

Environment-sensitive nucleotide analogs serve as good substrates and probes to study the activity of nucleic acid processing enzymes

विद्या वाचस्पति की
उपाधि की अपेक्षाओं की आंशिक पूर्ति में प्रस्तुत शोध प्रबंध

A thesis submitted in partial fulfillment of the requirements of the degree of
Doctor of Philosophy

द्वारा / By

पुलक घोष / Pulak Ghosh

पंजीकरण सं./Registration No.: 20162018

शोध प्रबंध पर्यवेक्षक/Thesis Supervisor:

प्रो. सीरगाज़ी जी. श्रीवत्सन / Prof. Seergazhi G. Srivatsan



भारतीय विज्ञान शिक्षा एवं अनुसंधान संस्थान पुणे

**INDIAN INSTITUTE OF SCIENCE EDUCATION
AND RESEARCH PUNE**

2024

CERTIFICATE

This is to certify that the work incorporated in the thesis entitled "*Environment-sensitive nucleotide analogs serve as good substrates and probes to study the activity of nucleic acid processing enzymes*" submitted by **Pulak Ghosh** (ID No.20162018) was carried out by the candidate at the Indian Institute of Science Education and Research (IISER), Pune, under my supervision. The work presented here or any part of it has not been included in any other thesis submitted previously for the award of any degree or diploma from any other University or Institution.



Date: 12/03/2024

Professor Seergazhi G. Srivatsan

Place: Pune

DECLARATION

I declare that the content presented entitled “**Environment-sensitive nucleotide analogs serve as good substrates and probes to study the activity of nucleic acid processing enzymes**” submission reflects my own thoughts expressed in my own language. In instances where I have incorporated external ideas, I have appropriately cited and referenced the original sources. I certify that I have not misrepresented, fabricated, or falsified any information in my submission, and I uphold the principles of academic honesty and integrity. I am aware that any violation of these principles may lead to disciplinary action by the institute and could result in penalties from sources that have not been accurately cited or from which proper permission has not been obtained.



Pulak Ghosh

Date: 11/03/2024

Place: Pune

Reg. No. 20162018

Acknowledgments

It brings me immense joy to convey my gratitude to Prof. Seergazhi G. Srivatsan, my esteemed research supervisor, for his exceptional mentorship during my MS-PhD journey at IISER Pune. I thank him for his invaluable guidance, unwavering support, and continuous motivation, which are pillars of my academic pursuit. I highly regard his innovative approach to research problems and his adeptness in initiating new research projects, which have provided me with a platform for growth as a successful researcher. On a personal level, his boundless energy and positive mindset have consistently inspired and encouraged the cultivation of fresh ideas. I am particularly indebted to Prof. Srivatsan for performing the crystallography experiment at Konstanz University, Germany, during his tenure as a senior Alexander Von-Humboldt project fellow. This opportunity significantly enriched my academic journey.

I thank my RAC members, Dr. Amrita Hazra and Dr. Janesh Kumar, for their valuable feedback and suggestions. I also want to thank my ACBI co-mentor, Prof. Motonari Uesugi from Kyoto University, for his valuable suggestions for my research work. I especially acknowledge Prof. Andreas Marx, Prof. Kay Diederichs, Dr. Karin Betz, Dr. Samra Ludmann, and Dr. Heike M. Kropp, for helping us perform all the X-ray crystallography experiments in their lab at Konstanz University, Germany. I also thank Dr. Saravanan Palani from IISC Bangalore for providing the TdT DNA polymerase. I would like to sincerely thank all the chemistry faculties at IISER Pune who taught me during the MS program. I thank Dr. Siddhesh S. Kamat for helping me share the lab space with us after the fire accident. I especially thank Dr. Sudha Rajamani for using the Typhoon gel imager in their lab and helping me during the fire accident. I would like to take this opportunity to thank Dr. Gayathri Pananghat, who helped me during the protein purification. I thank the former Director of IISER Pune, Prof. K.N. Ganesh, Prof. Jayant B. Udgoankar, and the current Director, Prof. Sunil S. Bhagwat, for the facilities provided at IISER Pune.

I am profoundly grateful to both my current and former colleagues in the lab, including Dr. Anupam, Dr. Pramod, Dr. Ashok, Dr. Vyankat, Dr. Sudeshna, Dr. Jerrin, Dr. Manisha, Dr. Saddam, Dr. Ragini, Dr. Manisha, Dr. Apeksha, Dr. Pankaj, Akanksha, Samikshakiran, Sarupa, Swagata, Amit, Arindam, Harsh, Vivek, and Punyaprava. Their collective contribution has created a delightful and memorable lab environment throughout this extensive journey. A special acknowledgment goes to all current lab members for their exceptional hard work in rebuilding the lab following the fire incident and sustaining its operations thereafter. I extend my sincere appreciation to Dr. Anupam, Dr. Jerrin, and Dr. Sudeshna for their senior guidance and support during my Ph.D. pursuit.

I thank members of Dr. Gayathri Pananghat's lab, especially Apurba, and Akhilesh, for helping me during the TdT purification. I also thank Bindu B. from Dr. Saravanan Palani's lab at IISC Bangalore

for giving me the TdT enzyme. I thank Rupam and Suman from Prof. Jayant B. Udgoankar's lab for their extensive help in the ESI-MS analysis of DNA. I express gratitude to the individuals who have contributed behind the scenes, including instrument operators, administrative personnel, and other non-teaching staff. I would also like to thank all my colleagues at IISER Pune, whose names I may have unintentionally omitted.

I am indebted to my friends Debanjan, Saikat, Krishnendu, Sukanya, Indra, Joy, Moushaki, Attek, and Prakash for making my life in IISER Pune memorable. I also thank the whole cricket team of Dadagiri and the football team FCP for giving me enjoyable PhD life.

Finally, I thank my extended family in IISER, Sarupa, Ragini, Anindita, and Writakshi for their patience and support throughout this journey. I extend special thanks to Sarupa for her unwavering support and encouragement, as she consistently assisted me from sharing lab space, discussing life events, and creating unforgettable memories to the completion of my thesis work, guiding me through all the ups and downs of my life.

I extend heartfelt appreciation to my parents and family for their boundless love and care, shaping me into who I am today. I owe a special debt of gratitude to my father, who served as an unsung hero throughout my entire academic journey. I wish to express my deepest gratitude to my mother, a constant source of inspiration, who has not only instilled in me the essence of patience and dedication in life but has also been a guiding light, encouraging me in my research endeavors. I thank all my school teachers who have played a pivotal role in shaping my personality and knowledge, from my early school days to my Ph.D. journey, without whom my current position would not have been possible. Finally, I offer my heartfelt gratitude to Lord Shiva for bestowing me the blessings of good health and fortune, enriching my daily life with divine favor and well-being.

Pulak Ghosh

Contents

Contents -----	i
Abbreviations -----	iv
Synopsis -----	vi
List of publications -----	xv

Chapter 1: Nucleotide analogs as substrates and probes for nucleic acid processing enzymes

1.1	Introduction	1
1.2	Enzymatic incorporation of different modified nucleotides and their applications	3
1.2.1	Sugar-modified nucleotides	4
1.2.2	Nucleobase-modified nucleotides	4
1.2.2.1	Incorporation of fluorescent tags	4
1.2.2.2	Incorporation of post-synthetic modifications	5
1.2.2.3	Metabolic labeling and cell imaging	6
1.2.2.4	Next-Generation sequencing	7
1.3	Template-independent labeling of nucleic acids	8
1.4	Numerous biophysical techniques and chemical probes to reveal the nucleic acid-protein interactions	9
1.4.1	X-ray crystallography	9
1.4.2	Nuclear Magnetic Resonance (NMR)	11
1.4.3	Fluorescence spectroscopy	13
1.4.3.1	label-free methods	13
1.4.3.2	Covalently incorporation of exogenous fluorophores	14
1.5	Base-modified fluorescence nucleoside analogs (FNAs) for studying nucleic acid-protein interactions	15
1.5.1	Size-expanded fluorescent nucleobase analogs	16
1.5.2	Size-extended fluorescent nucleobase analogs	17
1.5.3	Fluorescent poly aromatic hydrocarbon (PAH) and pteridine nucleobase analogs	19
1.5.4	Isomorphous nucleobase analogs	21
1.6	Motivation and outline of the thesis	23
1.7	References	26

Chapter 2: Microenvironment-sensitive fluorescent nucleotide probes from benzofuran, benzothiophene and selenophene as substrates for DNA polymerases

2.1	Introduction	39
2.2	Result and discussions	41
2.2.1	Design and synthesis of fluorescent nucleotide analogs	41
2.2.2	DNA polymerases show varying incorporation efficiency	43

2.2.3	Single nucleotide incorporation and site-specific labeling.	47
2.2.4	Responsive fluorescent BFdU provides insights on the enzyme– primer•template duplex complex formation	49
2.2.5	Nucleotide substrate itself serves as a probe to report its incorporation by KlenTaq	54
2.2.6	X-ray structure of the ternary complex sheds light on the incorporation behavior	56
2.2.7	Structural data correlates with KOD and KlenTaq incorporation and elongation behavior	59
2.3	Conclusion	61
2.4	Experimental sections	62
2.5	References	72
2.6	Appendix-I NMR and mass spectra of synthesized compounds	77

Chapter 3: Probing the processing of environment-sensitive nucleotide analogs by DNA polymerases using fluorescence, ¹⁹F NMR and X-ray crystallography

3.1	Introduction	90
3.2	Result and discussions	92
3.2.1	Synthesis of FBFdUTP 3	92
3.2.2	DNA polymerases incorporate the nucleotide analog with varying efficiency	94
3.2.3	Site-specific labeling of FBFdUTP 3	95
3.2.4	Binary complex (duplex-enzyme) formation by fluorescence	97
3.2.5	Nucleotide substrate fluorescently reports its incorporation into DNA ONs	103
3.2.6	Binary complex (duplex-enzyme) formation and nucleotide incorporation by ¹⁹ F NMR	104
3.2.7	X-ray crystal structures of ternary complexes provide insights on the incorporation mechanism	108
3.2.7.1	Processivity of KlenTaq depends on the size of the heterocycle- modified nucleotides	112
3.2.8	Size-dependency in incorporating multiple modifications in PCR reactions by KlenTaq and KOD DNA polymerases	114
3.3	Conclusion	116
3.4	Experimental sections	117
3.5	References	125
3.6	Appendix-II NMR and Mass spectra of synthesized compounds	129

Chapter 4: Microenvironment-responsive fluorescent nucleotide probes serve as good substrates to investigate the activity of terminal deoxynucleotidyl transferase (TdT)

4.1	Introduction	133
4.2	Result and discussions	136
4.2.1	Modified nucleotide substrates (1–5)	136

4.2.2 Incorporation of modified deoxynucleotides by TdT	137
4.2.3 Incorporation of modified dideoxynucleotides depends on different initiator primers	139
4.2.4 Competition between natural and modified nucleotide incorporation	140
4.2.5 Insight into the binary complex formation by fluorescence	143
4.2.6 Detection of TdT activity by fluorescence	146
4.2.7 Monitoring the competition with ddTTP by fluorescence	148
4.3 Conclusion	149
4.4 Experimental sections	150
4.5 References	159
4.6 Appendix-III NMR and Mass spectra of synthesized compounds	163
 Overall conclusion and outlook	 183
 Copyrights and Permission	 185

Abbreviations

ϵ	molar extinction coefficient	EDTA	ethylenediaminetetraacetic acid
μL	microliter	<i>em</i>	emission
μg	microgram	EMSA	electrophoretic mobility shift assay
μM	micro-molar	ESI-MS	electrospray ionization mass spectrometry
A	adenine	<i>ex</i>	excitation
AAC	azide-alkyne cycloaddition	EPR	electron paramagnetic resonance
ACN	acetonitrile	FAM	fluorescein amidite
AMP	adenosine monophosphate	FRET	fluorescence resonance energy transfer
AMdU	5-azidomethyl deoxyuridine	FPLC	fast protein liquid chromatography
ATP	adenosine triphosphate	G	guanosine
C	cytosine	GTP	guanosine triphosphate
CD	circular dichroism	HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
CTP	cytidine triphosphate	HPLC	high performance liquid chromatography
CuAAC	copper(I)-catalyzed alkyne-azide cycloaddition	HFIPA	hexafluoro-2-propanol
CuSO ₄	copper sulphate	<i>in vitro</i>	outside living organism
Cy2	cyanine 2	<i>in vivo</i>	inside living organism
Cy3	cyanine 3	LB	lysogeny broth
Cy5	cyanine 5	K _d	dissociation constant
dA	2'-deoxy adenosine	LB	lysogeny broth
dATP	2'-deoxy adenosine triphosphate	MALDI-TOF	matrix assisted laser desorption ionisation-time of flight
dC	2'-deoxy cytidine	<i>max</i>	maximum
dCTP	2'-deoxy cytidine triphosphate	MeOH	methanol
ddATP	2',3'-dideoxy adenosine triphosphate	(MeO) ₃ PO	trimethylphosphate
ddCTP	2',3'-dideoxy cytidine triphosphate	mg	milligram
ddTTP	2',3'-dideoxy thymidine triphosphate	MHz	megahertz
dG	2'-deoxy guanosine	mM	millimolar
dGTP	2'-deoxy guanosine triphosphate	MST	microscale thermophoresis
DMAP	4-dimethylaminopyridine	ng	nanogram
DMF	dimethylformamide	nm	nanometer
DMSO	N, N-dimethyl sulfoxide	nmol	nanomoles
DMT	dimethoxytrityl	nM	nanomolar
DNA	deoxyribonucleic acid	NMR	nuclear magnetic resonance
dNTP	2'-deoxy nucleoside triphosphate	OD ₆₀₀	optical density at 600 nm
ds	double stranded	ON	oligonucleotides
DTT	dithiothreitol	PAGE	polyacrylamide gel electrophoresis
dT	2'-deoxy thymidine		
dTTP	2'-deoxy thymidine triphosphate		

PCR	polymerase chain reaction
PDB	protein data bank
Pd	palladium
PEX	primer extension reaction
pmol	picomoles
POCl ₃	Phosphorous oxychloride
ppm	parts per million
R _f	retention factor
RNA	Ribonucleic acids
ss	single-stranded
SDS	sodium dodecyl sulphate
SPAAC	strain-promoted alkyne-azide cycloaddition
T	thymine
TdT	terminal deoxynucleotidyl transferase
TEAA	triethylammonium acetate
TEAB	triethylammonium bicarbonate
THF	tetrahydrofuran
TLC	thin layer chromatography
T _m	thermal melting
TMEDA	N,N,N',N'- tetramethylethylenediamine
Tris	tris(hydroxymethyl)amino methane
U	uridine
UTP	uridine triphosphate
UV	ultraviolet
WC	Watson-Crick

Synopsis

Environment-sensitive nucleotide analogs serve as good substrates and probes to study the activity of nucleic acid processing enzymes

Background and Aim

Enzymatic functionalization of nucleic acids with biophysical tools and probes is regularly used in various biotechnological applications.¹⁻⁴ This approach greatly depends on the capability of both native and engineered DNA/RNA polymerases to process unnatural nucleotides, particularly those that are base-modified.⁵⁻⁹ Various types of labeled nucleotides, including those tagged with dyes, environment-sensitive fluorescent probes, as well as NMR- and EPR-active labels, along with nucleotides equipped with handles for post-polymerization modification, have been extensively utilized.¹⁰⁻¹⁵ However, the precise molecular mechanism underlying the processing of these analogs is not yet fully elucidated. In a recent advancement, a platform has been established that integrates molecular biology and X-ray crystallography to provide structural insights into how DNA polymerases accept and incorporate certain modified nucleotides.^{16,17} In this approach, researchers successfully co-crystallized a DNA polymerase with various 2'-deoxynucleoside triphosphate (dNTP) surrogates containing different aliphatic chains and aromatic moiety.^{18,19} These studies offer valuable insights into the structure-activity relationship of DNA polymerases, paving the way for new experimental strategies in the rational design of modified nucleotides for diverse applications. However, most of these experiments are performed under non-equilibrium conditions and seldom provide real-time data.^{20,21}

Fluorescence-based techniques have been extensively employed for investigating nucleic acid structure, dynamics and ligand binding in solution state.^{22,23} In particular, fluorescent nucleobase analogs (FNAs) have emerged as a preferred tool for exploring nucleic acid structure and their interactions with small molecules and proteins. This has been particularly emphasized in the utilization of environment-sensitive fluorescent nucleoside analogs.^{24,25} These analogs, formed by attaching heterocyclic moieties to pyrimidine and purine bases, have proven highly valuable in unraveling nucleic acid conformation, dynamics and recognition properties. In our research efforts, we have pioneered the development of a series of multifunctional ribonucleoside probes with high conformational sensitivity, achieved by attaching heterobicycles at the C5 position of pyrimidine nucleosides. Leveraging the conformation-sensitivity of these probes, we have established assays to investigate various aspects, including nucleic acid lesions, interactions between nucleic acids and drugs, and the

exploration of therapeutically relevant non-canonical nucleic acid structures.^{26–29} These endeavors contribute to a deeper understanding of nucleic acid behaviors and pave the way for applications in diverse research areas. The present thesis focuses on establishing a biochemical and biophysical pipeline by combining fluorescence, ¹⁹F NMR, and X-ray crystallography techniques to understand the incorporation mechanism of template-dependent and template-independent DNA polymerases. This approach elucidates how polymerases incorporate these modified nucleotides into DNA oligomers. Such an initiative not only promises a deeper understanding of the mechanisms of polymerases but also holds the potential to guide the design of nucleotide probes that are compatible for enzymatic functionalization of nucleic acids.

The thesis is divided into Chapters 1–4.

Chapter 1: Nucleotide analogs as substrates and probes for nucleic acid processing enzymes

In this chapter, we provided a concise overview of the utilization of polymerases and transferases for labeling nucleic acids with diverse modifications. Additionally, we explored biochemical techniques and X-ray crystallography methods designed to unravel the intricate processes associated with incorporating modified nucleotides. It is noteworthy that many of these methods are conducted under non-physiological conditions and lack real-time analysis capabilities. To address this gap, various fluorescence techniques have been developed, offering insights into nucleic acid-protein interactions. The utility of fluorescence methods to study these processes and recent efforts to uncover enzymatic incorporation mechanisms by using fluorescence nucleotide analogs are described. Subsequently, the motivation for the present study is elaborated.

Chapter 2: Microenvironment-sensitive fluorescent nucleotide probes from benzofuran, benzothiophene and selenophene as substrates for DNA polymerases

In Chapter 2, our investigation focused on the incorporation of heterocycle-modified nucleotides, specifically **BFdUTP**, **BTdUTP**, and **SedUTP**, by utilizing family A and family B DNA polymerases, namely KlenTaq and KOD/Vent(exo-) (Figure 1A). The efficiency of incorporation of these nucleotides varies based on the size of the heterocycle modification and the DNA polymerase sequence family employed. KlenTaq exhibited sensitivity to the heterocycle size attached to the nucleobase, accepting only one heterobicyclic-modified nucleotide (**BFdUTP/BTdUTP**) into the oligonucleotide (ON) while efficiently incorporating the selenophene-modified nucleotide analog at multiple sites. In contrast, KOD and Vent(exo)

displayed high substrate tolerance, successfully incorporating all three modified nucleotides and generating ONs with multiple labels. Using the flexibility in the tolerance of the modifications by KlenTaq and KOD DNA polymerase we were able to develop a site-specific labeling technique (Figure 1B). Noteworthy is the efficiency of **BFdUTP** as a reporter, providing fluorescence readouts for monitoring enzymatic incorporation and estimating binding events in the reaction cycle. A thorough analysis of the incorporation profile, fluorescence data and the crystal structure of the ternary complex with KlenTaq DNA polymerase, **BFdUTP** and duplex, contributes to a detailed understanding of the incorporation mechanism of heterocycle-modified nucleotides (Figure 1C).

