$- in compound (**4**) and in their products vs time is plotted. The integral plots of mole fraction, (x) vs Time were used to calculate the rate constant. After seeing these plots, we propose that lactam **4** is in equilibrium with its hydrolysed form (**4-OCD$_3$**) and their rate laws are shown as. [2-4]



$$A \underset{k_-}{\overset{k_+}{\rightleftharpoons}} B$$

$$K = \frac{k_+}{k_-} = \frac{[B]}{[A]} \qquad (4)$$

K = equilibrium constant

k_+ = rate constant for forward reaction

k_- = rate constant for backward reaction

[A] = concentration of **4** (in mole fraction, calculated from ^{1}H-NMR)

[B] = concentration of **4-OCD$_3$** (in mole fraction, calculated from ^{1}H-NMR)

Table S4. Integration of the –O-C**H**$_2$- at different time points from their ^{1}H-NMR spectra

S.No	Time (min)	*x* of sm	*x* of prdt
1	0	1.00	0.00
2	5	0.98	0.02
3	16	0.88	0.12
4	23	0.74	0.26
5	29	0.67	0.33
6	36	0.61	0.39
7	40	0.55	0.45
8	47	0.52	0.48
9	53	0.49	0.51
10	58	0.46	0.54
11	65	0.43	0.57
12	70	0.41	0.59
13	76	0.40	0.60
14	80	0.38	0.62
15	86	0.37	0.63
16	91	0.36	0.64
17	96	0.35	0.65
18	101	0.34	0.66
19	105	0.33	0.67
20	110	0.33	0.67
21	115	0.32	0.68
22	122	0.32	0.68
23	128	0.31	0.69
24	134	0.31	0.69
25	141	0.30	0.70
26	153	0.30	0.70
27	163	0.29	0.71
28	176	0.29	0.71
29	186	0.29	0.71
30	195	0.29	0.71
31	203	0.28	0.72
32	351	0.28	0.72
33	446	0.27	0.73
34	508	0.28	0.72

sm = starting material, intd = intermediate, prdt = product, *x* = mole fraction

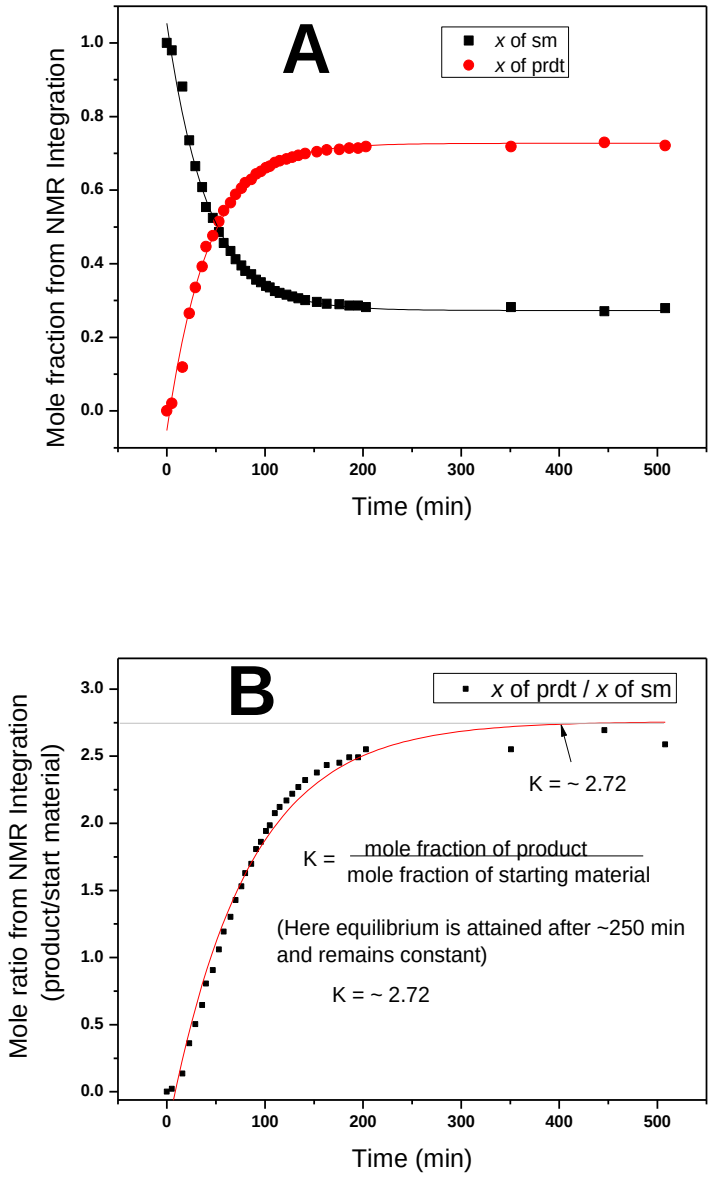
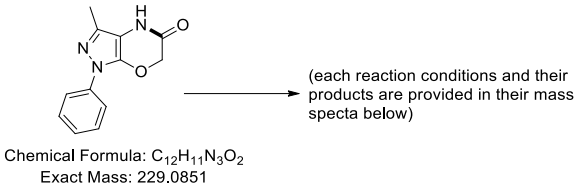
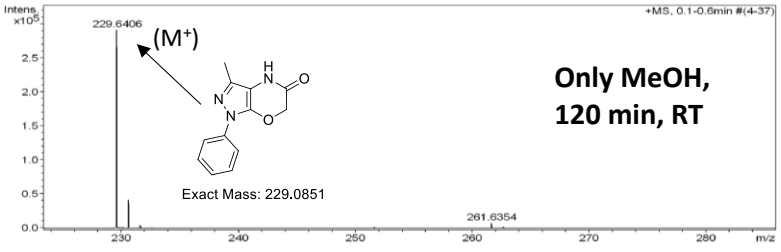