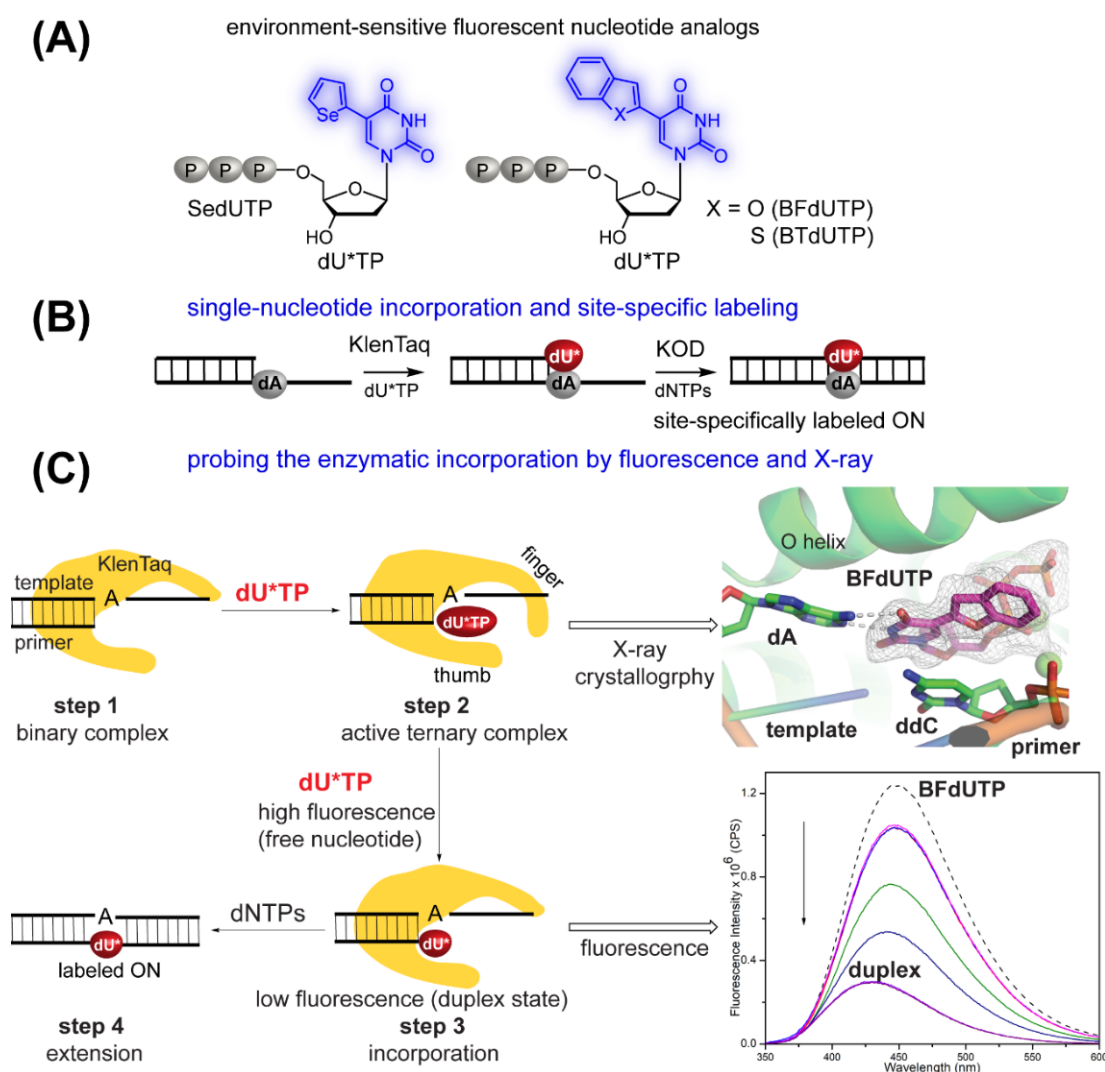


Figure 1. (A) microenvironment-sensitive fluorescent nucleotide analogs. (B) the schematic diagram for the two-step site-specific labeling technique (C) probing the ternary complex formation by X-ray crystallography and enzymatic incorporation using fluorescence spectroscopy.

Chapter 3: Probing the processing of environment-sensitive nucleotide analogs by DNA polymerases using fluorescence, ^{19}F NMR and X-ray crystallography

In Chapter 3, we introduced a nucleotide featuring a fluorobenzofuran moiety at the C5 position of 2'-deoxyuridine (**FBFdUTP**). The nucleotide serves as a dual-responsive probe composed of both responsive fluorescence and ^{19}F NMR labels (Figure 2A). Notably, it displays variations in incorporation efficiency when utilized with KlenTaq and KOD DNA polymerases. The introduction of the fluorine atom in the benzofuran ring does not impair the overall incorporation efficiency by DNA polymerases. This study established a two-channel spectroscopic read-out method to examine the binding event and the enzymatic incorporation process of **FBFdUTP** (Figure 2A, D). By fluorescence spectroscopy, we could detect the differences in the binding efficiency of KlenTaq and KOD DNA polymerase. To gain deeper insights, we conducted a comparative analysis of three crystal structures with KlenTaq DNA polymerase ternary complexes containing **FBFdUTP**, **SedUTP**, and **BFdUTP**, along with the respective primer-template duplex (Figure 2B, 2C). It has been demonstrated that the incorporation efficiency of the KlenTaq DNA polymerases highly depends on the size of the heterocyclic modifications. We observed that the flexibility of Arg660 and Arg587 near the active site and the stacking interaction between incoming nucleotides and the 3'-primer end nucleobase are important factors controlling the incorporation of the nucleotides containing the heterocyclic moiety. The heterobicyclic modification **BFdUTP** and **FBFdUTP** perfectly accommodate in the enzyme pocket and participates in the stacking interaction with 3'-end base of the primer. However, to accommodate this modification in the active site, Arg660 and Arg587 residues rearrange their conformation to make a crevice for the heterocyclic ring. On the other hand, the smaller heterocyclic moiety **SedUTP** successfully contributes to the stacking interaction without disturbing the active site conformation and thus, it has been proven to be a better substrate than the natural analog. Finally, comparing the multiple incorporation profile of modified nucleotides by primer extension reaction using KlenTaq and KOD DNA polymerase lead to us two main differences in the between the two enzymes. The protein channel volume and the conformations of the DNA near the active site. In KOD DNA polymerase the distance between thumb domain and finger domains is more and provide a larger space to accommodate multiple bulkier modifications. Whereas, KlenTaq DNA polymerase has a narrow channel volume and fails to incorporate multiple modified nucleotides. Secondly, the conformation of DNA near the active site of KOD DNA polymerase is B-form, which provides larger major groove for holding multiple C5 modifications. Whereas, A-form of DNA is present near KlenTaq DNA polymerase, which has a narrow major

groove for accommodating multiple C5 modifications. This comparative examination yielded valuable information regarding the mechanistic aspects of incorporating heterocyclic moieties and elucidated the chemical space tolerance exhibited by DNA polymerases.

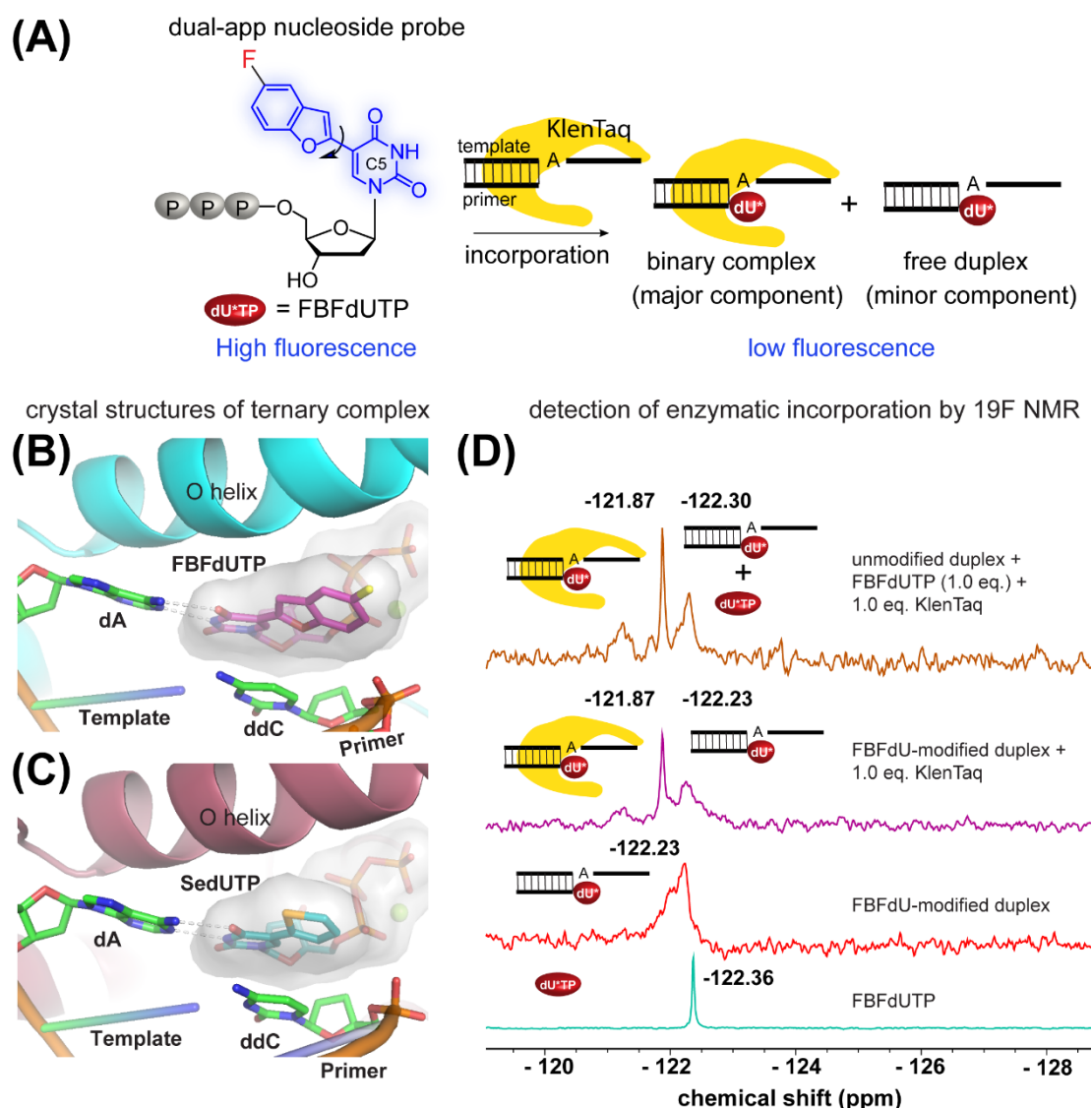


Figure 2. (A) Schematic illustration for incorporation of fluorobenzofuran-modified dual app nucleotide (B) and (C) crystal structure of the ternary complex of the primer-template duplex, KlenTaq and FBFdUTP and SedUTP, respectively. (D) detection of the enzyme activity using ^{19}F NMR spectroscopy.

Chapter 4: Microenvironment-responsive fluorescent nucleotide probes serve as good substrates to investigate the activity of terminal deoxynucleotidyl transferase (TdT)

Terminal deoxynucleotidyl Transferase (TdT) is a unique template-independent DNA polymerase with various biotechnological and immunological applications.^{30,31} TdT overexpression is associated with several leukemia diseases and is used as a biomarker and therapeutic target. Therefore, detecting TdT activity is crucial for developing a screening platform for finding suitable inhibitors for the TdT enzyme.^{32,33} Traditional detection

techniques are time-consuming, expensive and cannot provide real-time data. Recently, fluorescent nucleobase probes have shown promise for directly analyzing TdT activity.³⁴ Hence, our idea is to establish a simple and efficient platform to understand the mechanism of TdT and develop a real-time fluorescence-based detection technique to monitor the TdT activity using environment-sensitive nucleotide analogs. Specifically, we conducted an incorporation assay of C5-heterocycle-modified analogs by TdT, resulting mainly in double-modification products. Larger heterocyclic modifications like **BFdUTP**, **BTdUTP**, and **FBFfUTP** show increased efficiency in distributive incorporation compared to smaller modifications like **SedUTP** or natural thymidine analogs. We proposed that stacking interaction between incoming nucleotide and the primer's terminal base during the pre-catalytic step contributes to variations in TdT activity. To explore this, we utilized benzofuran-modified-2',3'-dideoxyuridine triphosphate (**BFddUTP**) and 2',3'-dideoxy thymidine triphosphate (ddTTP) and compared their incorporation by TdT DNA polymerase. The modified analogs exhibited superior substrate efficiency for enzymatic incorporation by TdT, primarily due to stacking interactions between incoming nucleotide and the 3'-termini of the primer strand. Next, we aimed to establish a simple and efficient detection method to monitor TdT activity. By using the inherent responsiveness of the environment-sensitive nucleotide analog **BFddUTP**, we designed a fluorescence-based detection method to detect the TdT activity. And monitor the competition between **BFddUTP** and ddTTP. This will help develop a screening platform for investigating novel TdT inhibitors (Figure 3A, B).

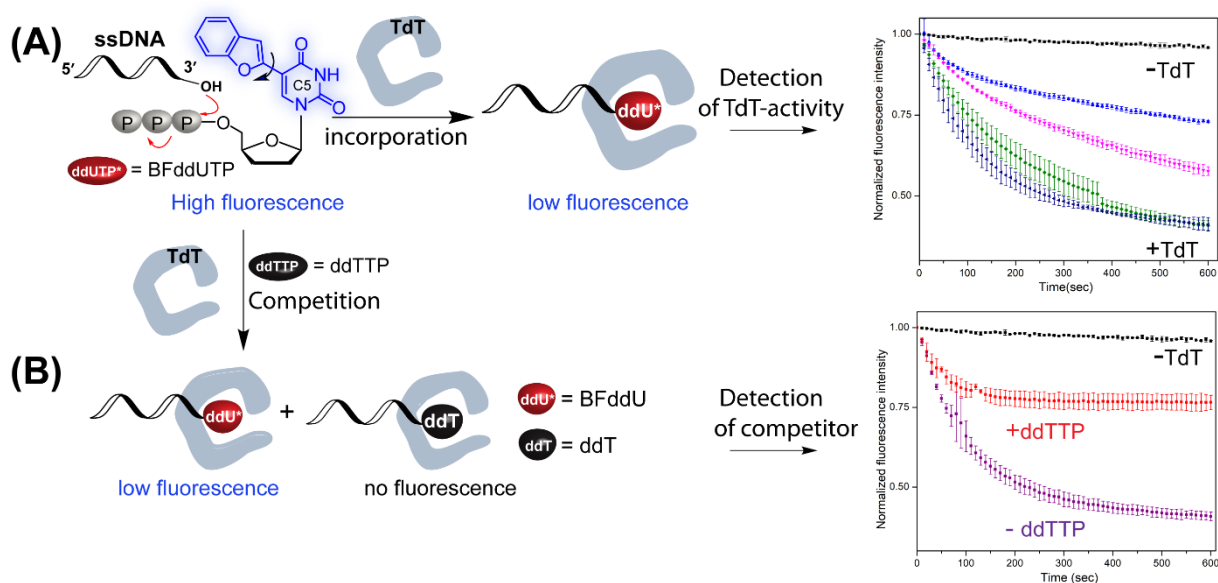


Figure 3. (A) Schematic illustration of incorporating BFddUTP by real-time fluorescence. (B) Detection of the Competition assay by fluorescence.

Taken together, we report the incorporation of a series of C5-heterocycle-modified nucleotide analogs by DNA polymerases and a terminal nucleotide transferase (TdT). These modifications are amenable to enzymatic incorporation by template-dependent (family A and B) and template-independent (TdT) DNA polymerases. The nucleotides themselves act as probes to report the catalytic events in equilibrium conditions by fluorescence and ^{19}F NMR spectroscopy. The results indicate that the tested enzymes could be used to generate probe-functionalized DNA oligonucleotides for various applications. Importantly, by combining solution state and crystallography data, we gained valuable insights into the enzymatic incorporation mechanism, which should greatly aid in designing newer heterocycle-modified nucleotide probes compatible with enzymatic labeling for nucleic acids.

References

1. Chen, T.; Hongdilokkul, N.; Liu, Z.; Thirunavukarasu, D.; Romesberg, F. E. The expanding world of DNA and RNA. *Curr. Opin. Chem. Biol.* **2016**, *34*, 80–87.
2. Aschenbrenner, J.; Marx, A. DNA polymerases and biotechnological applications. *Curr. Opin. Biotechnol.* **2017**, *48*, 187–195.
3. Hocek, M. Enzymatic synthesis of base-functionalized nucleic acids for sensing, cross-linking, and modulation of protein–DNA binding and transcription. *Acc. Chem. Res.* **2019**, *52*, 1730–1737.
4. Ashley, J.; Potts, I. G.; Olorunniji, F. J. Applications of Terminal Deoxynucleotidyl Transferase Enzyme in Biotechnology. *ChemBioChem.* **2023**, *24*, e2022005.
5. Ramsay, N.; Jemth, A.-S.; Brown, A.; Crampton, N.; Dear, P.; Holliger, P. CyDNA: synthesis and replication of highly Cy-dye substituted DNA by an evolved polymerase. *J. Am. Chem. Soc.* **2010**, *132*, 5096–5104.
6. Hollenstein, M. Nucleoside triphosphates-building blocks for the modification of nucleic acids. *Molecules.* **2012**, *17*, 13569.
7. Hocek, M. Synthesis of base-modified 2'-deoxyribonucleoside triphosphates and their use in enzymatic synthesis of modified DNA for applications in bioanalysis and chemical Biology. *J. Org. Chem.* **2014**, *79*, 9914–9921.
8. Winz, M. L.; Linder, E. C.; André, T.; Becker, J.; Jäschke, A. Nucleotidyl transferase assisted DNA labeling with different click chemistries. *Nucleic Acids Res.* **2015**, *43*, No. e110.
9. Klöcker, N.; Weissenboeck, F. P.; Rentmeister, A. Covalent labeling of nucleic acids. *Chem. Soc. Rev.* **2020**, *49*, 8749–8773.
10. Obeid, S.; Yulikov, M.; Jeschke, G.; Marx, A. Enzymatic synthesis of multiple spin-labeled DNA. *Angew. Chem. Int. Ed.* **2008**, *47*, 6782–6785.
11. Jarchow-Choy, S. K.; Krueger, A. T.; Liu, H.; Gao, J.; Kool, E.T. Fluorescent xDNA nucleotides as efficient substrates for a template-independent polymerase. *Nucleic Acids Res.* **2011**, *39*, 1586–1594.
12. Dziuba, D.; Jurkiewicz, P.; Cebecauer, M.; Hof, M.; Hocek, M. A rotational BODIPY nucleotide: an environment-sensitive fluorescence-lifetime probe for DNA interactions and applications in live cell microscopy. *Angew. Chem., Int. Ed.* **2016**, *55*, 174–178.
13. Manna, S.; Srivatsan, S. G. Synthesis and enzymatic incorporation of a responsive ribonucleoside probe that enables quantitative detection of metallo-base pairs. *Org. Lett.* **2019**, *21*, 4646–4650.