Figure S36. Rate constant for lactam **4**

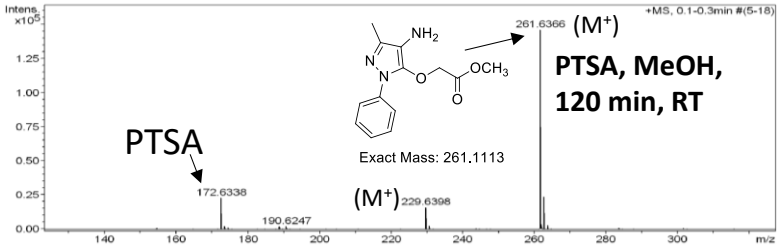
29.Mass study of amide bond cleavage in lactam **4**



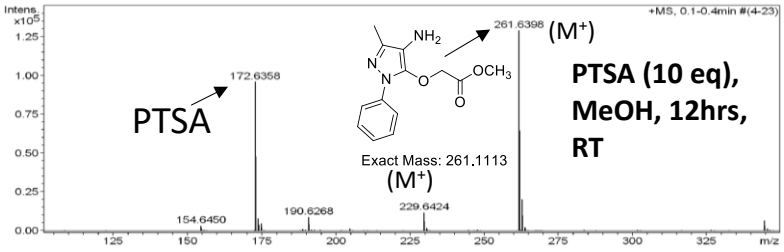
a



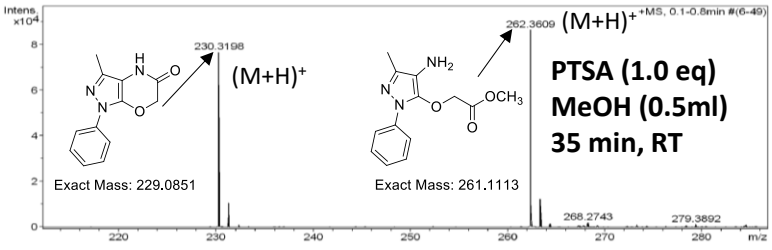
b



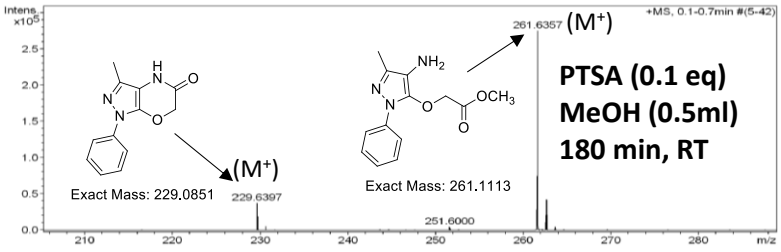
c



d



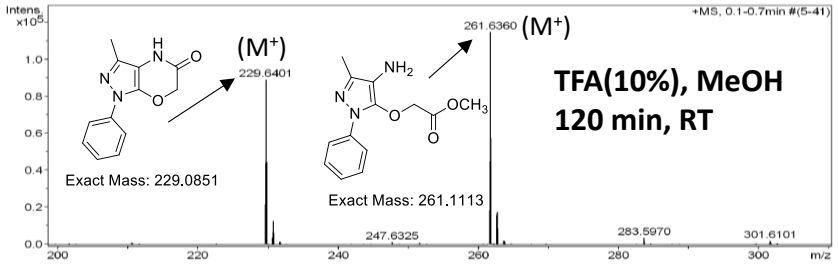
e



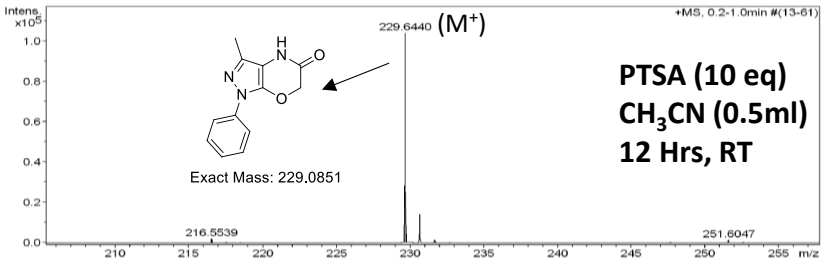
b

(Continued)

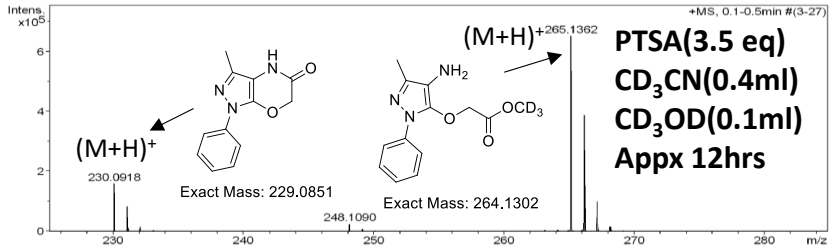
f



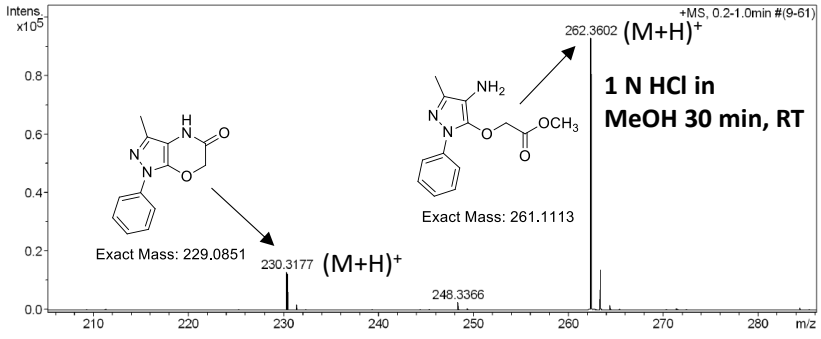
g



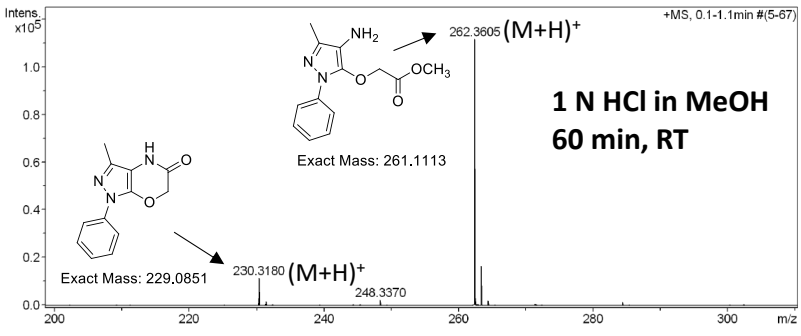
h



i



j



(Continued)

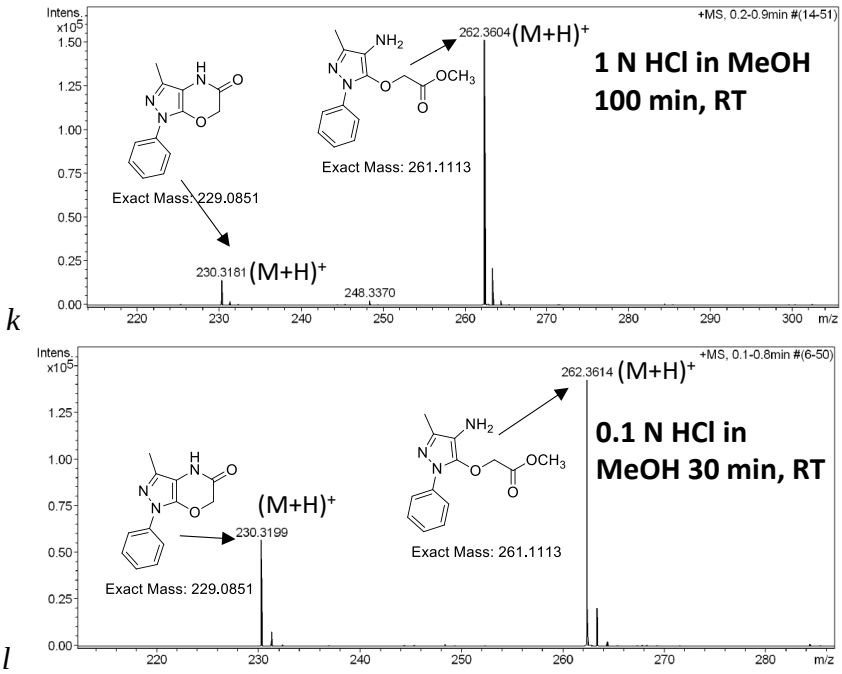


Figure S37. Reactivity of Amide bond of lactam (**4**) in various reaction conditions by mass spectra.

30.Crystallography

2-3 mg of compound **4** was dissolved in EtOAc, for solubility a drop of MeOH was added and left for slow evaporation at room temperature to obtain off-white crystals. A suitable single crystal of each compound **4** was carefully selected under a polarizing microscope. Single crystal structure determination by X–ray diffraction was performed on a Siemens SMART-CCD diffractometer equipped with a normal focus, 2.4 kW sealed-tube X-ray source (*Mo–Kα* radiation, λ = 0.71073Å) operating at 50 kV and 30 mA. Structures were solved by the direct method using SHELXT 2014 and refined on *F*2 by a full–matrix least-squares technique using the SHELXL 2014 programs package. An empirical absorption correction based on symmetry equivalent reflections was applied using SADABS. The graphic programs DIAMOND3.2 was used to draw the structures. Non-hydrogen atoms were refined anisotropically. In the refinement, hydrogens were treated as riding atoms using the SHELXL default parameters. The crystal structure of compound **1**, **3b** and **3h** are deposited to the Cambridge crystallographic Data centre (CCDC) with CCDC with numbers CCDC 1831989, CCDC 1831991 and CCDC 1831992 respectively.[1]