14. Samanta, A.; Krause, A.; Jäschke, A. A modified dinucleotide for site-specific RNA-labelling by transcription priming and click chemistry. *Chem. Commun.* **2014**, 50, 1313–1316.
15. Reisacher, U.; Ploschik, D.; Röncke, F.; Cserép, G. B.; Kele, P.; Wagenknecht, H.-A. Copper-free dual labeling of DNA by triazines and cyclopropenes as minimal orthogonal and bioorthogonal functions. *Chem. Sci.* **2019**, 10, 4032–4037.
16. Hottin, A.; Marx, A. Structural insights into the processing of nucleobase-modified nucleotides by DNA polymerases. *Acc. Chem. Res.* **2016**, 49, 418–427.
17. Vyas, R.; Zahurancik, W. J.; Suo, Z. Structural basis for the binding and incorporation of nucleotide analogs with L-stereo chemistry by human DNA polymerase λ . *Proc. Natl. Acad. Sci. U. S. A.* **2014**, 111, E3033–E3042.
18. Bergen, K.; Steck, A. L.; Strutt, S.; Baccaro, A.; Welte, W.; Diederichs, K.; Marx, A. Structures of KlenTaq DNA polymerase caught while incorporating C5-modified pyrimidine and C7-modified 7-deazapurine nucleoside triphosphates. *J. Am. Chem. Soc.* **2012**, 134, 11840–11843.
19. Obeid, S.; Bußkamp, H.; Welte, W.; Diederichs, K.; Marx, A. Interactions of non-polar and “Click-able” nucleotides in the confines of a DNA polymerase active site. *Chem. Commun.* **2012**, 48, 8320–8322.
20. Pagano, J.M.; Clingman, C.C.; Ryder, S. P. Quantitative approaches to monitor protein-nucleic acid interactions using fluorescent probes. *RNA.* **2011**, 17, 14–20.
21. Su, X.; Xiao, X.; Zhang, C.; Zhao, M; Nucleic acid fluorescent probes for biological sensing *Appl. Spectrosc.* **2012**, 66,1249–1261.
22. Brown, M. P.; Royer, C. Fluorescence spectroscopy as a tool to investigate protein interactions. *Curr. Opin. Biotechnol.* **1997**, 8, 45–49.
23. Sinkeldam, R. W.; Greco, N. J.; Tor, Y. Fluorescent analogs of bio molecular building blocks: design, properties, and applications. *Chem.Rev.* **2010**, 110, 2579–2619.
24. Sholokh, M.; Sharma, R.; Shin, D.; Das, R.; Zaporozhets, O. A.; Tor, Y.; Mély, Y. Conquering 2-aminopurine’s deficiencies: highly emissive isomorphous guanosine surrogate faithfully monitors guanosine conformation and dynamics in DNA. *J. Am. Chem. Soc.* **2015**, 137, 3185–3188.
25. Dziuba D. Environmentally sensitive fluorescent nucleoside analogues as probes for nucleic acid - protein interactions: molecular design and biosensing applications. *Methods Appl Fluoresc.* **2022**, 10, 044001.
26. Tanpure, A. A.; Srivatsan, S. G. A microenvironment-sensitive fluorescent pyrimidine ribonucleoside analogue: synthesis, enzymatic incorporation, and fluorescence detection of a DNA abasic site. *Chem. – Eur. J.* **2011**, 17, 12820–12827.
27. Manna, S.; Sarkar, D.; Srivatsan, S. G. A dual-app nucleoside probe provides structural insights into the human telomeric overhang in live cells. *J. Am. Chem. Soc.* **2018**, 140, 12622–12633.
28. Nuthanakanti, A.; Ahmed, I.; Khatik, S. Y.; Saikrishnan, K.; Srivatsan, S. G. Probing G-quadruplex topologies and recognition concurrently in real time and 3D using a dual-app nucleoside probe. *Nucleic Acids Res.* **2019**, 47, 6059–6072.
29. Tanpure, A. A.; Srivatsan, S. G. Conformation-sensitive nucleoside analogues as topology-specific fluorescence turn-on probes for DNA and RNA G-quadruplexes. *Nucleic Acids Res.* **2015**, 43, e149.
30. Fowler, J. D.; Suo, Z. Biochemical, structural, and physiological characterization of Terminal deoxynucleotidyl transferase. *Chem. Rev.* **2006**, 106, 2092–2110.
31. Ashley, J.; Potts, I. G.; Olorunniji, F. J. Applications of terminal deoxynucleotidyl transferase enzyme in biotechnology. *ChemBioChem*, **2023**, 24, e2022005.

32. Hoffbrand, A. V.; Ganeshaguru, K.; Janossy, G.; Greaves, M. F.; Catovsky, D.; Woodruff, R. K. Terminal deoxynucleotidyl-transferase levels and membrane phenotypes in diagnosis of acute leukaemia. *Lancet* **1977**, 310, 520–523.
33. Kung, P. C.; Long, J. C.; McCaffrey, R. P.; Ratliff, R. L.; Harrison, T. A.; Baltimore, D. Terminal deoxynucleotidyl transferase in the diagnosis of leukemia and malignant lymphoma. *Am. J. Med.* **1978**, 64, 788–794.
34. Pals, M. J.; Velema, W. A. Aquenched size-expanded nucleotide reports activity of the leukemia biomarker Terminal deoxynucleotidyl transferase (TdT) *Angew. Chem. Int. Ed.* **2023**, 62, e202302796.

List of publications

Based on the work presented in this thesis

1. **Ghosh, P.**; Kropp, H. M.; Betz, K.; Ludmann, S.; Diederichs, K.; Marx, A.; Srivatsan, S. G. Microenvironment-sensitive fluorescent nucleotide probes from benzofuran, benzothiophene, and selenophene as substrates for DNA polymerases. *J. Am. Chem. Soc.* **2022**, *144*, 10556–10569.
2. **Ghosh, P.**; Betz, K.; Gutfreund, C.; Marx, A.; Srivatsan, S. G. Probing the processing of environment-sensitive nucleotide analogs by DNA polymerases using fluorescence, ¹⁹F NMR and X-ray crystallography (*manuscript under preparation*).
3. **Ghosh, P.**; Bindu, B.; Palani, S.; Srivatsan, S. G. Microenvironment-responsive fluorescent nucleotide probes serve as good substrates to investigate the activity of terminal deoxynucleotidyl transferase (TdT) (*manuscript under preparation*).

Other publication during MS-PhD

1. Malla, J. A.; Upadhyay, A.; **Ghosh, P.**; Mondal, D.; Mondal, A.; Sharma, S.; Talukdar, P. Chloride Transport across Liposomes and Cells by Nontoxic 3-(1H-1,2,3-Triazol-1-yl) benzamides. *Org. Lett.* **2022**, *24*, 4124–4128.

Chapter 1

Nucleotide analogs as substrates and probes for nucleic acid processing enzymes

1.1 Introduction

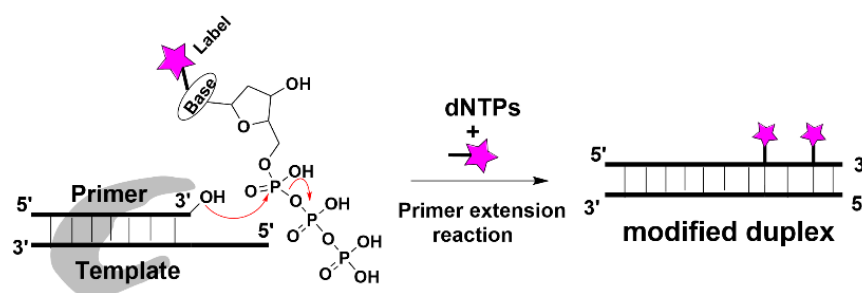
Nucleic acids serve indispensable roles in the storage and transfer of genetic information, the regulation of gene expression and the catalysis of chemical reactions through their interactions with nucleic acids, proteins and small molecule metabolites.¹⁻⁸ As the building block of nucleic acids, the nucleotides lack inherent labels, making it imperative to introduce extrinsic modifications for investigating their structure, dynamics, and interactions with biomolecules. Labeling nucleic acids with modified nucleotides becomes crucial to expand their functionality from fundamental research to various biotechnological applications, diagnostics tools, and biosensing platforms.⁹⁻¹⁴ Functionalization of nucleic acid can be achieved by a chemical method using solid-phase oligonucleotide (ON) synthesis or enzymatically using different DNA and RNA polymerases and transferases (Figure 1).¹⁵⁻²¹ In solid phase synthesis, protected nucleoside phosphoramidite building blocks are added stepwise to the 5'-end. The solid-phase synthesis method has been employed for several decades due to its robustness, scalability, and flexibility in accepting different modifications in specific positions of nucleic acids.^{15,16} However, the harsh chemical conditions for certain modifications and the challenges in synthesizing lengthy oligomers constrain the utility of solid-phase synthesis.¹⁷⁻¹⁹ In contrast, the enzymatic synthesis method involves using modified nucleotides as substrates for DNA and RNA polymerases, leading to their incorporation into the nascent DNA or RNA strands by natural enzymatic machinery. This method offers the advantage of synthesis of long oligomers and is performed under physiological conditions, which is suitable for sensitive modifications.¹⁷⁻²¹ However, nucleic acid polymerases have evolved to incorporate natural nucleosides.²²⁻²⁴ Therefore, the modified nucleotides need to fulfill certain criteria to be incorporated by these polymerases. Nevertheless, a large number of modified nucleotide analogs have been incorporated by enzymatic methods for various biotechnological applications. For example, sterically demanding modifications (such as fluorescent dyes and affinity tags) attached by long flexible linkers to the nucleobases are well accepted by the DNA polymerases, and this strategy is widely employed for designing fluorescent reversible

terminators for next-generation sequencing.⁹ Even the purine base bearing small modifications at the C8 positions²⁵ and a wide range of C5-modified pyrimidine²⁶ analogs have been incorporated by enzymatic reaction. In recent studies, it has been observed that the polymerases do not well tolerate nucleotides bearing the modifications at the phosphate chain and sugar moiety^{27,28}, as they project toward the minor groove of the DNA helix and perturb the DNA-protein interactions. However, the modifications at the nucleobases are comparatively well accepted and that highly depends on the position of the modifications in the nucleobase. Favorable modification positions were observed at the C5/C6-site of pyrimidine and the C8/C7 positions of (deaza-)purine bases, known as the “Hoogsteen face”.^{17,29,30} These modifications were preferred as they do not disturb the Watson-Crick base pairing essential for complementary base interactions. In the context of the double-stranded helix, modifications at these positions orient away from the helical core, extending toward the major groove where steric constraints are relatively less stringent compared to the minor groove.²⁹⁻³³ Despite identifying suitable modification sites, the underlying reasons for accepting these modified nucleotides by high-fidelity DNA polymerases remained elusive for an extended time. However, gaining insights into the mechanistic aspects of incorporating these nucleotides would allow researchers to design polymerase-compatible modifications and predict the likelihood of successful incorporation.

Several biochemical and biophysical techniques including X-ray crystallography have been developed to understand these complex processes. These methods offer structural insights into the incorporation mechanism of modified nucleotides, aiding in the design of novel nucleotides.³⁴⁻³⁶ However, most of these methods are performed under non-physiological conditions and incapable of providing real-time analysis. Various fluorescence techniques have been developed throughout the years to fulfill this need. Several commercially available fluorescent probes have been commonly linked covalently to the nucleobase, phosphate backbone, and sugar moiety, representing well-established methods. Incorporating such probes into nucleic acids has found extensive utility in studies involving FRET, fluorescence anisotropy, and single-molecule fluorescence experiments, revealing nucleic acid-protein interactions.³⁷⁻⁴⁰ However, a notable drawback of these probe systems is their tendency to disrupt the native structures significantly upon binding or incorporation into nucleic acid strands. Consequently, the applications of these probes are confined to monitoring the overall conformational changes of the nucleic acid during interactions with proteins.⁴¹ The conformational alterations occurring during this catalytic cycle encompass changes at both the global and nucleotide levels. Alterations in conformations of nucleotides involve alternation in

their local environment and their interactions with neighboring bases.^{37,38,41} Hence, incorporating a conformation-responsive and minimally perturbing fluorescent reporter at the site of interest and investigating the conformational changes would provide a better understanding of these incorporation processes.⁴²⁻⁴⁴ In this context, several such fluorescent nucleobase analogs have been developed and utilized to study nucleic acid-protein interactions in real-time by monitoring the changes in their photophysical properties such as quantum yield, fluorescence intensity, emission maximum and anisotropy.^{44,45} In this chapter, a detailed discussion on the use of different polymerases and transferases to functionalize nucleic acids is provided. This chapter also discussed the methods that have been developed to understand the polymerisation process and nucleic acid-protein interactions. Finally, the motivation for the present thesis work and how we have unraveled the enzymatic incorporation mechanism using environment-sensitive fluorescence nucleotide analog probes are described.

(A) Incorporation by template-dependent DNA polymerases



(B) Incorporation by template-independent DNA polymerases

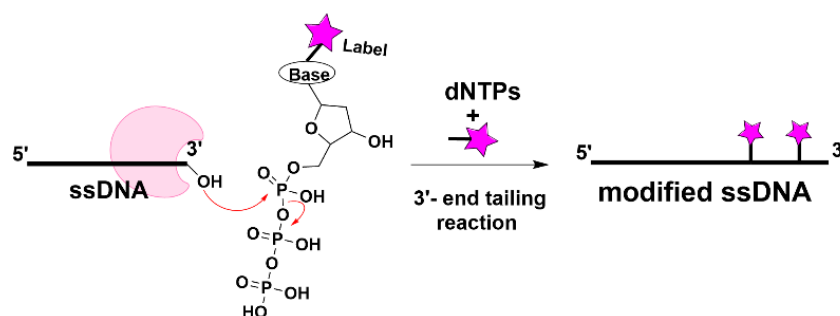


Figure 1. (A) Template-dependent enzymatic incorporation and (B) template-independent enzymatic incorporation of modified nucleotides.

1.2 Enzymatic incorporation of different modified nucleotides and their applications

Polymerases have evolved to accept canonical nucleotides, but they can also surprisingly accommodate a diverse array of modified nucleotides. This expands the capabilities of nucleic acids and allows for their use in various biological applications. Typically, all components of

natural nucleotides, including the ribose ring, phosphate backbone, and nucleobases, can be modified as needed.

1.2.1 Sugar-modified nucleotides

Recent studies of enzymatic labeling have shown that modifications at the sugar ring are the least preferred substrates for the enzymes. Modifications in the ribose ring point to the minor groove and additionally, polymerases typically rely on effective triphosphate orientation facilitated by interaction with the 3'-OH, making modifications at this position less readily incorporated by natural polymerases. Despite this difficulty of incorporation, several base-modified nucleotides have been incorporated by DNA polymerases. For instance, 2'-fluoro and 2'-OMe-modified nucleotides have been incorporated by DNA polymerases, which provide the stability of the oligonucleotides under biological conditions.⁴⁶ Numerous engineered DNA polymerases have demonstrated the ability to synthesize a series of sugar-modified oligonucleotides, such as 2'-fluoroarabinonucleic acids (FANA)⁴⁷, LNA⁴⁸, 7', 5'-bicyclo-DNA⁵⁰ and (3'-2') α -L-threose nucleic acid (TNA)⁵⁰. Renders et. al reported the incorporation of 2'-phosphomethyl 5'-3'-phosphonomethyl-dNTPs.⁵¹ Even 2'-deoxy-2'-isonucleoside (iNTPs)⁵², 3'-deoxyribonucleotide 3'-triphosphates (apio NTPs)⁵⁴ have been utilized as substrates for the Terminator DNA polymerases.

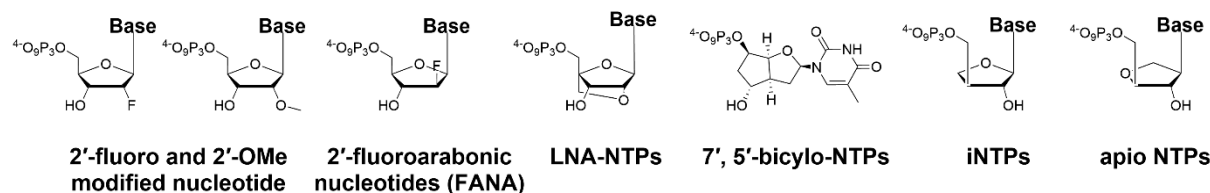


Figure 2. Examples of several sugar-modified nucleotides used in enzymatic incorporation.

1.2.2 Nucleobase-modified nucleotides

1.2.2.1 Incorporation of fluorescent tags

In 2001, Angerer and colleagues reported the incorporation of a complete set of modified 2'-deoxy uridine triphosphates composed of a coumarin/fluorescein residue at the C5 position using various DNA polymerases. They were successful in synthesizing modified DNA strands up to 40 base pairs in length using primer extension reactions (PEX) and even obtained full-length products for some of the modifications in PCR reactions. Polymerases from family B (e.g., Vent exo⁻ and Tgo exo⁻) incorporate labeled dNTPs more efficiently than Taq polymerase belonging to family A DNA polymerase.⁵⁴ Although fluorophores are usually linked to nucleobases using a long alkyl chain an amide⁵⁴ and a polyethylene glycol linker,⁵⁵ recent studies have shown that a phenylethyne linker is more efficient in terms of incorporation

compared to a small linker. They used TaqFS DNA polymerase to incorporate various fluorescently labeled dTTP and ddTTP analogs, where the fluorescein group is connected to the C5 position via different lengths of a conjugated linker. The results showed that the modified nucleotides with longer, conjugated linkers have better extinction coefficients, and their absorption maximum is shifted towards a higher wavelength. The researchers hypothesized that the linker serves as a donor, which transfers more energy to the fluorophore during absorption. This leads to an improvement in the fluorescence intensity of the labeled nucleic acids.⁵⁶ The Hocek group has reported the incorporation of fluorescent molecular rotor-based environment-sensitive fluorescent nucleotide analogs by various DNA polymerases (see section 1.5.2). Certain cytidine analogs that have been modified with BODIPY, such as **dC^{bdp}TP** and **dC^{Tbdp}TP**, have been successfully incorporated using KOD DNA polymerase.⁵⁷ Additionally, dCTP analogs with fluorene attachments (**dC^{FL}TP**)⁵⁸ and tryptophan-based fluorophores (**dC^{Trp}TP**)⁵⁹ have been used as substrates for KODXL and Bst DNA polymerases. These analogs were not only accepted by primer extension reactions but also successfully incorporated by PCR reactions.

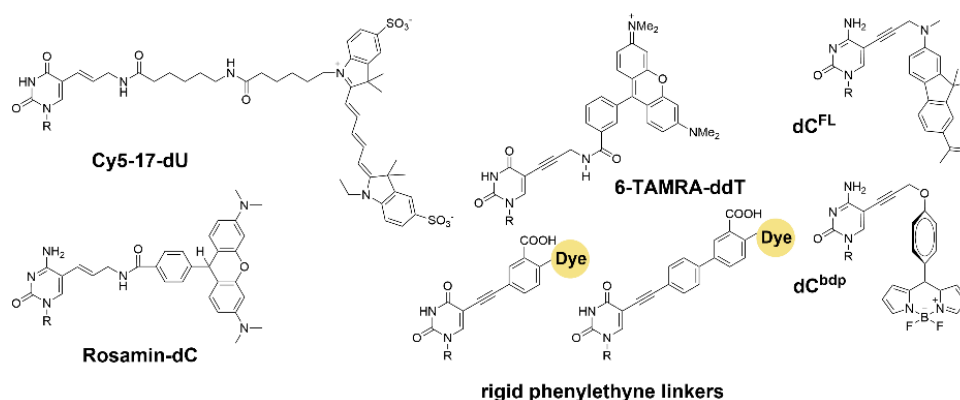


Figure 3. Examples of different fluorescently-labeled nucleotides. R= 2'- deoxyribose sugar.

1.2.2.2 Incorporation of post-synthetic modifications

Incorporating fluorescence reporter molecules into oligonucleotides using direct enzymatic labeling can be difficult because the modification is relatively large and not easily accepted by polymerases. Therefore, it is advantageous to incorporate small reactive handles that are more readily accepted by DNA polymerases, as they can facilitate the installation of larger fluorophore moieties for click reactions.⁶⁰ Several such post-synthetic reactive handles have been developed throughout the years (Figure 4). Jäschke and team investigated incorporating C5 norbornene-modified pyrimidine and deaza-purine-nucleotides into DNA using polymerases. They demonstrated the incorporation of **dUTP^{Norb}** by Taq DNA polymerase.

Whereas, Phusion protein accepted all three nucleotides efficiently in primer extension reactions but failed in the PCR reactions. KOD-XL DNA polymerase successfully incorporates multiple norbornene moieties in PCR reactions.⁶¹ Using PEX reactions, several C2-modified dATP analogs were introduced into the DNA strand, such as **d^{vinyl}ATP** and **d^{ethynyl}ATP** (C2 modified with vinyl and ethynyl, respectively, Figure 4). Followed by CUAAC and thiol-ene reactions were performed for attaching fluorescent tags.⁶² Wagenknecht and his team have developed a new method to label DNA with 1-methylcyclopropene (1MCP)-modified dATP analog (**1MCP-dzATP**) by employing a DNA polymerase-based approach. After that, a tetrazine-bearing fluorescent dye was inserted through click chemistry to generate fluorescently labeled DNA. Moreover, they transfected HeLa cells with the modified DNA oligonucleotides, and a click reaction was performed under cellular conditions to visualize the modified DNA inside living cells by confocal microscopy.⁶³ Recently, several flexible alkyne-linked thymidine triphosphates such as 5-(1,7-octadiynyl)-dUTP (**dU^oTP**), PEG-alkyne modified (**dU^{peg}TP**), O-undecynyl modified (**dU^{un}TP**), were incorporated into the DNA by PEX and PCR reactions by using KOD XL and Bst DNA polymerase.⁶⁴

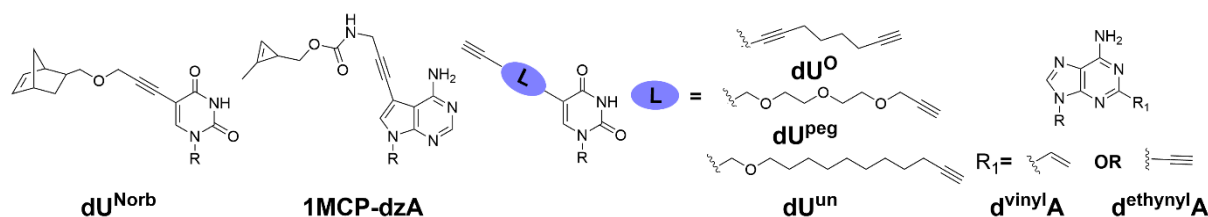


Figure 4. Different post-synthetic modifications. R = 2'-deoxyribose sugar.

1.2.2.3 Metabolic labeling and cell imaging

Metabolic functionalization of nucleic acids is a highly effective technique that involves the introduction of modified nucleotides into newly synthesized RNA or DNA during cellular processes. This approach is widely used in the fields of biochemistry and molecular biology for various applications, including the study of nucleic acid synthesis, transcription, translation, and nucleic acid-protein interactions. Metabolic labeling uses cell-permeable nucleotide precursors such as nucleosides or nucleobase that enter the salvage pathway and are processed to (deoxy) nucleoside triphosphates (dNTPs) which can be further processed by polymerases. Hence, small modifications are preferred as they don't interfere with the metabolism. One such modification is the 5-ethynyl-2'-deoxyuridine (**EdU**), which is a popular nucleoside used for DNA labeling. This modification is efficiently taken up by cells and metabolized to the corresponding nucleotide, which can be accepted by the polymerases.⁶⁵ Azido-modified

fluorophores can be attached to **EdU** by a click reaction for fixed-cell imaging. Stimulated Raman scattering has been used to image **EdU** directly in cells using alkyne vibration. However, this compound can perturb DNA functions and stability at high concentrations.⁶⁶⁻⁶⁸ An alternative modification, the arabinofuranosyl derivative (**F-ara-EdU**), is comparatively less toxic and has been used for metabolic incorporation in cells and zebrafish⁶⁹ Similar to **EdU**, other analogs such as cytidine analog, **EdC**,⁷⁰ and purine analog (**EdA**)⁷¹ are also used for metabolic labeling (Figure 5). Azide-modified nucleotide analogs, such as 5-(azidomethyl)-2'-deoxyuridine (**AmdU**), is used in metabolic labeling to evaluate copper-free click reactions known as SPAAC reactions. **AmdU** demonstrates high intracellular stability for metabolic labeling of DNA using SPAAC reactions. It allows for robust three-color staining of nuclei in fixed HeLa cells when combined with **F-ara-EdU** and **BrdU** (Figure 5).⁷² Similarly, 2'-azide modified arabinose analog of cytosine (**AzC**) has been utilized as metabolic labeling and even this molecule acts as an anticancer prodrug.⁷³ Leudke et al. reported the incorporation of 5-vinyl-2'-deoxyuridine (**VdU**) by endogenous enzymes. The **VdU** can be visualized after the IEDDA reaction with a fluorescent tetrazine.⁷⁴ For cellular DNA synthesis, a new nucleoside known as **VTdT** was introduced, which underwent a hetero-Diels-Alder reaction with an o-quinoline quinone methide to enable DNA labeling.⁷⁵

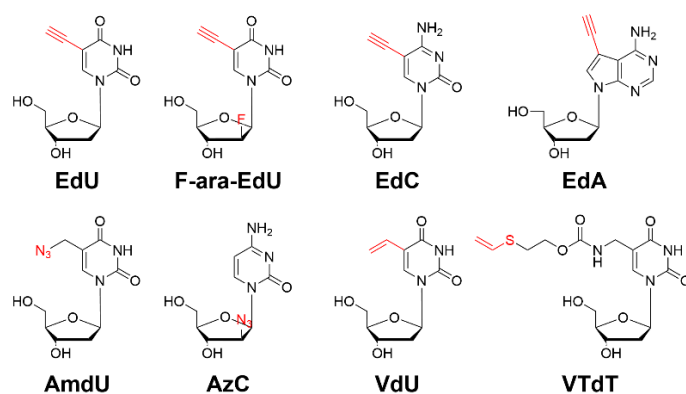


Figure 5. Examples of 2'-deoxynucleosides successfully used for metabolic labeling.

1.2.2.4 Next-Generation sequencing

DNA polymerases are crucial in DNA sequencing. The Sanger DNA sequencing method was first used to sequence the human genome in 2001 and has since become a standard and widely used method to determine DNA sequence (Figure 6).^{76,77} With advancements, new next-generation sequencing methods (NGS) have increased output and lowered costs. Reversible chain terminators and removable fluorophores have advanced sequencing chemistry, leading

to faster, high-throughput NGS methods.^{78,79} The Illumina sequencing method employs nucleotide substrates modified with 3'-azido methylene to terminate DNA synthesis. This method enables fluorescent readout after each nucleotide incorporation. Followed by the removal of blocking groups by chemical methods, and the cycle will continue.⁸⁰ Engineered DNA polymerases also have been used to incorporate various types of reversible terminators. For instance, mutated Taq polymerases have incorporated 3'-O-amino-protected fluorescent nucleotides.⁸¹ A novel category of nucleotides with unblocked 3'-OH, termed "Lightning Terminators™," has been reported. These terminators feature a terminating 2-nitrobenzyl group attached to hydroxymethylated nucleobases. The molecular tuning effect of this photocleavable group enhances single-base termination, leading to improved nucleotide fidelity. In comparison to other reversible terminators like 3'-O-amino-TTP and 3'-O-azidomethyl-TTP, these new terminators exhibit superior enzymatic performance by Terminator™ DNA polymerase.⁸²

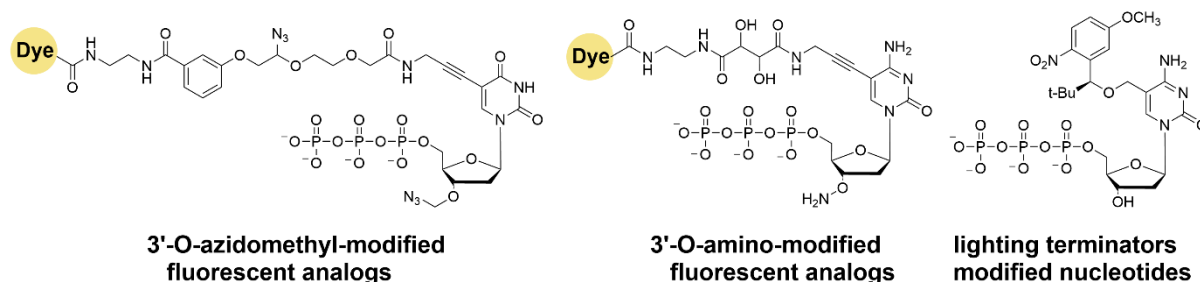


Figure 6. Structure schematic for reversible terminator used in next generation sequencing (NGS).

1.3 Template-independent labeling of nucleic acids

Unlike typical template-dependent DNA polymerase reaction, certain enzyme incorporates modified functionalities in a template-independent manner. Terminal deoxynucleotidyl transferase (TdT), which belongs to the X family, exhibits a distinctive ability to incorporate modified nucleotides at the 3'-end of single-stranded DNA in a template-independent fashion.⁸³ This enzyme plays a crucial role in vertebrate immune system development by introducing random nucleotides to single-stranded DNA at 3'-termini during V(D)J recombination, thereby enhancing T-cell receptor diversity. The biotechnological applications of TdT have expanded in the past two decades due to its remarkable substrate promiscuity.⁸⁵ In 2012 and 2015, Jäschke and colleagues demonstrated TdT's capability to incorporate various click chemistries into DNA and RNA, enabling the conjugation of diverse functional groups (such as biotin, fluorescent dyes, etc.) through the CUAAC reaction.^{86,87} In 2017, researchers utilized end-

labeled DNA with a histag mimic for immobilizing DNA.⁸⁸ Scientists have been able to modify DNA using different types of nucleotide analogs, including nucleotides that are tagged with fluorescent markers and nucleotides that are chemically altered to resist nuclease degradation.⁸⁹ These modified nucleotides are effective in labeling the ends of DNA strands using the enzyme TdT. Furthermore, redox-labeled nucleotides have been successfully incorporated into DNA and RNA oligonucleotides at the 3'-end.^{90,91}

The development of enzymatic methods for *de novo* synthesis, especially using polymerase chemistry such as TdT, provides a green chemistry methodology for industry.⁹² In 2018, Palluk and colleagues reported a method for synthesizing 10-20 nucleotides long single-stranded DNA in a sequence-specific manner. The researchers used TdT polymerase-nucleotide conjugates to add nucleotides step by step. They attached each TdT polymerase molecule to a single dNTP using a cleavable bond, which was then added to a primer. Once the dNTP was incorporated, the 3'-end of the primer remained covalently attached to TdT and could not be accessed by other TdT-dNTP molecules. To continue the extension process, the bond between TdT and the incorporated nucleotide had to be cleaved, which released the primer.⁹³ Matthews and coworkers presented 3'-O-modified dNTPs equipped with photolabile protecting groups, allowing selective removal after the addition of each nucleotide. This method provides precise control over the sequence-specific synthesis of oligonucleotides.⁹⁴ Similarly, Flamme et al. reported the use of nucleotides with benzoyl and pivaloyl protecting groups for sequence-sequence controlled synthesis of DNA and XNA.⁹⁵ In another example, a borate complex with 2' and 3'-hydroxyl groups of nucleotides were utilized as a reversible terminator for controlled oligo synthesis. The TdT DNA polymerase well tolerates this modification.⁹⁶ In 2021, Flamme et al. detailed the synthesis of Locked Nucleic Acid (LNA) using TdT. They employed various nuclease-resistant LNA-dTTP analogs as substrates for end labeling of DNA and observed efficient incorporation by the TdT enzyme for both LNA-TTP and α -thio-dTTP.⁹⁷

1.4 Numerous biophysical techniques and chemical probes to reveal the nucleic acid-protein interactions

1.4.1 X-ray crystallography

X-ray crystallography is an essential tool for deciphering the biological functions of nucleic acid-protein interactions. It provides high-resolution structural insights into individual complexes and allows us a better understanding of the biological processes at the atomic level. It has been very helpful in studying the mechanism of replication, transcription, translation

regulation and nucleic acids-protein interactions. In particular, Crystal structures of active ternary complex (duplex-enzyme-modified nucleotide) provide insights into the incorporation mechanism of certain modified nucleotides. Consequently, studying the incorporation of modified nucleotides by processing enzymes or the interaction of modified oligonucleotides with the nucleic acid binding proteins expands the scope for designing modified nucleotides for myriad biotechnological applications and nucleic acid therapeutics, as we discussed in the previous section.^{30,98}

In this context, Marx et al. conducted a study on the acceptance of certain C5-modified pyrimidine analogs and C7-modified purine analogs by KlenTaq DNA polymerase, which is a family A DNA polymerase. They used an incorporation assay and X-ray crystal structures analysis to evaluate the acceptance of these analogs. The study started by examining the effect of rigid nonpolar modifications by crystallizing KlenTaq that was bound to the spin-labeled nucleotide **dT^{spin}TP** and primer-template duplex.⁹⁹ Although **dT^{spin}TP** showed low incorporation efficiency compared to natural dTTP, it caused only minor disturbances in the overall confirmation. The study found that enzymes can effectively recognize the modification through usual Watson-Crick base pairing and pi-stacking interactions with the primer. The O-helix is appropriately aligned in a closed state similar to the natural one. However, certain significant differences were observed, such as the positioning of the Arg660 residue in the active site. In the case of modification, it moves away to encompass the nitroxide modification, whereas, in natural dTTP, it makes hydrogen bonding with the phosphate backbone of the primer strand.⁹⁹ Similarly, they have studied the acceptance of flexible bulky modification, **dT^{dend}TP**, where the flexible polyamide branches stay outside of the active site and do not hamper the usual nucleotide binding interactions.¹⁰⁰ It has been observed that Arg660 residue is slightly displaced but maintains its interactions with the primer 3'-end and the amide moiety of the propargylamide linker. Further, they have solved the crystal structures of KlenTaq polymerase in complex with rigid nonpolar aromatic modifications, **dC^{aryl}TP** and **dT^{aryl}TP**.¹⁰¹ They observed similar results as described earlier: Arg660 makes room for the aromatic moiety by losing contact with the primer, and additionally, Arg587 and Lys663 stabilized the modification by cation-pi interaction. These interactions are reflected in the incorporation efficiency, where the aryl-modified nucleotide, **dT^{aryl}TP**, exhibits an efficient substrate than the **dT^{spin}TP** (Figure 7).

However, a very small number of crystal structures of the ternary complex of family B DNA polymerases have been reported to date. Recently, Marx Lab reported an active ternary complex of KOD DNA polymerase, where a 7-deaza-modified adenosine triphosphate, **dA*TP**

trapped in the nucleotide-binding site.¹⁰² Chaput lab reported the reverse transcriptase activity of Bst DNA polymerase while processing sugar-modified template strands using X-ray crystallography.¹⁰³

The Weinhold group has successfully crystallized an amino purine-modified DNA complexed with DNA adenine methyl transferase, Mtaq1. Where Amino purine modification perfectly fits into the active site as similar to the natural adenine without disturbing the activity of the Mtaq1.¹⁰⁴ Recently, Egli's lab demonstrated that the replacement of natural thymidine analogs by a modified nucleotide increases the thrombin binding affinity and successfully crystallized aptamers modified with either **W** or **K** bound to thrombin, which provides insight into the reason for the increased affinities (Figure 7).¹⁰⁵

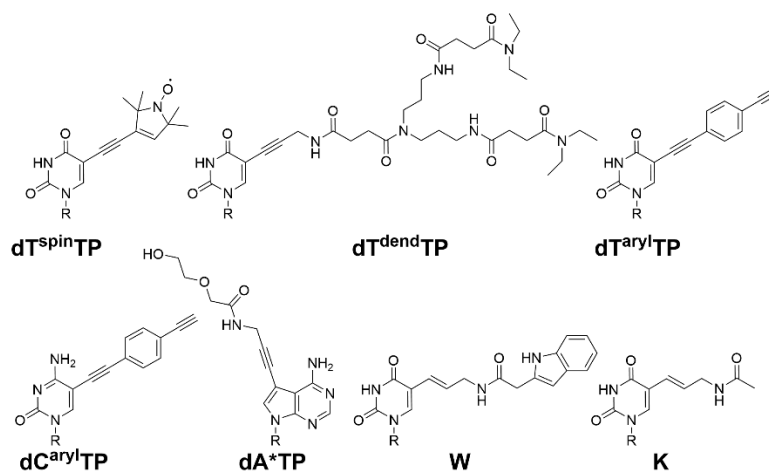


Figure 7. Chemical structure of modified nucleotides used in X-ray crystallography. R= 2'-deoxy ribose sugar with 5'-triphosphate.

1.4.2 Nuclear Magnetic Resonance (NMR)

NMR is one of the most important to study the structure and dynamics of nucleic acids and proteins in solution state. Studying nucleic acids in solution becomes challenging due to spectral overlap of nucleobases, the building blocks of DNA molecules.¹⁰⁶ Recently, site-specific labeling of isotope-labeled (¹³C, ¹⁵N isotopes) nucleosides into the nucleic acids facilitates to study nucleic acids-protein interaction. The structure of nucleic acids-protein complexes can be solved using standard NMR techniques or by sophisticated 2D NMR methods where either the protein or the NA component of the complex is labeled with ¹³C/¹⁵N isotopes.^{107,108} Williamson and coworkers reported an isotope labeling strategy by enzymatic pathway, where they incorporate several purine bases into the HIV-2 transactivation region of RNA (Figure 8).^{109,110} Recently, to synthesize an isotope-labeled dsDNA, a PCR method named

endonuclease sensitive repeat amplification (ESRA) has been developed.¹²⁷ By employing these labeling techniques isotope-labeled dsDNA and RNA were synthesized for the investigation of nucleic acid-protein complexes.¹¹²⁻¹¹⁵

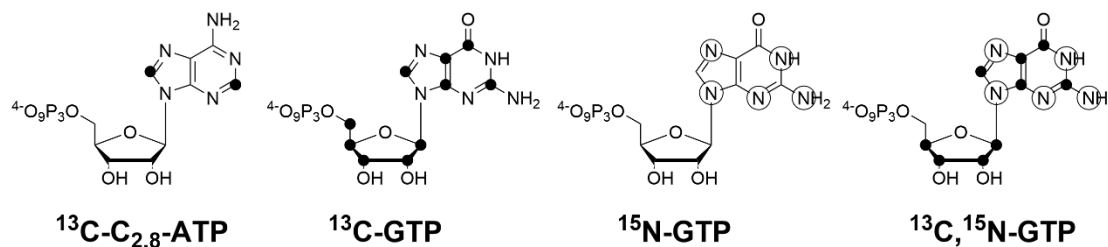


Figure 8. Examples of ^{13}C and ^{15}N labeled NTPs to synthesis isotope-labeled oligonucleotides to monitor RNA-protein interactions.

In this context, fluorine labels offer an efficient system to monitor nucleic acid structures and their binding with proteins. Fluorine-labeled nucleic acids are an effective tool for biophysical studies due to the high gyromagnetic coefficient and sensitivity toward changes in the local chemical environment. This is because most biomolecules do not contain fluorine atoms. The ^{19}F NMR studies on such nucleic acids are highly efficient and have been widely used for this purpose (Figure 9).¹¹⁶⁻¹¹⁸ Fluorine-modified nucleotides have been inserted into oligonucleotide strands to examine their structure, dynamics, and interactions with small molecules in solution or cellular conditions.¹¹⁹⁻¹²⁴ The Micura group synthesized a 2'-SCF₃-modified RNA and monitored the binding interaction with the RNA binding domain of nuclear ribonucleoprotein U1A (U1A-RBD).¹²⁵ In 2017, Hock et al. demonstrated the enzymatic incorporation of Trifluoroacetophenone-Linked Nucleotide (**dA^{TAP}TP**) that shows distinct ^{19}F NMR peaks for the free duplex and the binary complex of DNA-p53 proteins, respectively.¹²⁶ Yan Xu and coworkers introduced a 3,5-bis(trifluoromethyl) benzene moiety into 5' termini of RNA strand using phosphoramidite chemistry to monitor the complex formation between RNA G quadruplex and the telomeric protein TRF2.¹²⁷ Recently, 2'-fluoro-substituted deoxynucleotide has been used as a substrate for the thymine DNA glycosylase (TDG) enzymes to investigate flipping of nucleotide in the free and protein-bound states.¹²⁸

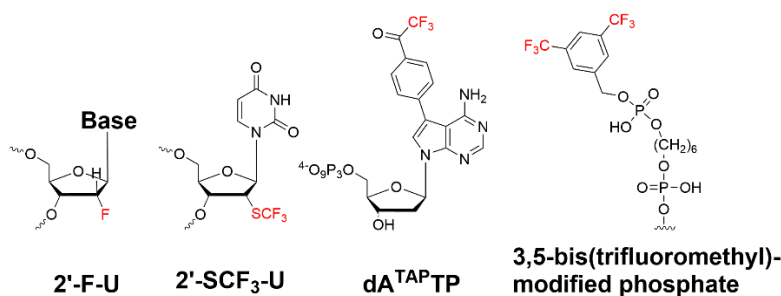


Figure 9. Examples of fluorine labeled probed successfully utilized to monitor NA-protein interactions.