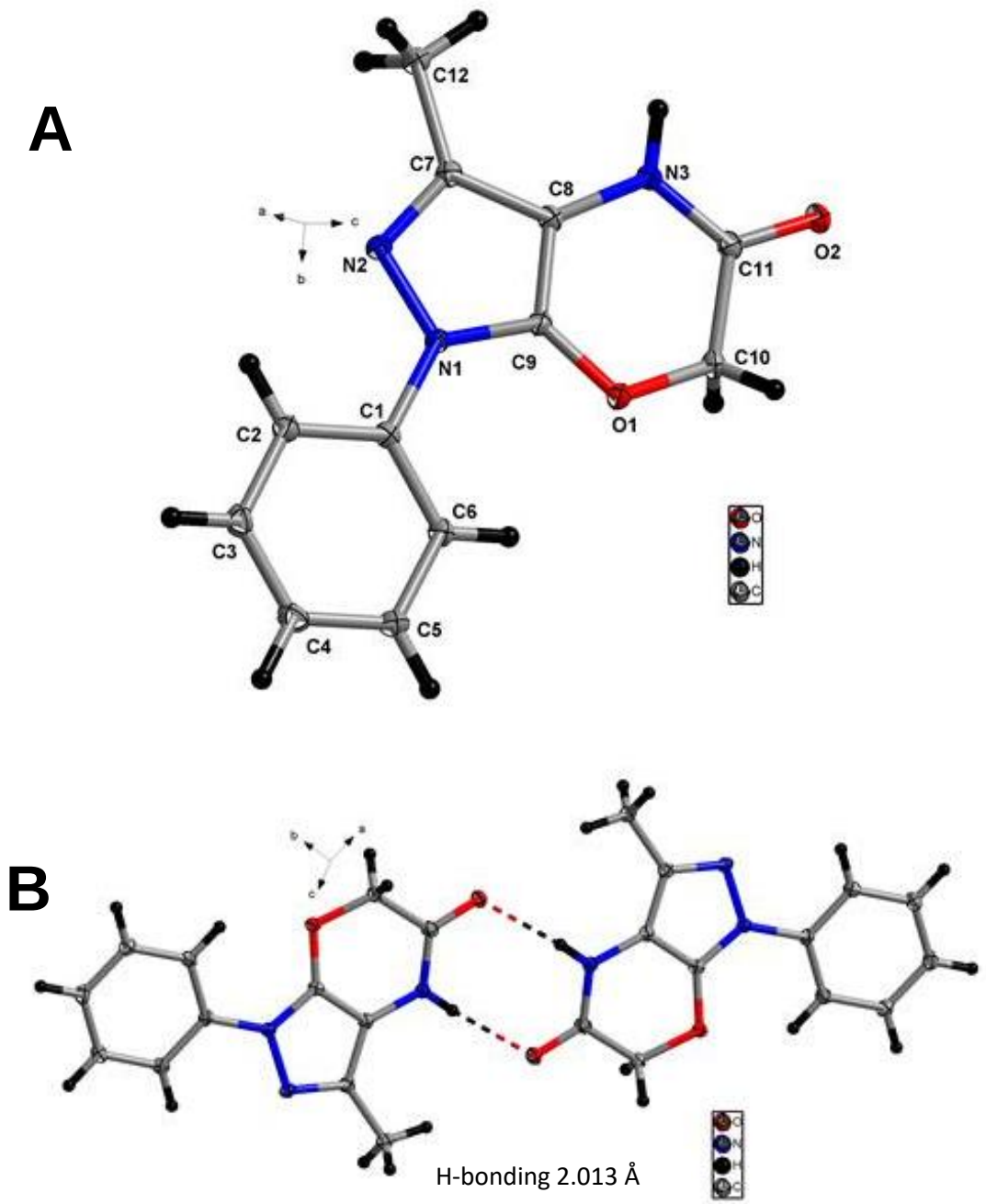


Figure S38. Single crystal X-ray structure of lactam **4**. Thermal ellipsoids with **3**0% ellipsoid probability (a), H-bonding (b).

Table S5 Crystallographic table for lactam **4**

Parameter	lactam **4**
Formula	C12 H11 N3 O2
*M*r	229.24
crystal system	Monoclinic
space group	*C 2/c*
a/Å	18.5240(6)
b/Å	10.8785(6)
c/Å	13.2463(5)
α/°	90°.
β/°	129.107(4).
γ/°	90°.
V/Å^3	2071.30(19)
Z	8
ρ_{calc} (gcm^{-3})	1.470
μ/ mm^{-1} μ(cm^{-1})	0.104
θ range (deg)	2.3 to 26.4°.
reflections collected	13877
Unique reflections[R(int)]	2133 [0.0359]
data/restraints/ parameters	2133 / 0 / 154
GOF on F^2	1.083
*R*1 and *R*2 [$I>2\sigma(I)$]	0.0450, 0.1200
*R*1 and *R*2 (all data)	0.0536, 0.1306
largest residual peaks (e.Å^{-3})	0.454 and -0.402
CCDC no	1890879

31.Structural parameters of peptides from crystal structures of APA derivatives

Table S6 Torsion angles and structural parameters of compounds **3h**/**3b**/**1**/**4**

Compound	phi (ϕ)	psi (ψ)	tau (τ)	Sum of bond angle at "N"	Sum of bond angle at "C=O"	C-N Bond length (in A)	C-O Bond length (in A))
3h	ϕ_1 (144.345)	ψ_1 (26.561)	5.538	360	359.99	1.305	1.196
	ϕ_2 (-135.735)	ψ_2 (139.39)	0.468	360	360	1.360	1.182
	ϕ_3 (-142.304)	ψ_3 (-169.368)					
3b[a]	ϕ_1 (133.536)	ψ_1 (-12.46)	5.5389(A) & 3.847 (B)	359.093 (A) & 359.663(B)	359.992 (A)	1.333	1.231
	ϕ_2 (-139.577)	ψ_2 (58.43)			& 359.99 (B)	1.322	1.228
1	ϕ_1 (-165.135)	ψ_1 (3.889)					
4	ϕ_1 (21.52)	ψ_1 (-12.303)	3.985	360	359.98	1.334	1.233