1.4.3 Fluorescence spectroscopy

Fluorescence-based techniques are widely utilized for understanding structural, and mechanistic aspects of nucleic acid-protein interactions. The distinct advantages of fluorescence spectroscopy over alternative spectroscopic methods are: (i) fast responsiveness, provides rapid analysis for biomolecular interactions. (ii) remarkable sensitivity, enabling the tracking of a single molecule through its fluorescence; (iii) selectivity, allowing for the monitoring of a fluorescent-labeled nucleic acids or a protein in its natural environment within a living cell; (iv) high sensitivity of a fluorescent molecule to detect the changes in the surrounding environment and (v) capability to observe a broad range of dynamic processes in real-time over a wide timescale, from femtoseconds to seconds.^{129,130} Therefore, introducing fluorescent probes into the nucleic acids becomes the method of choice to study their structural changes and interactions with proteins and small molecules in real time with high temporal resolution. To monitor this interaction, these fluorescent probes can be incorporated either non-covalently or covalently into the nucleic acid strands.

1.4.3.1 label-free methods

A non-covalent labeling approach relies on the use of fluorogenic dyes, such as YOYO and ethidium bromide, which show differences in binding to nucleic acids with different conformation states. These dyes can exhibit a discernible optical signal for structural changes of nucleic acids due to their interactions with the protein.¹³¹⁻¹³⁴ For example, a non-covalent DNA binder YOYO dye was utilized to study the kinetics of DNA digestion by k-exonuclease and the endonuclease EcoRV using the differences in the fluorescent intensities between single and double-stranded DNA-YOYO complex.¹³¹ An alternative approach, emissive nanomaterials, such as metallic nanoparticles or conjugated polymers, have been strategically employed to develop fluorescence-based assays for detecting nucleic acid binding protein interactions.^{135,136} Utilizing unlabeled oligonucleotides in label-free methods is more cost-effective than employing fluorescently labeled oligonucleotides, making them more important

for low-cost applications and high-throughput detection. However, this method falls short due to the non-specific binding of the multiple fluorophore molecules, which sometimes destabilizes the native nucleic acid-protein interactions and lacks multiplex detection capability.

1.4.3.2 Covalently incorporation of exogenous fluorophores

A more advanced method for detecting nucleic acid-protein interactions involves attaching exogenous fluorophores covalently into nucleic acid sequences. This approach offers advantages over non-covalent labeling, as it minimizes destabilization of the native nucleic acid structure and their biomolecular interactions. Numerous extrinsic fluorophores (such as fluorescein, rhodamine, Alexa, etc.) have been incorporated at specific positions, either chemically or enzymatically, by replacing or modifying the natural base pairs for different applications.¹³⁷⁻¹³⁹ These modifications possess high fluorescence quantum yields and are significantly photostable, they are mainly used for fluorescence detection, such as DNA sequencing, real-time PCR detection, and single-nucleotide mutation screening.¹⁴⁰⁻¹⁴³ For instance, nucleic acid end-labeled with a single fluorophore shows an increase in fluorescence anisotropy upon interacting with proteins, and this method is one of the most widely used methods to date. Additionally, multi-labeling techniques have become convenient for investigating certain DNA-protein interactions. This method uses two fluorescent molecules to create an excimer or a fluorophore-quenching pair. When the target protein binds, the global changes in the structure of nucleic acids hamper the interaction between the fluorescent molecules, causing a noticeable change in the fluorescence signal. One of the early examples of this method was investigating TATA-binding protein (TBP) binding to its target promoter sequence and the consequential DNA deformation.¹⁴⁴ DNA was labeled with fluorescein and rhodamine fluorophores as a RET pair. TBP-induced bending in the promoter duplex causes a reduction in the donor-acceptor distance, resulting in an enhancement in the RET efficiency. The bending event was monitored by observing the changes in fluorescence intensities of the donor and acceptor.¹⁴⁴ Additionally, the molecular beacon containing a fluorophore-quencher pair at the 5' and 3' ends was used to study the binding of a single-stranded binding protein from *E. coli*. An increase in fluorescence intensity was observed upon binding the SSB protein in the loop region.¹⁴⁵ In a similar approach, the binding of thrombin (a serine protease) to a non-catalytic site on single-stranded DNA, which forms a unimolecular quadruplex structure, was detected. The sensor demonstrated a 2.5-fold fluorescence enhancement upon thrombin binding, with an apparent K_d value of around 10 nM when utilizing a fluorescein–dabcyl pair.¹⁴⁶

1.5 Base-modified fluorescence nucleoside analogs (FNAs) to probe nucleic acid-protein interactions

Despite the broader range of applications of these extrinsic fluorophore systems, a significant drawback of most labeled probes is that when they are incorporated into the nucleic acids, they often disturb their native structure. Therefore, in many cases, a long linker is used to connect the probe molecule to the oligonucleotide, aiming to reduce the disruptions to the native structure. In addition, FRET experiments exhibit a distance-dependent nature, which faces Challenges in studying interactions where global changes of the target nucleic acid are not significant. During the folding and recognition of nucleic acids, conformational changes occur at both the global and nucleotide levels. Usually, the conformational changes at the nucleotide level involve alternating its local environment, such as polarity, viscosity, stacking, and hydration, and their interactions with neighboring bases or interactions with the protein binding sites. To better understand these interactions, environment-sensitive nucleoside probes have been introduced at the site of interest. These report the local conformational changes by showing changes in its fluorescence properties, such as quantum yield, emission maximum, lifetime, and anisotropy. The advantage of these Fluorescence nucleoside analogs (FNAs) is that they share structural similarities with natural nucleosides and are minimally perturbed in hydrogen bonding with complementary bases. Hence, Fluorescence nucleoside analogs (FNAs) have become a preferred method for real-time monitoring of nucleic acid-protein interactions. Several such fluorescent nucleotide analogs have been incorporated into the nucleic acid strand chemically or enzymatically at the internal or terminal positions by replacing non-emissive natural nucleosides. Depending on variations in the chemical structures, different classes of FNAs have been reported, such as size-expanded, size-extended, polycyclic aromatic hydrocarbon (PAH), pteridine, and minimally perturbing isomorphous nucleobase analogs (Figure 10).¹⁴⁷

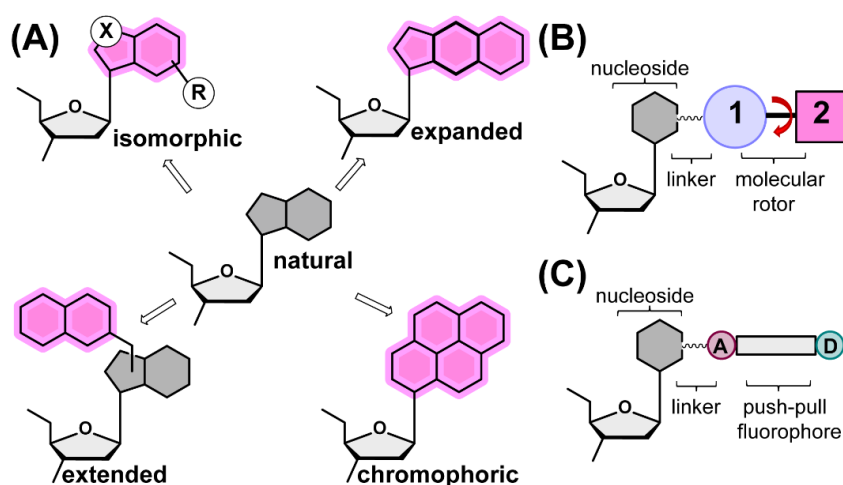


Figure 10. (A) Classification of the fluorescent nucleobase analogs (FNAs). Schematic representation of the (B) FNA containing molecular rotor element shows changes in photophysical properties due to changes in viscosity. (C) FNA containing push-pull fluorophore exhibits changes with solvent polarity. Linker distances vary with different modifications, or sometimes the probe is directly attached to nucleobase.

1.5.1 Size-expanded fluorescent nucleobase analogs

Size-expanded fluorescent nucleobase analogs are constructed by fusing or inserting an additional aromatic ring or a heterocyclic moiety onto the natural nucleobases (Figure 11).¹⁵² Many such nucleobase analogs are emissive and retain the perfect hydrogen base pairing with the complementary bases. Numerous expanded natural base analogs with fluorescence have been constructed and applied in investigating enzymatic incorporations. Due to their larger size and shape than natural bases, these analogs exhibit different activity and specificity towards certain enzymes. Nevertheless, they serve as valuable tools to assess the tolerance and preference of specific enzymes for their substrates, shedding light on the intricate details of the catalytic process. For instance, Wilhelmsson and coworkers reported the incorporation of ring fused cytidine modifications (**tC**, **tC⁰**) using Klenow fragment (Kf) DNA polymerase, where they exhibit slightly higher efficiency of incorporation than the natural cytidine analogs (Figure 11). Further, using the fluorescence anisotropy property of these analogs, they determined the dissociation constant for binding the nucleotides to the DNA polymerase active site.¹⁵¹ A **tC**-labeled DNA as fluorophore and a fluorescent quencher (TEMPO) modified KF DNA polymerase have been employed to detect conformation changes of the DNA-protein complex involved during the incorporation process.¹⁵² Kool et al. reported the incorporation of xDNA pairs by Kf (family A) and Dpo4 (Y-family) DNA Polymerases. The findings indicated that, like Kf, Dpo4 generally favored the accurate selection of the appropriately paired partner for each xDNA (**xA**, **xG**, **xC**, **xT**) base. However, the efficiency was compromised in most instances due to the increased size of the base pairs.¹⁵³ A series of xDNA analogs have been incorporated into a DNA strand using a template-independent DNA polymerase TdT. The study reported that the modified xDNA monomers exhibit similar efficiencies to natural nucleotides.¹⁵⁴ This benzo-expanded adenosine nucleotide analog has recently been employed to develop a detection platform to monitor TdT activity.¹⁵⁵

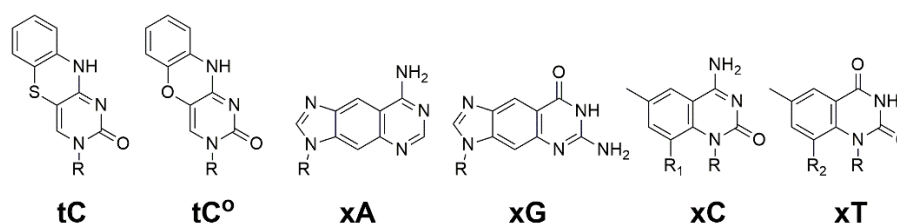


Figure 11. Examples of size-expanded fluorescent nucleobase analogs R= ribose or 2'-deoxyribose.

1.5.2 Size-extended fluorescent nucleobase analogs

Size-extended fluorescent nucleobases are typically designed by attaching an extrinsic fluorophore to the nucleobase through a rigid or flexible linker (Figure 12).^{156,157} As the extended pi-conjugated systems are connected by a flexible bond or a single rotatable bond, these nucleoside analogs show sensitivity towards viscosity. In low-viscosity environments, the intramolecular rotation at flexible bonds leads to internal conversion, resulting in low fluorescence quantum yields. Conversely, within high-viscosity environments or confined spaces like nano compartments and nanocavities, it inhibits internal rotation and enhances the fluorescence of these dyes. The term fluorescent molecular rotors (FMRs) is used to classify such viscosity-sensitive behaviors.¹⁵⁸⁻¹⁶²

The Hocek group has reported green fluorescence protein-based extended fluorescent nucleobase analogs in this context. They synthesized two derivatives of the GFP fluorophore **dC^{FBI}** and **dC^{MBI}** by attaching 3,5-difluoro-4-hydroxy-benzylideneimidazolinone (**FBI**) and 3,5-bis(methoxy)-4-hydroxy-benzylideneimidazolinone (**MBI**) cores at the C5 position of the cytidine base using a short propargylic respectively.¹⁶³ The nucleosides were transformed into the corresponding 5'-triphosphates, **dC^{MBI}TP** and **dC^{FBI}TP**, proving to be effective substrates for two recombinant DNA polymerases, KOD XL and Vent(exo-). Consequently, this enabled the synthesis of single-stranded (ss) and double-stranded (ds) DNA labeled with multiple GFP-like molecular rotor fluorophores at specific nucleotide positions through primer extension (PEX) reactions. Further, they have titrated fluorophore-labeled oligonucleotides with DNA-binding proteins such as SSB and the DNA-binding domain of transcription factor p53 and observed a 2 to 3-fold increase in the fluorescence emission in the presence of proteins.¹⁶³ From these observations, the authors hypothesize that the increase in fluorescence results from restriction in the free rotation through the flexible bond in the enzyme pockets. Tryptophan-based imidazolinone fluorophore-modified cytidine (**dC^{Trp}**) analogs were reported from the same group.¹⁶⁴ This modification also exhibited sensitivity towards viscosity and has been incorporated into the DNA using polymerase-based labeling techniques. It also senses the

protein binding and shows a 2-fold enhancement in the fluorescence intensity. Again, the Hock group has reported another improved FMRs nucleoside, **dC^{VP}** and **dC^{VJ}**, based on the benzylidene cyanoacetamide scaffold. The modifications have been introduced at the C5 positions of the pyrimidine bases by a propargylic linker and subjected to primer extension reaction to synthesize labeled DNA using KOD XL and Vent(exo-) DNA polymerase. Meanwhile, **dC^{VP}** exhibited better incorporation efficiencies than the **dC^{VJ}** analog due to the bulkiness in the modification. Further, **dC^{VP}**-labeled DNA was titrated with single-stranded binding protein (SSB), and it showed a 4-fold increase in the fluorescence intensity upon binding with the protein.¹⁶⁵

Recently, Seo and coworkers synthesized two new deoxyuridine derivatives, **dU^{CN2}**, and **dU^{Py}**, by attaching p-vinylaniline-derived molecular rotor moiety at the C5 position of the deoxyuridine. These molecules show substantial sensitivity towards changes in the viscosity, and 5'-triphosphates of these analogs demonstrate efficient substrate properties in primer extension reactions. By using a combination of primer extension reactions and followed by degradation of the template strand using lambda-exonuclease, they were able to synthesize a single-stranded primer labeled with modified nucleotide. A **dU^{CN2}**-labeled DNA demonstrates a 1.7 to 3-fold increase in fluorescence intensity upon binding with a single-strand binding SSB protein. Also, these new molecular FMR-based nucleosides were shown to effectively explore interactions with nucleolin proteins, which are known for stabilizing nucleolar G-quadruplexes and playing multiple roles in transcriptional regulation. The introduction of modified nucleosides into the c-myc ssDNA sequence, forming a quadruplex structure, did not hinder nucleolin binding, as confirmed by the EMSA assay. The emission intensity of the **dU^{Py}** reporter remained nearly the same for labeled ssDNA alone, in double-stranded and G-quadruplex forms. However, upon binding to nucleolin, there was a 2-fold increase in emission intensity, indicating the stabilization of the quadruplex motif.¹⁶⁶

Also, another class of size-extended fluorescence nucleoside analogs has been developed, where the protein binding process was monitored by fluorescence lifetime.¹⁶⁷ A BODIPY fluorophore was attached to position 5 of deoxycytidine using a minimal propargylic linker (**dC^{bdp}**) and showed a positive relationship between mean fluorescence lifetime and viscosity. Subsequently, enzymatic DNA labeling was carried out with the modified nucleoside triphosphate **dC^{bdp}TP** using PEX (Figure 12). Remarkably, this bulk hydrophobic fluorophore modification does not disturb the DNA structures and binding with the p53 protein. Upon interaction with the protein, it demonstrates changes in the mean lifetime range of 0.5–2.2 ns.¹⁶⁸ For this purpose, some of the size-extended base-modified nucleotides, which showed

changes in emission spectra with changing local polarity, have been reported. 5-(dimethylamino)-naphthalene-1-sulfonyl (Dansyl) conjugated nucleotide **dU^{DNSTP}** and **dU^{AEDANS-STP}** showed an early demonstration of such systems (Figure 12). These nucleotides were utilized to investigate interactions with the Klenow fragment of DNA polymerase I. In the modified duplex-protein binary complex, there is an observed rise in fluorescence intensity accompanied by a blue shift in the emission maximum. This shift is attributed to the increased hydrophobic environment within the protein.^{169,170} On the other hand, **dC^{API}TP**, which has a high quantum yield, is incorporated into the ssDNA and dsDNA using primer extension reactions. The labeled oligonucleotides were subjected to fluorescence titration with SSB and p53 as representative DNA-binding proteins, indicating a modest increase in fluorescence emission (2–3 fold) coupled with a slight blue shift in the emission spectrum upon binding. However, this light-up response was not universal, as other tested DNA-binding proteins did not exhibit similar effects.¹⁷¹

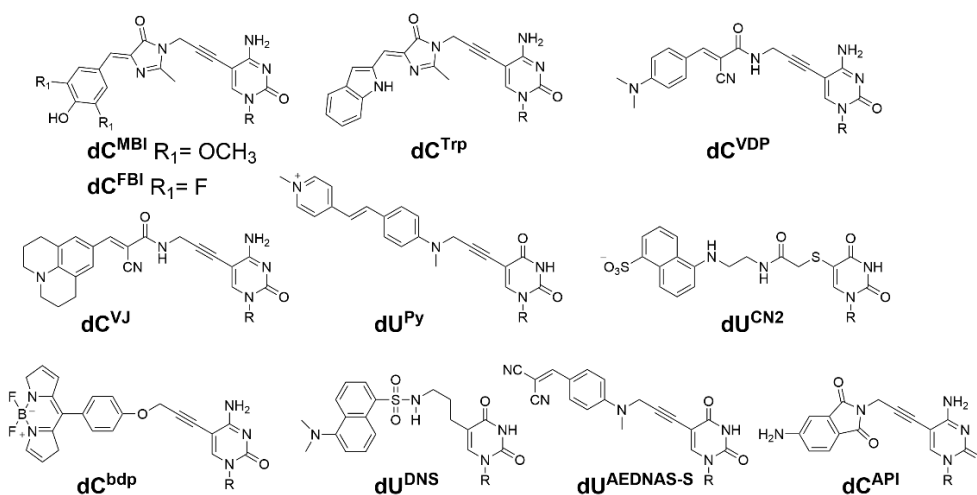


Figure 12. Size-extended fluorescent nucleobase analogs. R= 2'-deoxyribose sugar.

1.5.3 Fluorescent poly aromatic hydrocarbon (PAH) and pteridine nucleobase analogs

Substituting natural nucleobases with recognized fluorescent polycyclic aromatic hydrocarbons (PAHs) results in a distinctive group of polycyclic aromatic hydrocarbons (PAH) base analogs that do not possess the Watson-Crick (WC) hydrogen bonding interface (Figure 13). This kind of nucleoside analogs shows large stokes shifts, a diverse range of quantum yields, and emission wavelengths spanning the visible spectrum. Kool and coworkers have synthesized “oligodeoxyfluorosides” (ODFs) by incorporating multiple PAH fluorophores into a DNA. These multi-chromophoric oligomers form densely packed emissive assemblies, displaying intricate energy transfer phenomena among the chromophores. This makes them highly responsive to their surroundings, leading to complex fluorescence emissions across the

visible spectrum. Further, these have been used for detecting toxic gases, metal ions, and cellular imaging.¹⁷²⁻¹⁷⁵ The same group has reported a fluorescence-based assay using the PAH nucleoside **Y** for detecting the activity of ALKBH3, RNA demethylase enzymes that eliminates the epitranscriptomic mark N1 -methyladenosine (m1 A) from mRNAs and tRNAs. They placed this modification **Y** next to the m1A in the DNA, where it shows low fluorescence but displays 10-fold high fluorescence upon treatment with the enzyme. This probe demonstrated specificity for ALKBH3 over its closely related homolog ALKBH2 and is also employed to monitor the enzyme activity in living cells by fluorescence microscopy.¹⁷⁶ One of the most successful PAH nucleosides is Coumarin 102 (**C102**), which was reported two decades ago.¹⁷⁷ Computational and spectroscopic data revealed that incorporating this modification into the DNA strands does not disrupt the helical structure of DNA. This probe was employed to investigate the interaction mechanism between DNA and the APE1 endonuclease. Further, a biochemical assay was performed to validate the recognition and cleavage of artificial oligonucleotide substrates containing **C102** as a base pair by the enzyme.¹⁷⁸ Similarly, several pteridine nucleoside analogs have been developed by mimicking the natural purine bases with the planar heterobicyclic aromatic compounds. One such fluorescent pteridine analog **6MI** was utilized by Shamoo and coworkers, to examine the interactions involving UP1, an unwinding protein, and a part of heterogeneous ribonucleoprotein (hnRNP A1), which plays a role in RNA processing and chromosome maintenance. UP1 usually binds to the purine-rich nucleic acid sequence and unwinds the human telomeric G-quadruplexes. The authors selectively inserted the modification **6MI** into the nucleic acid sequence where it will not hamper the secondary structure formation and unstacked upon binds with protein. Fluorescence titrations unveiled a substantial enhancement in the steady-state emission of **6MI**, facilitating direct observation of the protein binding and subsequent unfolding of the quadruplex. Further, they co-crystallized the **6MI**-labeled sequence with the UP1 to prove that the modification does not disturb the protein-DNA recognition and folding process. Another adenosine analog, **6MAP**, was employed as a fluorescence reporter to investigate the base-flipping activity conducted by DNA photolyase (PL). This enzyme is responsible for repairing cyclobutylpyrimidine dimers found in photodamaged DNA.¹⁷⁹

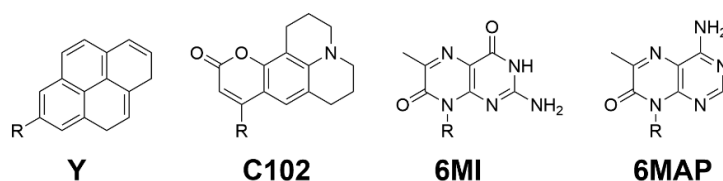


Figure 13. Poly aromatic hydrocarbons (PAH) and pteridine fluorescent nucleobase analogs. R= 2'-deoxyribose sugar.

1.5.4 Isomorphic nucleobase analogs

Isomorphic nucleobase analogs are a type of heterocycles that have a close resemblance to their corresponding natural nucleobases.¹⁸⁰⁻¹⁸² These analogs offer a distinct advantage due to their strong similarity to native nucleosides and structurally minimally invasive nature, especially when compared to various other analogs discussed previously (Figure 14). These analogs exhibit a notable quantum yield and possess the capability to investigate the alterations in the microenvironment by monitoring changes in photophysical properties. A quintessential example of isomorphic nucleobase analogs is 2-amino purine (**2AP**), which is extensively employed for analyzing nucleic acid structures, dynamics, and binding. It is highly fluorescent, environment-sensitive, non-perturbing, and an isoster of the adenine. Free **2AP** is highly emissive in the aqueous medium, but when incorporated into the DNA or RNA strand, its fluorescence is dramatically quenched due to stacking and the electron transfer process. This behavior makes **2AP** a good reporter molecule to monitor the DNA-protein interactions, which involves the distortions of the stacking interactions and alternations in the local environment of the nucleic acids binding site. To date, 2-amino purine (**2AP**) stands out as the most extensively employed fluorescent nucleotide analog (FNA) in enzymatic and biochemical investigations of DNA polymerases. For example, **2AP** has been applied in studies of different classes of DNA polymerases, such as the Klenow fragment of Escherichia coli DNA Polymerase I¹⁸³⁻¹⁸⁶, DNA Polymerase β ²⁰³, and bacteriophage T4 DNA Polymerase.²⁰⁴⁻²⁰⁵ Further, **2AP** was used to study DNA repair enzymes, such as M.HhaI and M.TaqI methyl transferase, and the EcoP15I type III restriction-modification enzyme.¹⁹⁰⁻¹⁹² Usually, methyl transferase interacts with the base-flipped nucleic acid substrates, which leads to destabilization of the base stacking interaction. **2AP** has been utilized as a fluorescent probe to observe the disruption of local base pairing and the base flipping of a nucleotide residue that facilitates enhanced substrate accessibility to the enzyme. **2AP** is also employed to study the junction revolving DNA polymerases CCE1 and thrombin. It has been incorporated into the DNA strand and demonstrated a 39-fold increase in the case of CCE1, which majorly distorts the DNA base pairing upon binding. Meanwhile, only a 10-fold fluorescence increase has been observed for thrombin binding, resulting in minimum geometric distortions.¹⁹³ **2AP** was directly integrated into an RNA sequence to explore the adenosine deamination activity mediated by the RNA-editing enzyme adenosine deaminase (ADAR). Placing **2AP** at the editing site resulted in a 3.3-fold increase in fluorescence intensity. Utilizing a **2AP**-labeled RNA duplex unveiled

specific details of the catalytic process.^{194,195} Although **2AP** has been used in a broad range of applications, the excitation and emission maximum wavelengths are in the UV region, low pH sensitivity in the medium to high pH range, and the drastic quenching in fluorescence emission upon incorporation in the oligonucleotides caused limitations in its applications, specifically in cell microscopy.¹⁹⁶⁻¹⁹⁸

Another fluorescent analog, pyrrolo-deoxycytidine (**pC**), has been introduced into the DNA strand, and steady-state fluorescence analysis revealed the binding process of DNA-reverse transcriptase.¹⁹⁹ Further, this analog has been utilized to monitor the structural changes in the promoter region during the transition from initiation to elongation steps by T7 RNA polymerase.^{200,201} Hudson and coworkers used the 6-Phenylpyrrolocytidine (**PhpC**), a structurally conservative and highly fluorescent cytidine analog, to probe the cleavage of **PhpC**-labeled RNA by HIV-1 ribonuclease H. This approach demonstrated superior responsiveness and rapidity compared to a molecular beacon method.²⁰² In this regard, Tor and coworkers obtained a series of isomorphous fluorescent nucleobases (**thA**, **thG**, **^{tz}A**, **^{tz}G**) by mutating the five-member heterocyclic ring of the purine bases with thieno [3,4-d]-pyrimidine and isothiazolo [4,3-d]-pyrimidine scaffold.²⁰³⁻²⁰⁵ Among these analogs, **thG** appears as an excellent emissive fluorescent reporter, with a long fluorescence lifetime and highly sensitive to the local environment. This analog has been incorporated into the RNA and DNA using transcription reaction and primer extension reactions by replacing the guanidine residue, respectively, and used to study DNA mismatching.²⁰⁶⁻²⁰⁸ **thA** also has been demonstrated in several applications to investigate enzymatic processes. For instance, adenosine deaminase enzyme (ADA), which converts adenosine to inosine moiety, recognized **thA** as a substrate. It converts **thA** to **thI**, which shows substantial changes in fluorescence intensity with a blue shift in the emission maximum. Even this conversion kinetics can be monitored in real time by measuring the changes in the fluorescence intensity. ADA is a drug target for numerous diseases, this technique could establish a high-throughput screening platform to find ADA inhibitors.²⁰⁹

Another one of the more successful isomorphous analogs has been developed by the Tor group by attaching the 5-membered-heterocyclic moiety to the C5 positions of the pyrimidine bases. Additionally, these compounds closely mimicked the natural pyrimidines in terms of molecular size and the Watson-Crick hydrogen bonding interface. They have explored the spectroscopic properties of various combinations of the nucleoside core (2'-deoxyuridine, 2'-deoxycytidine) with five-membered heterocycles (furan, thiophene, oxazole, and thiazole). The studies revealed that derivatives like 5-(fur-2-yl)-uracil (**^{Fu}dU** and **^{Fu}U**) displayed optimal

photophysical properties, making them valuable probes for DNA and RNA. ¹⁹F-dU, sensitive to polarity and rigidity, emerged as an effective tool for investigating nucleic acid structures and interactions with small molecules and proteins. These molecules proved useful in distinguishing between alternative folds or hybridization states through intensity- and wavelength-based fluorescence sensing.^{210,211}

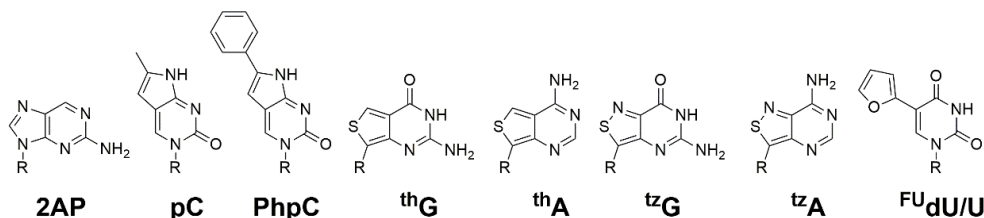


Figure 14. Examples of Isomorphous base analogs. R= 2'-deoxy ribose sugar.

1.6 Motivation and outline of the thesis

Nucleic acid processing enzymes incorporate a wide range of modifications, and this process highly depends on the accepting capability of the DNA and RNA polymerases and transferases. However, the underlying mechanism responsible for incorporating different modifications still needs to be elucidated. Hence, predicting substrate tolerance by polymerases poses significant challenges. Therefore, developing and refining methods for studying interactions between nucleic acids and proteins are crucial for designing appropriately modified nucleotides for myriad biological applications. Recent developments in high-throughput biopolymer crystallography and biomolecular NMR spectroscopy has helped in gaining structural insights into protein-DNA/RNA complex formation. Notably, X-ray crystallography-based methods typically provide still shots of exceptionally dynamic biomolecular complexes. Therefore, alternative experimental techniques are essential to investigate nucleic acid-protein interactions in real time. In this context, fluorescent nucleobase analogs have become the point of attraction for monitoring such interactions under equilibrium conditions. It is also becoming increasingly clear that a detailed understanding of an enzymatic system, such as DNA polymerases, could be gained only by using complementing biophysical techniques. While each technique provides an important piece of information, correlating these informations using uniquely labeled ONs becomes difficult. In light of the merits and drawbacks of each technique, the challenge lies in establishing an efficient platform that seamlessly correlates fluorescence, NMR, and X-ray crystallography techniques. Such a system could offer a comprehensive understanding of the steps involved in the enzymatic incorporation mediated by different polymerases and transferases.

To address these unmet needs, our lab has developed several environment-sensitive fluorescent nucleoside analogs by conjugating heterocyclic moiety at the C5 positions of pyrimidine bases (Figure 15).⁴³ These molecules are minimally invasive and possess moderate to good quantum yields. Notably, the presence of a connecting biaryl bond between the heterocyclic moiety and uridine ring endows molecular rotor system in the nucleoside analog. Hence, high sensitivity to local viscosity, polarity, and stacking interaction with neighboring bases make these analogs highly valuable in probing nucleic acid folding and recognition processes. However, these probes share an excitation wavelength at the UV region and are unsuitable for analysis in cellular environment. In order to develop in-cell compatible nucleoside analog probes, ¹⁹F atom(s) were introduced at an appropriate site without affecting their conformation sensitivity. Based on these considerations, a series of dual-app probes were developed in our group.¹³⁹ A few of these responsive analogs are good substrates for enzymatic incorporation by polymerases and transferases and greatly aid in probing nucleic acid lesions, nucleic acid-small molecules interactions.²¹²⁻²¹⁴ Further these analogs serve as good biochemical tools to investigate conformation and recognition properties of therapeutically important nucleic acid motifs such as (i) G-quadruplexes and i-motifs,^{123,124} (ii) bacterial ribosomal decoding site-antibiotic interactions²¹⁵ and (iii) RNA-protein interactions.²¹⁶

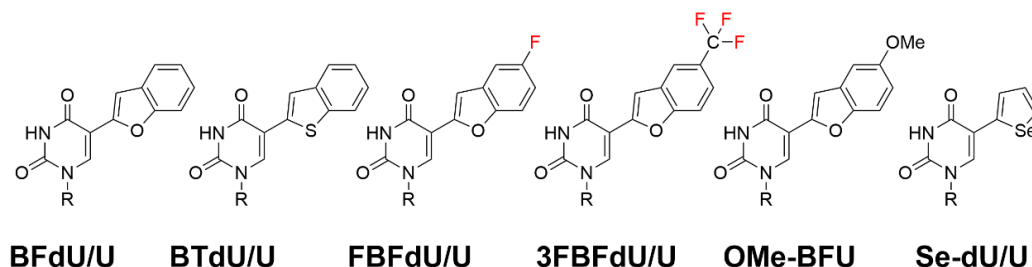


Figure 15. Examples of heterocyclic-modified microenvironment-sensitive fluorescent probe.

The aim of this doctoral work is to create a simple and efficient platform using these environment-sensitive nucleotide substrates, X-ray crystallography, and biophysical methods to investigate the individual steps of the catalytic cycle during the enzymatic incorporation process (Figure 16). The advantages of such probe systems are three-fold: firstly, the substrate itself can function as a fluorescent probe, which could provide a direct fluorescence readout for studying events in the catalytic cycle under equilibrium conditions. Secondly, the efficiency of incorporation and binding data, assessed through biochemical, and fluorescence assays, and NMR spectroscopy, can be directly correlated with the structural information obtained through X-ray crystallography. Thirdly, these smart probe systems could pave ways for developing

screening platforms for developing new inhibitors against disease-causing nucleic acid processing enzymes (e.g., viral polymerases).

The workflow of the thesis is divided into four chapters including the present introductory chapter. In Chapter 2, we studied the incorporation of heterocycle-modified nucleotides, **BFdUTP**, **BTdUTP**, and **SedUTP**, by using family A and family B DNA polymerases namely, KlenTaq and KOD/Vent(exo-).²¹⁷ These nucleotides exhibit variations in the incorporation efficiency depending on the size of the heterocycle modification and the DNA polymerase sequence family employed. In particular, KlenTaq demonstrated sensitivity towards the size of the heterocycle attached to the nucleobase, where it only accepts one heterobicycle-modified nucleotide (**BFdUTP/BTdUTP**) into the ON while efficiently incorporating selenophene-modified nucleotide analog at multiple sites.

Microenvironment-sensitive fluorescent probes (dU*TP)

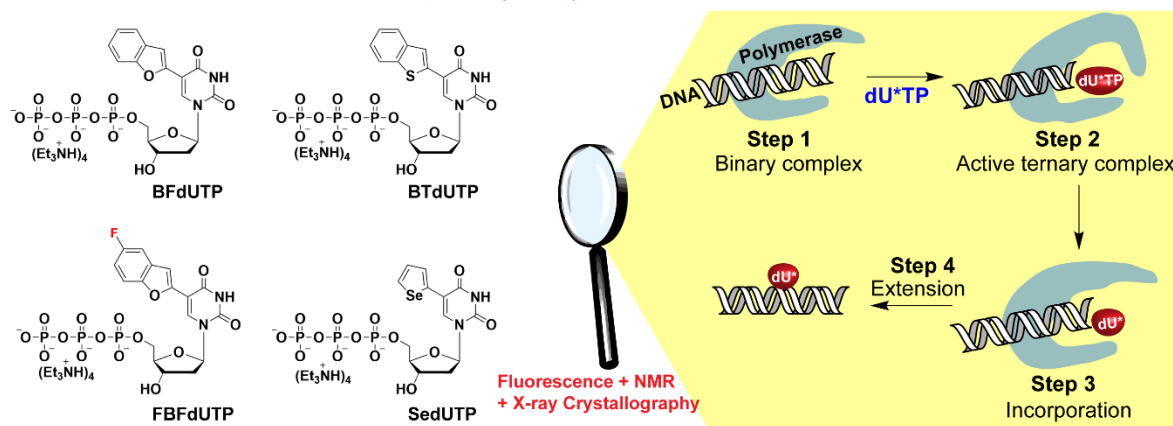


Figure 16. Design strategy to monitor each step of enzymatic incorporation of microenvironment-sensitive probes by DNA polymerases.

On the other hand, KOD and Vent(exo-) exhibit high substrate tolerance and incorporate all three modified nucleotides, resulting in ONs with multiple labels. Notably, **BFdUTP** serves as an efficient reporter, offering real-time fluorescence readouts for monitoring enzymatic incorporation and estimating binding events in the reaction cycle. A comprehensive analysis, including incorporation profiles, fluorescence data, and crystal structure of the ternary complex of KlenTaq DNA polymerase, **BFdUTP**, and primer-template duplex, provides a detailed understanding of the incorporation mechanism of heterocycle-modified nucleotides.

In Chapter 3, we introduced a nucleotide that contains a fluorobenzofuran moiety at the C5 position of 2'-deoxyuridine (**FBFdUTP**). This nucleotide acts as a responsive fluorescent and ¹⁹F NMR probe, exhibiting differences in incorporation efficiency for KlenTaq and KOD

DNA polymerases. Further, we developed a two-channel spectroscopic platform by utilizing fluorescence and ^{19}F NMR techniques to investigate the binding events and enzymatic incorporation process. We also compared three crystal structures of the KlenTaq DNA polymerase ternary complexes containing **FBFdUTP/SedUTP/BFdUTP** and primer-template duplex. This comparison provided valuable insights into the mechanism of incorporation of heterocyclic moieties and the chemical space tolerance of DNA polymerases.

Finally, in Chapter 4, our investigation was extended to the incorporation of these modified nucleotides by terminal deoxynucleotidyl transferase (TdT), a member of the family X DNA polymerases with a unique capability to incorporate nucleotides at the 3'-end of single-stranded DNA (ssDNA) ONs in a template-independent manner. Our modified nucleotide analogs exhibited high substrate efficiency for enzymatic incorporation by TdT compared to thymidine. Consequently, the versatility of **BFdUTP** was broadened as it not only contributed to the understanding of the catalytic cycle but also could prove useful in devising a screening platform to identify potential TdT inhibitors.

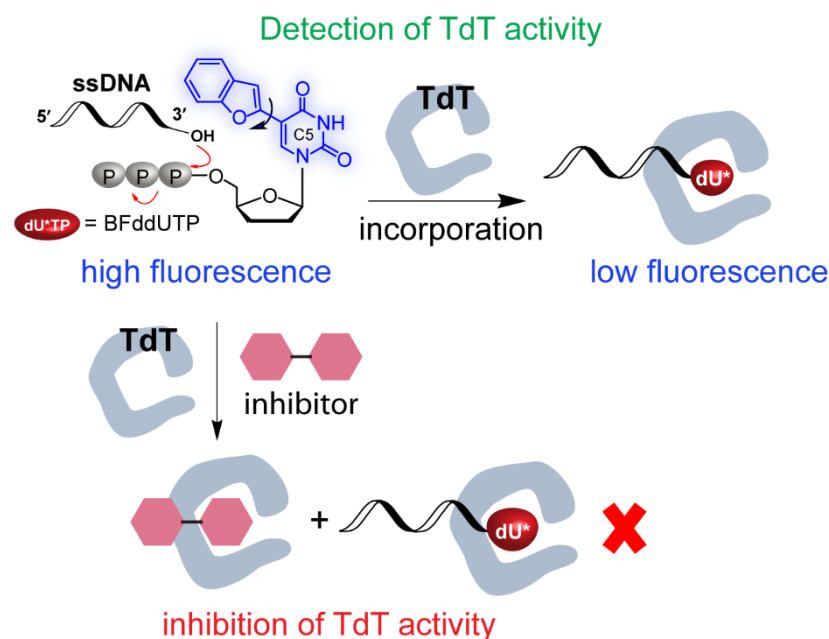


Figure 17. The schematic diagram for detecting TdT activity and monitoring the inhibition process.

1.7 References

1. Crick, F.H.C.; Barnett, L.; Brenner, S.; Watts-Tobin, R.J. General nature of the genetic code for proteins. *Nature*, **1961**, 192, 1227–1232.
2. Hermann, T.; Patel, D. J. RNA bulges as architectural and recognition motifs. *Structure*, **2000**, R47–R54.