[a] unit cell has two molecules

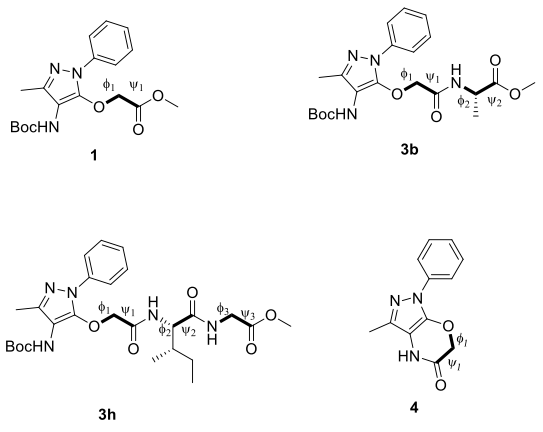


Figure S39. Torsion angles and structural parameters in compounds **1**, **3a**, **3h** and **4** from their crystal structures.

32.UV-spectra of lactam **4**

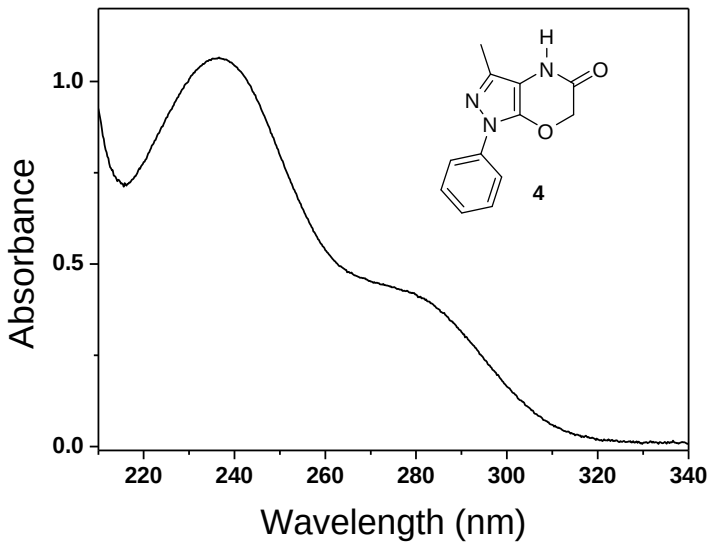
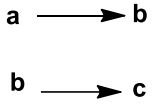


Figure S40. UV-spectra of lactam **4** in MeOH.

33.COPASI Modelling Procedure and Results

Kinetic modelling was done by using COPASI - COmplex PAthway SImulator software.[5] The model was created assuming two sequential first order kinetics. The observed rate constants were used to generate the concentration versus time fitting curves in the model, with initial volume of 0.5 ml and initial time, in minutes, as zero (t = 0) and used with following reaction model:



Plots were generated using Deterministic (LSODA) method in Time Course.

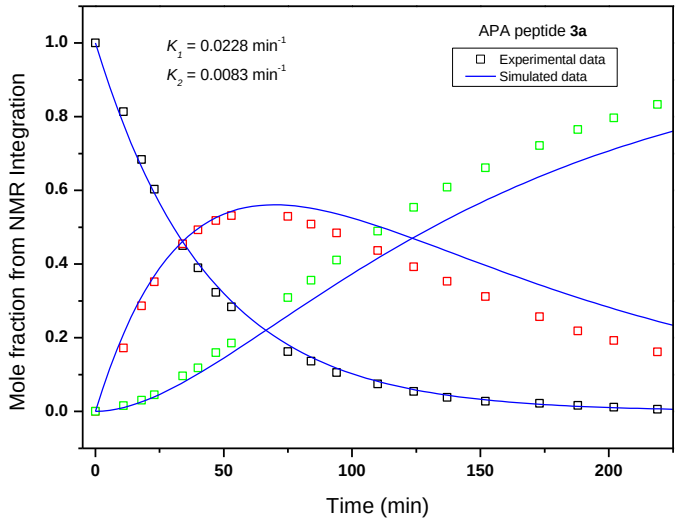


Figure S41. Kinetic modelled and experimental data fitting for APA peptide **3a**

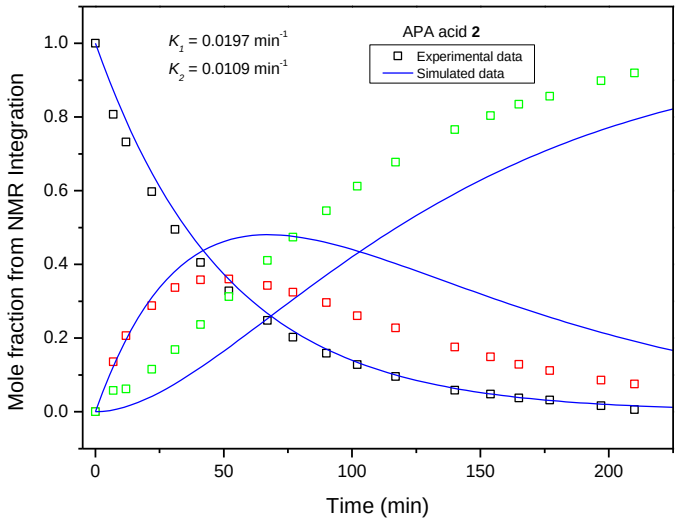


Figure S42. Kinetic modelled and experimental data fitting for APA acid **2**

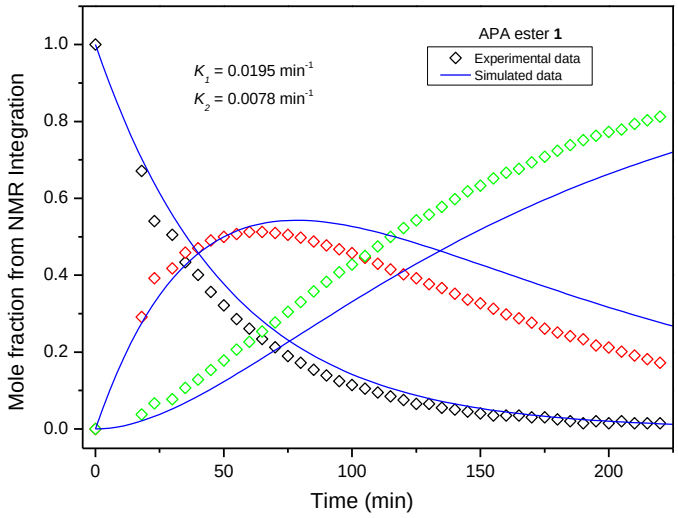


Figure S43. Kinetic modelled and experimental data fitting for APA ester **1**

ChemPubSoc Europe

DOI: 10.1002/cbic.201800822

CHEMBIOCHEM Full Papers

Tropolone-Conjugated DNA: Fluorescence Enhancement in the Duplex

Amarnath Bollu[a, b] and Nagendra K. Sharma*[a, b]

Dedicated to Prof. Krishna N. Ganesh, IISER-Tirupati, on his 65th birthday

Tropolone (2-hydroxycyclohepta-2,4,6-triene-1-one and tautomer) is a non-benzenoid bioactive natural chromophore with pH-dependent fluorescence character and extraordinary metal binding affinities, especially with transition-metal ions Cu^{2+}/Zn^{2+}/Ni^{2+}. This report describes the syntheses and biophysical studies of a new tropolonyl thymidine [(4(5)-hydroxy-5(4)-oxo-5(4)*H*-cyclohepta-1,3,6-trienyl)thymidine] (*tr*-T) nucleoside and of corresponding tropolone-conjugated DNA oligonucleotides that form B-form DNA duplex structures with a complementary DNA strand, although their duplex structures are less stable than that of the control. Furthermore, the stabilities of those DNA duplex structures are lowered by the presence of increasing numbers of *tr*-T residue or by decreasing pH of their environments. Most importantly, these duplex structures are made fluorescent because of the presence of the tropolone moieties conjugated to the thymidine residues. The fluorescence behavior of those duplex structures exhibits pH dependence, with stronger fluorescence at lower pH and weaker fluorescence at high pH. Importantly, the fluorescence characters of *tr*-DNA oligonucleotides are significantly enhanced by nearly threefold after duplex structure formation with their complementary control DNA oligonucleotide. Further, the fluorescence behavior of these *tr*-DNA duplex structures is also dependent on the pH conditions. Hence, tropolonyl-conjugated DNA represents a class of new fluorescent analogues that might be be employed for sensing DNA duplex formation and provide opportunities to improve fluorescence properties further.

Introduction

Tropolone is an α-hydroxy derivative of tropone, a non-benzenoid seven-membered-ring aromatic compound containing a carbonyl functional group.[1] Tropolone derivatives, tropolonoids, occur in nature mostly in plants and fungi.[2] Tropolone and related derivatives display an enormous range of biological activities, such as antibacterial,[3] antifungal,[4] antitumor,[5] and antiviral[6] properties. Recent highlights have included the demonstration that tropolones could be potent and selective inhibitors for enzymes dependent on zinc as a cofactor.[7] Tropolone has also been reported to be a distinctive metal-chelating scaffold, especially for Cu^{2+}/Zn^{2+}/Ni^{2+}, because of the α-hydroxycarbonyl functional group.[8]

The aromatic structure of tropolone exhibits interesting photophysical properties (absorptions and emission), and tropolone is considered a potential structural moiety for the design of small-molecule-derived chromophores. Absorption spectra of tropolone exhibit two peaks in the wavelength ranges of 200–260 nm and 260–380 nm, due to two types of electronic transitions: π–π* and n–π* transitions.[9]

Further, the pK_a of tropolone has been determined as pK_a = 6.5, and it is reported to be a weakly fluorescent compound.[10] Its fluorescence properties are described in terms of the existence of the (n–π*) singlet state. Martin and co-workers showed that the fluorescence properties of tropolone, at excitation wavelength (λ) 315 nm, are dependent on the solvent polarity and pH conditions, with quantum yields of tropolone having been measured as 0.005 in cyclohexane (neutral), 0.003 in acidic water (pH 3.0), and 0.005×10^{-2} under basic conditions (pH 12.0).[11] These variations are due to the different forms of tropolone structures, as neutral or anion forms, due to hydroxy protonation and deprotonation processes under different pH conditions (Scheme 1 A).

We thus realized that tropolone systems might be interesting scaffolds to incorporate into peptides and nucleic acids in order to tune their physical and structural properties. Recently we reported amino acids/peptides with tropolonyl motifs and established a role of the tropolonyl carbonyl system in conformational changes of peptides.[12] However, we have found that tropolone systems play roles in unusual cleavage reactions of adjacent amide bonds under mild pH conditions and reversible amidation under mild basic conditions.[12–13] To explore the photophysical properties of tropolone-related compounds, we also explored tropolone-derived BODIPY analogues and dem-

[a] A. Bollu, Dr. N. K. Sharma
School of Chemical Science
National Institute of Science Education and Research (NISER) Bhubaneswar
Jatani 752050, Odisha (India)
E-mail: nagendra@niser.ac.in

[b] A. Bollu, Dr. N. K. Sharma
Homi Bhabha National Institute (HBNI)
Training School Complex, Anushakti Nagar, Mumbai (India)

Supporting information and the ORCID identification numbers for the authors of this article can be found under https://doi.org/10.1002/cbic.201800822: ESI-HRMS and $^1H/^{13}C/^{31}P$ NMR spectra of all new compounds, HPLC of oligonucleotides, and single-crystal X-ray data for compound 4.

ChemPubSoc Europe

DOI: 10.1002/ejoc.201801640

EurJOC European Journal of Organic Chemistry

Full Paper

Amino Acids and Peptides

Synthesis and Conformational Analysis of *Aminopyrazolonyl* Amino Acid (APA)/Peptides

Amarnath Bollu[a] and Nagendra K. Sharma*[a]

Abstract: Pyrazole, pyrazolone, and aminopyrazolone derived molecules are bioactive molecules and considered as potential therapeutic drug candidates because of their unique structural properties. These molecules have abilities to interact with several bio-macromolecules via non-covalent interactions such as hydrogen bonding and π–π interactions. In structural organization of dipeptides, pyrazole containing aromatic amino acid/dipeptides have been explored and considered as potential amino acid residue. In repertoire of unnatural aromatic amino acids, this report describes the synthesis of 4-aminopyrazolonyl containing amino acids and their crystal structures. The incorporation of 4-aminopyrazolonyl at N-terminal of native amino acid/dipeptides influences the conformational changes of respective peptide which induces the formation of distinctive supramolecular self-assembly structures such as β-sheet and α-helices in their solid-state crystal. The structural conformation of those peptides, here, are also demonstrated in solution phase by ^{1}H-NMR (1D/2D) and [D$_6$]DMSO titration methods which support the formation of inter-/intramolecular hydrogen bonding in solution. Hence, these unnatural amino acid analogues can tune the secondary structure of natural amino acid/peptides by introducing at *N*-terminal via amide bond.

Introduction

Pyrazole, pyrazolone and aminopyrazolone are constituents of several natural products, synthetic drug molecules, agrochemical reagents, and metal chelating agents (Figure 1a).[1] For example, pyrazolone derivatives are well-known antipyretic and analgesic drugs for more than one century.[2] Many other pyrazolone derivatives are reportedly known as neuroprotective agents, HIV integrase inhibitors, phosphodiesterase inhibitors, and antibacterial agents.[3] Further the substituted pyrazolone derivative, 4-aminophenazone (Aminopyrene), and its metabolite (Ampyrone) have shown excellent analgesic, anti-inflammatory and antipyretic activities, though the some risk factors as agranulocytosis are also associated (Figure 1a).[4] The substituted pyrazole and pyrazolone derivatives have shown strong DNA binding affinities, and considered as potential anti-cancer agents.[5] For an example, the amino substituted pyrazolone molecule, ampyrone, has been modified as Schiff base derivatives and employed for metal complexation with metal ions (Co^{2+}/Ni^{2+}/Fe^{3+}) which are considered as antimicrobial agents.[6]

Further, ampyrone is explored as biochemical reagent for determination of phenol concentration including inside the cell.[7] Moreover, pyrazole and pyrazolone derivatives have abilities to form hydrogen bonding and have been explored in screening of their biological activities.[8] Thus pyrazole and pyrazolone

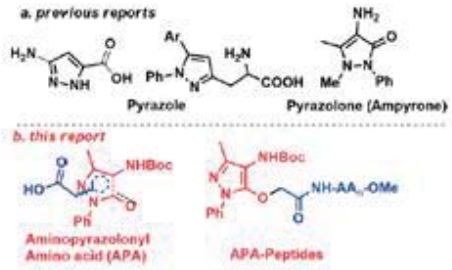


Figure 1. (a) Chemical structures of pyrazole/pyrazolone/aminopyrazolone; (b) rationally designed amino acids/peptides.

could be considered as potential scaffold for syntheses of unnatural aromatic amino acids for tuning the peptide structures. In literature, Schrader and co-workers have reported that aminopyrazole containing amino acids, are the artificial templates for stabilization of β-sheet conformation (Figure 1a).[9] Recently, Sutherland and co-workers have reported the new class of conjugated unnatural α-amino acid containing 5-arylpyrazole residue as probe of serine proteases (Figure 1a).[10] In repertoire of unnatural aromatic amino acids, herein we report the syntheses of rationally designed aminopyrazolonyl amino acid (APA) derivatives (Figure 1b) and their roles in structural organization of native dipeptides.

Results and Discussion

In Scheme 1, we began the syntheses of rationally designed unnatural aminopyrazolonyl amino acid (APA) from L-threonine.

[a] School of Chemical Sciences, National Institute of Science Education and Research (NISER)-Bhubaneswar, Jatni campus, Bhubaneswar-752050 (Odisha), India & HBNI-Mumbai, Mumbai, India
E-mail: nagendra@niser.ac.in
https://www.niser.ac.in/users/nagendra#profile-main

Supporting information and ORCID(s) from the author(s) for this article are available on the WWW under https://doi.org/10.1002/ejoc.201801640.

JOC The Journal of Organic Chemistry Cite This: J. Org. Chem. 2019, 84, 5596−5602 Article pubs.acs.org/joc

Cleavable Amide Bond: Mechanistic Insight into Cleavable 4-Aminopyrazolyloxy Acetamide at Low pH

Amarnath Bollu[†] and Nagendra K. Sharma*[,†,‡]

[†]School of Chemical Science, National Institute of Science Education and Research − (NISER) Bhubaneswar, Jatani 752050, Odisha, India

[‡]HBNI-Mumbai, Mumbai 400094, India

Supporting Information

ABSTRACT: The cleavage of amide bonds under mild acidic conditions is a rare chemical event. *N*-Acetamide bond of peptides is extremely stable even under the strongest organic acid trifluoromethanesulfonic acid. This report mechanistically describes a new cleavable amide bond in 4-aminopyrazolyloxy acetamide peptide analogues under mild acidic conditions such as trifluoroacetic acid (10−20%) or HCl (0.1−4.0 N) at room temperature, and the formation of unusual lactam from 4-aminopyrazolyloxy acetic acid after evaporation of solvent. This is a rare chemical event in peptide bond, which could be explored as acid-sensitive protecting group of free amines.

INTRODUCTION

Natural amides are extremely stable, and their half-life for spontaneous hydrolysis is estimated to be 350−600 years at neutral pH and room temperature.[1] The carbonyl group of amide are poor electrophiles because of the typical resonance stability. Although amides can readily be cleaved/hydrolyzed by enzymes such as proteases, it is difficult to cleave the C−N bond of an amide selectively using synthetic chemistry. The cleavage/hydrolysis of amide bonds usually requires heating under strongly acidic or basic conditions. However, the cyclic amides, lactams, are easily cleavable compared to the linear amides because of ring-strained amides.[2] A number of highly strained lactams are synthesized and found that their amide bonds are cleavable under mild conditions. Since the C=O bond of reactive lactams amide is arising to the strong electrophile as "ketonic" carbonyl because of the resonance decoupling through N−C=O torsion. Brown and co-workers mechanistically confirmed that resonance decoupling induces a remarkable enhancement in hydrolysis rate by direct nucleophilic attack at the carbonyl of strained amide.[3] The most twisted amide is trimethylazatricyclodecanone, which is a highly strained lactam with large resonance decoupling through N−C bond, and is found as readily hydrolyzable. This twisted lactam also shows the dual reactivity of amide, i.e., the nucleophilic character of amine and the electrophilic nature of carbonyl. Booker-Milburn and co-workers reported the solvolysis/substitution reaction of acyclic synthetic amides at room temperature and neutral conditions via the formation of ketene intermediates.[4] In the literature, cleavage of few natural amide bonds is precisely demonstrated through the formation of thiazolinone derivatives (by Edman degradation), oxazolinium intermediates, and chemical modification of the amide backbone.[5] Most importantly, Edman degradation for cleavage of amide bond is the foundation for protein sequencing technology.[5a] The presence of zinc species in the active sites of metalloproteases has inspired synthetic chemists to cleave amide bond using Zn^{2+}. Mashima and co-workers reported the cleavage of amides bearing a *β*-hydroxyethyl group under mild conditions with Lewis acid $Zn(OTf)_2$.[6] Kostic and co-workers prepared an artificial peptidase, which undergoes cleavage of the sequence-specific amide bonds as Pro-Met/Pro-His segments of peptide under mild conditions with Pd catalyst.[7] In the repertoire of making reactive amide group, Hauk, Garg, and co-workers established the conversion of amide functional group into ester group by cleaving C−N bond of amide using Ni catalyst.[8] The conversion of an amide into another amide with different amine group, transamidation, is an important chemical transformation reaction for ligation or removal of amino-functionalized chemical moieties at carboxylate-functionalized molecules.[9] Gellman, Stahl, and co-workers explored the amide groups for transamidation reaction, between secondary amines and tertiary amines, using Zr/Hf-catalyst.[9a] Other metal-catalyzed and metal-free transamidation reactions are also reported.[9b,c,10] Recently, we have also established the transamidation reaction in tropolonylalkylglycinate peptide derivatives with different amino acid/peptides.

We have recently found the instability of amide bonds containing *N*-tropolonylaminoalkylglycinate (Trag) in short peptides under mild acidic conditions. The formation of stable tropolonium cations followed by the formation of ketene intermediate is the main cause for the cleavage of Trag-amide bonds.[11] In the repertoire of unnatural heterocyclic aromatic amino acids, we have explored the syntheses and supramolecular self-assembly studies of pyrazolone-containing amino acid analogues. One of those analogues is 4-amino-

Received: February 23, 2019
Published: April 10, 2019

DOI: 10.1021/acs.joc.9b00535
J. Org. Chem. 2019, 84, 5596−5602