3. Leulliot, N.; Varani, G. Current topics in RNA-protein recognition: control of specificity and biological function through induced fit and conformational capture. *Biochemistry*. **2001** *40*, 7947-56.
4. McCarty, M. Discovering genes are made of DNA. *Nature*, **2003**, *421*, 406.
5. Kemp, M. The Mona Lisa of modern science. *Nature*, **2003**, *421*, 416-420.
6. Roeder, R.G. The eukaryotic transcriptional machinery: complexities and mechanisms unforeseen. *Nat. Med.*, **2003**, *9*, 1239-1244.
7. Mandal, M.; Breaker, R. Gene regulation by riboswitches. *Nat Rev Mol Cell Biol*, **2004**, *5*, 451-463.
8. Kruglyak, L.; Stern, D.L. An embarrassment of switches. *Science*, **2007**, *317*, 758-759.
9. Guo, J. et al. Four-color DNA sequencing with 3'-O-modified nucleotide reversible terminators and chemically cleavable fluorescent dideoxynucleotides. *Proc. Natl. Acad. Sci.* **2008**, *105*, 9145-9150.
10. Chen, T.; Hongdilokkul, N.; Liu, Z.; Thirunavukarasu, D.; Romesberg, F. E. The expanding world of DNA and RNA. *Curr. Opin. Chem. Biol.* **2016**, *34*, 80-87.
11. Aschenbrenner, J.; Marx, A. DNA polymerases and biotechnological applications. *Curr. Opin. Biotechnol.* **2017**, *48*, 187-195.
12. Zhou, C. Y.; Alexander, S. C.; Devaraj, N. K. Fluorescent turn on probes for wash-free mRNA imaging via covalent site-specific enzymatic labeling. *Chem. Sci.* **2017**, *8*, 7169-7173.
13. Hocek, M. Enzymatic synthesis of base-functionalized nucleic acids for sensing, cross-linking, and modulation of protein-DNA binding and transcription. *Acc. Chem. Res.* **2019**, *52*, 1730-1737.
14. Ashley, J.; Potts, I. G.; Olorunniji, F. J. Applications of Terminal Deoxynucleotidyl Transferase Enzyme in Biotechnology. *ChemBioChem*. **2023**, *24*, e2022005.
15. Matteucci, M. D.; Caruthers, M. H. Synthesis of deoxyoligonucleotides on a polymer support. *J. Am. Chem. Soc.* **1981**, *103*, 3185-3191.
16. Beaucage, S. L.; Caruthers, M. H. Deoxynucleoside phosphoramidites-A new class of key intermediates for deoxypolynucleotide synthesis. *Tetrahedron Lett.* **1981**, *22*, 1859-1862.
17. Hocek, M. Synthesis of base-modified 2'-deoxyribonucleoside triphosphates and their use in enzymatic synthesis of modified DNA for applications in bioanalysis and chemical Biology. *J. Org. Chem.* **2014**, *79*, 9914-9921.
18. Klöcker, N.; Weissenboeck, F. P.; Rentmeister, A. Covalent labeling of nucleic acids. *Chem. Soc. Rev.* **2020**, *49*, 8749-8773.
19. McKenzie, L. K.; El-Khoury, R.; Thorpe, J. D.; Damha, M. J.; Hollenstein, M. Recent progress in non-native nucleic acid modifications. *Chem. Soc. Rev.* **2021**, *50*, 5126-5164.
20. Preston, B. D.; Loeb, L. A. Enzymatic synthesis of site-specifically modified DNA. *Mutat. Res.* **1988**, *200*, 21-35.
21. Hollenstein, M. Nucleoside Triphosphates-Building Blocks for the Modification of Nucleic Acids. *Molecules*. **2012**, *17*, 13569.
22. Kim, T. W.; Delaney, J. C.; Essigmann, J. M. et al. Probing the active site tightness of DNA polymerase in subangstrom increments, *Proc. Natl. Acad. Sci.* **2005**, *102*, 15803-15808.
23. Kool, E. T.; Sintim, H. O. The difluorotoluene debate—a decade later, *Chem. Commun.* **2006**, *36*, 65-3675.
24. Ganai, R. A.; Johansson, E. DNA Replication-A Matter of Fidelity. *Mol. Cell* **2016**, *62*, 745-755.

25. Cahova, H.; Pohl, R.; Bednarova, L.; Novakova, K.; Cvacka, J.; Hocek, M. Synthesis of 8-bromo-, 8-methyl- and 8-phenyl-dATP and their polymerase incorporation into DNA. *Org. Biomol. Chem.* **2008**, 6, 3657–3660.
26. Lam, C.; Hipolito, C.; Perrin, D. M. Synthesis and enzymatic incorporation of modified deoxyadenosine triphosphates. *Eur. J. Org. Chem.* **2008**, 2008, 4915–4923.
27. Gao, M.; Orlova, M. M.; Georgiadis, et al., Conferring RNA polymerase Activity to a DNA polymerase: A single residue in reverse transcriptase controls substrate selection, *Proc. Natl. Acad. Sci. U.S.A.* **1997**, 94, 407–411.
28. Weisbrod, S. H.; Marx, A. Novel strategies for the site-specific covalent labelling of nucleic acids. *Chem. Commun.* **2008**, 5675–5685.
29. Jäger, S.; Rasched, G.; Kornreich-Leshem, H. et al. A versatile toolbox for variable DNA functionalization at high density. *J. Am. Chem. Soc.* **2005**, 127, 15071–15082.
30. Hottin, A.; Marx, A. Structural Insights into the Processing of Nucleobase-Modified Nucleotides by DNA Polymerases. *Acc. Chem. Res.* **2016**, 49, 418–427.
31. Hocek, M.; Fojta, M. Cross-coupling reactions of nucleoside triphosphates followed by polymerase incorporation. Construction and applications of base-functionalized nucleic acids, *Org. Bio.Chem.* **2008**, 6, 2233–2241.
32. Boge, N.; Jacobsen, M. I.; Szombati, Z. et al. Synthesis of DNA strands site-specifically damaged by c8-arylamine purine adducts and effects on various DNA polymerases. *Chem. Euro. J.* **2008**, 14, 11194–11208.
33. Lee, S. E.; Sidorov, A.; Gourlain, T. et al. Enhancing the catalytic repertoire of nucleic acids: a systematic study of linker length and rigidity. *Nucleic Acids Res.* **2001**, 29, 1565–1573.
34. Vyas, R.; Zahurancik, W. J.; Suo, Z. Structural basis for the binding and incorporation of nucleotide analogs with L-stereo chemistry by human DNA polymerase λ . *Proc. Natl. Acad. Sci. U. S. A.* **2014**, 111, E3033–E3042.
35. Bergen, K.; Steck, A. L.; Strutt, S.; Baccaro, A.; Welte, W.; Diederichs, K.; Marx, A. Structures of KlenTaq DNA polymerase caught while incorporating C5-modified pyrimidine and C7-modified 7-deazapurine nucleoside triphosphates. *J. Am. Chem. Soc.* **2012**, 134, 11840–11843.
36. Wolberger, C. How structural biology transformed studies of transcription regulation *J. Biol. Chem.* **2021** 296, 10074.
37. Hilario, J.; Kowalczykowski, S. C. Visualizing protein DNA interactions at the single-molecule level. *Curr. Opin. Chem. Biol.* **2010**, 14, 15–22.
38. Pagano, J.M.; Clingman, C.C.; Ryder, S. P. Quantitative approaches to monitor protein-nucleic acid interactions using fluorescent probes. *RNA*. **2011**, 17, 14–20.
39. Su, X.; Xiao, X.; Zhang, C.; Zhao, M; Nucleic acid fluorescent probes for biological sensing *Appl. Spectrosc.* **2012**, 66,1249–1261.
40. Leung, C-H.; Chan, D.S-H.; He, H-Z.; Cheng, Z.; Yang, H.; Ma, D-L. Luminescent detection of DNA-binding proteins. *Nucleic Acids Res.* **2012**, 40, 941–955.
41. Michel B.Y.; Dziuba, D.; Benhida, R.; Demchenko, A. P.; Burger, A. Probing of nucleic acid structures, dynamics, and interactions with environment-sensitive fluorescent labels. *Front. Chem.* **2020**, 8, 112.
42. Sinkeldam, R. W.; Greco, N. J.; Tor, Y. Fluorescent analogs of biomolecular building blocks: design, properties, and applications. *Chem. Rev.* **2010**, 110, 2579–2619.
43. Tanpure, A. A.; Pawar, M. G.; Srivatsan, S. G. Fluorescent Nucleoside Analogs: Probes for Investigating Nucleic Acid Structure and Function. *Isr. J. Chem.* **2013**, 53, 366–378.
44. Xu, W.; Chan, K. M.; Kool, E. T. Fluorescent nucleobases as tools for studying DNA and RNA. *Nat Chem.* **2017**, 11, 1043–1055.

45. Dziuba, D. Environmentally sensitive fluorescent nucleoside analogues as probes for nucleic acid - protein interactions: molecular design and biosensing applications. *Methods Appl Fluoresc.* **2022** *10*, 044001.
46. Kasuya, T.; Hori, S.; Hiramatsu, H.; Yanagimoto, T. Highly accurate synthesis of the fully 2'-fluoro-modified oligonucleotide by Terminator DNA polymerases. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 2134–2136.
47. Peng, C. G.; Damha, M. J. Polymerase-directed synthesis of 2'-Deoxy-2'-fluoro- β -D-arabino nucleic acids. *J. Am. Chem. Soc.* **2007**, *129*, 5310–5311.
48. Veedu, R. N.; Vester, B.; Wengel, J. Enzymatic Incorporation of LNA Nucleotides into DNA Strands. *ChemBioChem*, **2007**, *8*, 490–492.
49. Diafa, S.; Evequoz, D.; Leumann, C. J.; Hollenstein, M. Enzymatic synthesis of 7',5'-bicyclo-DNA Oligonucleotides. *Chem.-Asian J.* **2017**, *12*, 1347–1352.
50. Ichida, J. K.; Horhota, A.; Zou, K. Y.; McLaughlin, L. W.; Szostak, J. W. High fidelity TNA synthesis by Terminator polymerase. *Nucleic Acids Res.* **2005**, *33*, 5219–5225.
51. Renders, M.; Emmerechts, G.; Rozenski, J.; Krecmerová, M.; Holý, A.; Herdewijn, P. Enzymatic synthesis of phosphonomethyl oligonucleotides by Terminator polymerase. *Angew. Chem. Int. Ed.* **2007**, *46*, 2501–2504.
52. Ogino, T.; Sato, K.; Matsuda, A. Incorporation of 2'-deoxy-2'-isonucleoside 5'-triphosphates (iNTPs) into DNA by A- and B-family DNA polymerases with different recognition mechanisms. *ChemBioChem*, **2010**, *11*, 2597–2605.
53. Kataoka, M.; Kouda, Y.; Sato, K.; Minakawa, N.; Matsuda, A. Highly efficient enzymatic synthesis of 3'-deoxyapionucleic acid (apioNA) having the four natural nucleobases. *Chem. Commun.* **2011**, *47*, 8700–8702.
54. Augustin, M. A.; Ankenbauer, W.; Angerer, B. Progress towards single-molecule sequencing: enzymatic synthesis of nucleotide-specifically labeled DNA. *J. Biotechnol.* **2001**, *86*, 289–301.
55. Thoresen, L. H.; Jiao, G. S.; Haaland, W. C.; Metzker, M. L.; Burgess, K. rigid, conjugated, fluoresceinated thymidine triphosphates: syntheses and polymerase mediated incorporation into DNA analogues. *Chem. – Eur. J.* **2003**, *9*, 4603–4610.
56. Dziuba, D.; Jurkiewicz, P.; Cebecauer, M.; Hof, M.; Hocek, M. A rotational BODIPY nucleotide: an environment-sensitive fluorescence-lifetime probe for DNA interactions and applications in live-cell microscopy. *Angew. Chem., Int. Ed.* **2016**, *55*, 174–178.
57. Guixens-Gallardo, P.; Humpolickova, J.; Miclea, S. P.; Pohl, R.; Kraus, T.; Jurkiewicz, P.; Hof, M.; Hocek, M. Thiophene-linked tetramethyl bodipy-labeled nucleotide for viscosity-sensitive oligonucleotide probes of hybridization and protein–DNA interactions. *Org. Biomol. Chem.* **2020**, *18*, 912–919.
58. Dziuba, D.; Pospíšil, P.; Matyášovský, J.; Brynda, J.; Nachtigallová, D.; Rulišsek, L.; Pohl, R.; Hof, M.; Hocek, M. solvatochromic fluorene-linked nucleoside and DNA as color-changing fluorescent probes for sensing interactions. *Chem. Sci.* **2016**, *7*, 5775–5785.
59. Kuba, M.; Pohl, R.; Hocek, M. Synthesis of 2'-deoxycytidine and its triphosphate bearing tryptophan-based imidazolinone fluorophore for environment sensitive fluorescent labelling of DNA *Tetrahedron*, **2018**, *74*, 6621–6629.
60. Sletten, E. M.; Bertozzi, C. R. From Mechanism to Mouse: A Tale of Two Bioorthogonal Reactions. *Acc. Chem. Res.*, **2011**, *44*, 666–676.
61. Schoch, J.; Jaeschke, A. Synthesis and enzymatic incorporation of norbornene-modified nucleoside triphosphates for Diels–Alder bioconjugation. *RSC Adv.*, **2013**, *3*, 4181–4183.

62. Matyasovsky, J.; Perlikova, P.; Malnuit, V.; Pohl, R.; Hocek, M. Thiol–ene reaction: synthetic aspects and mechanistic studies of an anti-Markovnikov-selective hydrothiolation of olefins. *Angew. Chem. Int. Ed.* **2016**, *55*, 15856–15859.
63. Ploschik, D.; Roenicke, F.; Beike, H.; Strasser, R.; Wagenknecht, H.-A. DNA primer extension with cyclopropenylated 7-deaza-2'-deoxyadenosine and efficient bioorthogonal labeling in vitro and in living cells. *ChemBioChem.* **2018**, *19*, 1949–1953.
64. Panattoni, A.; Pohl, R.; Hocek, M. Flexible alkyne-linked thymidine phosphoramidites and triphosphates for chemical or polymerase synthesis and fast postsynthetic DNA functionalization through copper-catalyzed alkyne–azide 1,3-dipolar cycloaddition. *Org. Lett.*, **2018**, *20*, 3962–3965.
65. Hong, S. L.; Chen, T.; Zhu, Y. T.; Li, A.; Huang, Y. Y.; Chen, X. Live-Cell stimulated raman scattering imaging of alkyne-tagged biomolecules. *Angew. Chem., Int. Ed.*, **2014**, *53*, 5827–5831.
66. Danenberg, P. V.; Bhatt, R. S.; Kundu, N. G.; Danenberg, K.; Heidelberger, C. Interaction of 5-ethynyl-2'-deoxyuridylate with thymidylate synthetase. *J. Med. Chem.* **1981**, *24*, 1537–1540.
67. Ross, H. H.; Rahman, M.; Levkoff, L. H.; Millette, S.; Martin-Carreras, T.; Dunbar, E. M.; Reynolds, B. A.; Laywell, E. D. Ethynyldeoxyuridine (EdU) suppresses in vitro population expansion and in vivo tumor progression of human glioblastoma cells. *J. Neuro-Oncol.* **2011**, *105*, 485–498.
68. Diermeier-Daucher, S.; Clarke, S. T.; Hill, D.; VollmannZwerenz, A.; Bradford, J. A.; Brockhoff, G. Cell type specific applicability of 5-ethynyl-2'-deoxyuridine (EdU) for dynamic proliferation assessment in flow cytometry. *Cytometry, Part A*, **2009**, *75a*, 535–546.
69. Neef, A. B.; Luedtke, N. W. Dynamic metabolic labeling of DNA in vivo with arabinosyl nucleosides. *Proc. Natl. Acad. Sci. U. S. A.*, **2011**, *108*, 20404–20409.
70. Qu, D.; Wang, G.; Wang, Z.; Zhou, L.; Chi, W.; Cong, S.; Ren, X.; Liang, P.; Zhang, B. 5-Ethynyl-2'-deoxycytidine as a new agent for DNA labeling: Detection of proliferating cells. *Anal. Biochem.* **2011**, *417*, 112–121.
71. Neef, A. B.; Samain, F.; Luedtke, N. W. Metabolic labeling of DNA by purine analogues in vivo. *ChemBioChem*, **2012**, *13*, 1750–1753.
72. Neef, A. B.; Luedtke, N. W. An azide-modified nucleoside for metabolic labeling of DNA. *ChemBioChem*, **2014**, *15*, 789–793.
73. Triemer, T.; Messikommer, A.; Glasauer, S. M. K.; Alzeer, J.; Paulisch, M. H.; Luedtke, N. W. *Proc. Natl. Acad. Sci. U. S. A.*, **2018**, *115*, E1366–E1373.
74. Rieder, U.; Luedtke, N. W. Alkene–tetrazine ligation for imaging cellular DNA. *Angew. Chem., Int. Ed.*, **2014**, *53*, 9168–9172.
75. Gubu, A.; Li, L.; Ning, Y.; Zhang, X. Y.; Lee, S.; Feng, M. K.; Li, Q.; Lei, X. G.; Jo, K. B.; Tang, X. J. Bioorthogonal Metabolic DNA Labelling using Vinyl Thioether-Modified Thymidine and o-Quinolinone Quinone Methide. *Chem. – Eur. J.*, **2018**, *24*, 5895–5900.
76. Lander, E. S.; Linton, L. M.; Birren, B.; Nusbaum, C.; Zody, M. C.; Baldwin, J.; et al. Initial sequencing and analysis of the human genome. *Nature.*, **2001**, *409*, 860–921.
77. Venter, J. C.; Adams, M. D.; Myers, E. W.; Li, P. W.; Mural, R. J.; Sutton, G. G.; et al. The sequence of the human genome. *Science.* **2001**, *291*, 1304–1351.
78. Chen, F.; Dong, M.; Ge, M.; Zhu, L.; Ren, L.; Liu, G.; Mu, R. The history and advances of reversible terminators used in new generations of sequencing technology. *Genom. Proteom. Bioinf.* **2013**, *11*, 34–40.
79. Goodwin, S.; McPherson, J. D.; McCombie, W. R. Coming of age: ten years of next-generation sequencing technologies. *Nat. Rev. Genet.* **2016**, *17*, 333–351.

80. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, Hall KP, Evers DJ, Barnes CL, Bignell HR et al.: Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **2008**, 456, 53-59.
81. Chen F, Gaucher EA, Leal NA, Hutter D, Havemann SA, Govindarajan S, Ortlund EA, Benner SA: Reconstructed evolutionary adaptive paths give polymerases accepting reversible terminators for sequencing and SNP detection. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, 107, 1948-1953.
82. Gardner, A. F.; Wang, J.; Wu, W.; Karouby, J.; Li, H.; Stupi, B. P.; Jack, W. E.; Hersh, M. N.; Metzker, M. L. Rapid incorporation kinetics and improved fidelity of a novel class of 3'-OH unblocked reversible terminators. *Nucleic Acids Res.* **2012**, 40, 7404–7415.
83. Bollum, F. J. Chemically Defined Templates and Initiators for Deoxypolynucleotide Synthesis. *Science*. **1964**, 144, 560.
84. Fowler, J. D.; Suo, Z. Biochemical, Structural, and Physiological Characterization of Terminal Deoxynucleotidyl Transferase. *Chem. Rev.* **2006**, 106, 2092–2110.
85. Motea, E. A.; Berdis, A. J.; Terminal deoxynucleotidyl transferase: the story of a misguided DNA polymerase. *Biochim Biophys Acta*. **2010**, 5, 1151–1166.
86. Winz, M. L.; Linder, E. C.; André, T.; Becker, J.; Jäschke, A. Nucleotidyl transferase assisted DNA labeling with different click chemistries. *NucleicAcidsRes.***2015**, 43, 1–10
87. Winz, M. L.; Samanta, A.; Benzinger, D.; Jäschke, A. Site-specific terminal and internal labeling of RNA by poly(A) polymerase tailing and copper-catalyzed or copper-free strain-promoted click chemistry. *NucleicAcidsRes.***2012**, 40, 1–13.
88. Röthlisberger, P.; Levi-Acobas, F.; Sarac, I.; Baron, B.; England, P.; Herdewijn, P.; Hollenstein, M. Facile immobilization of DNA using an enzymatic his-tag mimic. *Chem.Comm.***2017**, 53, 13031–13034.
89. Sarac, I.; Hollenstein, M. Terminal Deoxynucleotidyl Transferase in the Synthesis and Modification of Nucleic Acids. *ChemBioChem* **2019**, 20, 860–871.
90. Simonova, A.; Havran, L.; Pohl, R.; Fojta, M.; Hocek, M. Phenothiazine-linked nucleosides and nucleotides for redox labelling of DNA. *Org.Biomol.Chem.* **2017**, 15, 6984–6996.
91. Horáková, P.; MacÍková-Cahová, H.; Pivoková, H.; Paek, J.; Havran, L.; Hocek, M.; Fojta, M. ail-labelling of DNA probes using modified deoxynucleotide triphosphates and terminal deoxynucleotidyl tranferase. Application in electrochemical DNA hybridization and protein-DNA binding assays. *Org.Biomol.Chem.* **2011**, 9, 1366–1371.
92. Perkel, J. The race for enzymatic DNA synthesis heats up. *Nature* **2019**, 566, 9.
93. Palluk, S.; Arlow, D. H.; De Rond, T.; Barthel, S.; Kang, J. S.; Bector, R.; Baghdassarian, H. M.; Truong, A. N.; Kim, P. W.; Singh, A. K.; Hillson, N. J.; Keasling, J. D. De novo DNA synthesis using polymerase-nucleotide conjugates. *Nat.Biotechnol.* **2018**, 36, 645–650.
94. Mathews, A. S.; Yang, H.; Montemagno, C. Photo-cleavable nucleotides for primer free enzyme mediated DNA synthesis. *Org.Biomol.Chem.* **2016**, 14, 8278–8288.
95. Flamme, M.; Hanlon, S.; Iding, H.; Puentener, K.; Sladojevich, F.; Hollenstein, M. Evaluation of 3'-phosphate as a transient protecting group for controlled enzymatic synthesis of DNA and XNA oligonucleotides. *Bioorg. Med.Chem.Lett.* **2021**, 48, 12824.
96. Jang, E. K.; Son, R. G.; Pack, S. P. Novel enzymatic single-nucleotide modification of DNA oligomer: prevention of incessant incorporation of nucleotidyl transferase by ribonucleotide-borate complex. *Nucleic Acids Res.***2019**,47, E10.
97. Flamme, M.; Katkevica, D.; Pajuste, K.; Katkevics, M.; Sabat, N.; Hanlon, S.; Marzuoli, I.; Püntener, K.; Sladojevich, F.; Hollenstein, M.; Benzoyl and pivaloyl as efficient

- protecting groups for controlled enzymatic synthesis of DNA and XNA oligonucleotides. *Asian J. Org.Chem.* **2022**, *11*, e202200384.
98. Kropp, H. M.; Betz, K.; Wirth, J.; Diederichs, K.; Marx, A. Crystal structures of ternary complexes of archaeal B-family DNA polymerases. *PLoS One*. **2017**, *12*, No. e0188005.
99. Obeid, S.; Yulikov, M.; Jeschke, G.; Marx, A. Enzymatic Synthesis of Multiple Spin-Labeled DNA. *Angew. Chem., Int. Ed.* **2008**, *47*, 6782–6785.
100. Obeid, S.; Baccaro, A.; Welte, W.; Diederichs, K.; Marx, A. Structural basis for the synthesis of nucleobase modified DNA by *Thermus aquaticus* DNA polymerase. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 21327–21331.
101. Obeid, S.; Bußkamp, H.; Welte, W.; Diederichs, K.; Marx, A. Interactions of non-polar and 'Click-able' nucleotides in the confines of a DNA polymerase active site. *Chem. Commun.* **2012**, *48*, 8320–8322.
102. Kropp, H. M.; Diederichs, K.; Marx, A. The Structure of an Archaeal B-Family DNA Polymerase in Complex with a Chemically Modified Nucleotide. *Angew. Chem. Int. Ed.* **2019**, *58*, 5457–5461.
103. Jackson, L. N.; Chim, N.; Shi, C.; Chaput, J. C. Crystal structures of a natural DNA polymerase that functions as an XNA reverse transcriptase. *Nucleic Acids Res.* **2019**, *47*, 6973–6983.
104. Lenz, T.; Bonnist, E. Y. M.; Pljevaljčić, G.; Neely, R. K.; Dryden, D. T. F.; Scheidig, A. J.; Jones, A. C.; Weinhold, E. 2-Aminopurine flipped into the active site of the adenine-specific DNA methyl transferase M.TaqI: crystal structures and time-resolved fluorescence. *J. Am. Chem. Soc.* **2007**, *129*, 6240–624.
105. Dolot, R.; Lam, C. H.; Sierant, M.; Zhao, Q.; Liu, F-W.; Nawrot, B.; Egli, M.; Yang, X. Crystal structures of thrombin in complex with chemically modified thrombin DNA aptamers reveal the origins of enhanced affinity. *Nucleic Acids Res.* **2018**, *46*, 4819–4830.
106. Kellenbach, E. R.; Remerowski, M. L.; Eib, D.; Boelens, R.; van der Marel, G. A.; van den Elst, H.; van Boom, J. H.; Kaptein, R. Synthesis of isotope labeled oligonucleotides and their use in an NMR study of a protein-DNA complex. *Nucleic Acids Res.* **1992**, *20*, 653–657.
107. Szpotkowski, K.; Wojcik, K.; Kurzynska-Kokorniak, A. Structural studies of protein–nucleic acid complexes: A brief overview of the selected techniques. *Comput. Struct. Biotechnol. J.* **2023**, *21*, 2858–2872.
108. Dayie, T. K.; Olenginski, L. T.; Taiwo, K. M. Isotope Labels Combined with Solution NMR Spectroscopy Make Visible the Invisible Conformations of Small-to-Large RNAs. *Chem. Rev.* **2022**, *122*, 9357–9394.
109. Scott, L.G.; Tolbert, T. J.; Williamson, J. R. Preparation of specifically ^2H - and ^{13}C labeled ribonucleotides. *RNA-Ligand Interact. A* **2000**, *317*, 18–38.
110. Schultheisz, H. L.; Szymczyna, B. R.; Scott, L. G.; Williamson, J. R. Pathway Engineered Enzymatic de Novo Purine Nucleotide Synthesis. *ACS Chem. Biol.* **2008**, *3*, 499–511.
111. Nelissen, F.H.T.; Tessari, M.; Wijmenga, S. S.; Heus, H. A. Stable isotope labeling methods for DNA. *Progress in NMR Spectroscopy*. **2016**, *96*, 89–108.
112. Huang, K. Louis, J. M.; Donaldson, L.; Lim, F. L.; Sharrocks, A. D.; Clore, G. M. Solution structure of the MEF2A-DNA complex: structural basis for the modulation of DNA bending and specificity by MADS-box transcription factors. *EMBO J.* **2000**, *19*, 2615–2628.
113. Wojciak, J.M. Iwahara, J. Clubb, R.T. The Mu repressor-DNA complex contains an immobilized 'wing' within the minor groove. *Nat. Struct. Biol.* **2001**, *8*, 84–90.

114. C. Dominguez, M. Schubert, O. Duss, S. Ravindranathan, F.H.T. Allain, Structure determination and dynamics of protein–RNA complexes by NMR spectroscopy, *Prog. Nucl. Magn. Reson. Spectrosc.* **2011**, *58*, 1–61.
115. Wu, H.; Finger, L.D.; Feigon, J. Structure determination of protein/RNA complexes by NMR, *Methods Enzymol.* **2005**, *394*, 525–545.
116. Chen, H.; Viel, S.; Ziarelli, F.; Peng, L. ^{19}F NMR: a valuable tool for studying biological events. *Chem. Soc. Rev.* **2013**, *42*, 7971–7982.
117. Bao, H.-L.; Ishizuka, T.; Sakamoto, T.; Fujimoto, K.; Uechi, T.; Kenmochi, N.; Xu, Y. Characterization of human telomere RNA G-quadruplex structures in vitro and in living cells using ^{19}F NMR spectroscopy. *Nucleic Acids Res.* **2017**, *45*, 5501–5511.
118. Ishizuka, T.; Zhao, P.-Y.; Bao, H.-L.; Xu, Y. Oligonucleotides DNA containing 8-trifluoromethyl-2'-deoxyguanosine for observing Z-DNA structure. *Analyst* **2017**, *142*, 4083–4088.
119. Himmelstoß, M.; Erhardter, K.; Renard, E.; Ennifar, E.; Kreutz, C.; Micura, R. 2'-O-Trifluoromethylated RNA—a powerful modification for RNA chemistry and NMR spectroscopy. *Chem. Sci.* **2020**, *11*, 11322–11330.
120. Li, Q.; Chen, J.; Trajkovski, M.; Zhou, Y.; Fan, C.; Lu, K.; Tang, P.; Su, X.; Plavec, J.; Xi, Z.; Zhou, C. 4'-Fluorinated RNA: Synthesis, Structure, and Applications as a Sensitive ^{19}F NMR Probe of RNA Structure and Function. *J. Am. Chem. Soc.* **2020**, *142*, 4739–4748.
121. Bao, H.-L.; Liu, H.-S.; Xu, Y. Hybrid-type and two-tetrad antiparallel telomere DNA G-quadruplex structures in living human cells. *Nucleic Acids Res.*, **2019**, *47*, 4940–4947.
122. Baranowski, M. R.; Warminski, M.; Jemielity, J.; Kowalska, J.; 5'-fluoro(di)phosphate-labeled oligonucleotides are versatile molecular probes for studying nucleic acid secondary structure and interactions by ^{19}F NMR. *Nucleic Acids Res.* **2020**, *48*, 8209–8224.
123. Manna, S.; Sarkar, D.; Srivatsan, S. G. A Dual-App Nucleoside Probe Provides Structural Insights into the Human Telomeric Overhang in Live Cells. *J. Am. Chem. Soc.* **2018**, *140*, 12622–12633.
124. Khatik, S. Y.; Srivatsan, S. G. Environment-Sensitive Nucleoside Probe Unravels the Complex Structural Dynamics of i-Motif DNAs. *Bioconjugate Chem.* **2022**, *33*, 1515–1526.
125. Fauster, K.; Kreutz, C.; Micura, R. 2'-SCF₃ Uridine—A Powerful Label for Probing Structure and Function of RNA by ^{19}F NMR Spectroscopy. *Angew. Chem. Int. Ed.* **2012**, *51*, 13080–13084.
126. Olszewska, A.; Pohl, R.; Hocek, M. Trifluoroacetophenone-Linked Nucleotides and DNA for Studying of DNA–Protein Interactions by ^{19}F NMR Spectroscopy. *J. Org. Chem.* **2017**, *82*, 11431–11439.
127. Bao, H.-L.; Ishizuka, T.; Iwanami, A.; Oyoshi, T.; Xu, Y. A Simple and Sensitive ^{19}F NMR Approach for Studying the Interaction of RNA G-Quadruplex with Ligand Molecule and Protein. *ChemistrySelect.* **2017**, *2*, 4170–4175.
128. Dow, B. J.; Malik, S. S.; Drohat, A. C. Defining the Role of Nucleotide Flipping in Enzyme Specificity Using ^{19}F NMR. *J. Am. Chem. Soc.* **2019**, *141*, 4952–4962.
129. Brown, M. P.; Royer, C. Fluorescence spectroscopy as a tool to investigate protein interactions. *Curr. Opin. Biotechnol.* **1997**, *8*, 45–49.
130. Sinkeldam, R. W.; Greco, N. J.; Tor, Y. Fluorescent analogs of bio molecular building blocks: design, properties, and applications. *Chem. Rev.* **2010**, *110*, 2579–2619.
131. Fernández-Sierra, M.; Quiñones, E. Assays for the determination of the activity of DNA nucleases based on the fluorometric properties of the YOYO dye. *Arch. Biochem. Biophys.* **2015**, *570*, 40–46.

132. Öz, R.; Kk, S.; Westerlund, F. A nanofluidic device for real-time visualization of DNA–protein interactions on the single DNA molecule level *Nanoscale*, **2019**, *11*, 2071–2078.
133. Tripathy, D. R.; Dinda, A. K.; Dasgupta, S. A simple assay for the ribonuclease activity of ribonucleases in the presence of ethidium bromide *Anal. Biochem.* **2013**, *437*, 126–129.
134. Zhang, X.; Cheng, R.; Shi, Z.; Jin, Y. Label-free detection of EcoRI endonuclease activity via fluorescent DNA logic gate *Sensors Actuators B*, **2001**, *244* 387–92.
135. Pallares, R. M.; Sutarlie, L.; Thanh, N. T. K.; Su, X. Fluorescence sensing of protein–DNA interactions using conjugated polymers and graphene oxide *Sensors Actuators B*, **2018**, *271* 97–103.
136. Sha, L.; Zhang, X.; Wang, G. A label-free and enzyme free ultra-sensitive transcription factors biosensor using DNA-templated copper nanoparticles as fluorescent indicator and hairpin DNA cascade reaction as signal amplifier. *Biosens. Bioelectron.* **2016**, *82*, 85–92.
137. Ranasinghe, R. T.; Brown, T. Fluorescence based strategies for genetic analysis. *Chem. Commun. (Cam bridge, U.K.)* **2005**, 5487–5502.
138. Mart, A. A.; Jockusch, S.; Stevens, N.; Ju, J.; Turro, N. J. Fluorescent hybridization probes for sensitive and selective DNA and RNA detection. *Acc. Chem. Res.* **2007**, *40*, 402–409.
139. Davies, M. J.; Shah, A.; Bruce, I. J. Synthesis of fluorescently labelled oligonucleotides and nucleic acids *Chem. Soc. Rev.* **2000**, *29*, 97–107.
140. Dai, N.; Kool, E. T. Fluorescent DNA-based enzyme sensors *Chem. Soc. Rev.* **2011**, *40*, 5756–5770.
141. Heid, C. A.; Stevens, J.; Livak, K. J.; Williams, P. M. Real time quantitative PCR. *Genome Res.* **1996**, *6*, 986–994.
142. Tan, W.; Wangand, K.; Drake, T. J. Molecular beacons. *Curr. Opin. Chem. Biol.* **2004**, *8*, 547–553.
143. Li, Y.; Zhou, X.; Ye, D. Molecular beacons: An optimal multifunctional biological probe. *Biochem. Biophys. Res. Commun.* **2008**, *373*, 457–461.
144. Parkhurst, K. M.; Brenowitz, M.; Parkhurst, L. J. Simultaneous binding and bending of promoter DNA by the TATA binding protein: real time kinetic measurements. *Biochemistry.* **1996**, *357*, 459–465.
145. Li, J. J.; Fang, X.; Schuster, S. M.; Tan, W. Molecular beacons: a novel approach to detect protein–DNA interactions. *Angew. Chem. Int. Ed.* **2000**, *39*, 1049–1052.
146. Hamaguchi, N.; Ellington, A.; Stanton, M. Aptamer beacons for the direct detection of proteins. *Anal. Biochem.* **2001**, *294*, 126–131.
147. Wilson, J. N.; Kool, E. T. Fluorescent DNA base replacements: reporters and sensors for biological systems. *Org. Biomol. Chem.* **2006**, *4*, 4265–4274.
148. Secrist III, J. A.; Barrio, J. R.; Leonard, N. J.; Weber, G. Fluorescent modification of adenosine-containing coenzymes. Biological activities and spectroscopic properties. *Biochemistry.* **1972**, *11*, 3499–3506.
149. Gao, J., Liu, H., and Kool, E. T. Expanded-size bases in naturally sized DNA: evaluation of steric effects in Watson–Crick pairing. *J. Am. Chem. Soc.* **2004**, *126*, 11826–11831.
150. Liu, H.; Gao, J.; Maynard, L.; Saito, Y. D.; Kool, E. T. Toward a new genetic system with expanded dimensions: size-expanded analogues of deoxyadenosine and thymidine. *J. Am. Chem. Soc.* **2004**, *126*, 1102–1109.
151. Sandin, P.; Stengel, G.; Ljungdahl, T.; Borjesson, K.; Macao, B.; Wilhelmsson, L. M. Highly efficient incorporation of the fluorescent nucleotide analogs tC and tCO by Klenow fragment. *Nucleic Acids Res.* **2009**, *37*, 3924–3933.

152. Tahmassebi, D. C.; Millar, D. P. Fluorophore–quencher pair for monitoring protein motion. *Biochem. Biophys. Res. Commun.* **2009**, *380*, 277–280.
153. Lu, H.; Krueger, A. T.; Gao, J.; Liu, H.; Kool, E. T. Toward a designed genetic system with biochemical function: polymerase synthesis of single and multiple size-expanded DNA base pairs. *Org. Biomol. Chem.* **2010**, *8*, 2704–2710.
154. Jarchow-Choy, S. K.; Krueger, A. T.; Liu, H.; Gao, J.; Kool, E. T. Fluorescent xDNA nucleotides as efficient substrates for a template-independent polymerase. *Nucleic Acids Res.* **2011**, *39*, 1586–1594.
155. Pals, M. J.; Velema, W. A. A quenched size-expanded nucleotide reports activity of the leukemia biomarker terminal deoxynucleotidyl transferase (TdT). *Angew. Chem. Int. Ed.* **2023**, *62*, e202302796.
156. Sinkeldam, R. W.; Greco, N. J.; Tor, Y. Fluorescent analogs of biomolecular building blocks: design, properties, and applications. *Chem. Rev.* **2010**, *110*, 2579 – 2619.
157. Wilson, J. N.; Kool, E. T. Fluorescent DNA base replacements: reporters and sensors for biological systems. *Org. Biomol. Chem.* **2006**, *4*, 4265 – 4274.
158. Haidekker, M. A.; Theodorakis, E. A. Environment sensitive behavior of fluorescent molecular rotors. *J. Biol. Eng.* **2010**, *4*, 11
159. Kuimova, M. K. Mapping viscosity in cells using molecular rotors. *Phys. Chem. Chem. Phys.* **2012**, *14*, 12671
160. Vyšniauskas, A.; Kuimova, M. K. A twisted tale: measuring viscosity and temperature of microenvironments using molecular rotors. *Int. Rev. Phys. Chem.* **2018**, *37*, 259–85.
161. Ma, C.; Sun, W.; Xu, L.; Qian, Y.; Dai, J.; Zhong, G.; Hou, Y.; Liu, J.; Shen, B. A minireview of viscosity-sensitive fluorescent probes: design and biological applications. *J. Mater. Chem. B.* **2020**, *89*, 642–651.
162. Sinkeldam, R. W.; Wheat, A. J.; Boyaci, H.; Tor, Y. Emissive nucleosides as molecular rotors. *ChemPhysChem.* **2011**, *12*, 567–570.
163. Riedl, J.; Ménová, P.; Pohl, R.; Orság, P.; Fojta, M.; Hocek, M. GFP-like fluorophores as DNA labels for studying DNA–protein interactions. *J. Org. Chem.* **2012**, *77*, 8287–8293.
164. Kuba, M.; Pohl, R.; Hocek, M. Synthesis of 2'-deoxycytidine and its triphosphate bearing tryptophan-based imidazolinone fluorophore for environment sensitive fluorescent labelling of DNA. *Tetrahedron.* **2018**, *74*, 6621–6629.
165. Dziuba, D.; Pohl, R.; Hocek, M. Polymerase synthesis of DNA labelled with benzylidene cyanoacetamide-based fluorescent molecular rotors: fluorescent light-up probes for DNA-binding proteins. *Chem. Commun.* **2015**, *51*, 4880–4882.
166. Ravi Kumara, G. S.; Seo, Y. J. Polymerase-mediated synthesis of p-vinylaniline-coupled fluorescent DNA for the sensing of nucleolin protein– c-mycG-quadruplex interactions *Org. Biomol. Chem.* **2021**, *19*, 5788–93.
167. Dziuba, D.; Pohl, R.; Hocek, M. Bodipy-labeled nucleoside triphosphates for polymerase synthesis of fluorescent DNA *Bioconjugate Chem.* **2014**, *25*, 1984–1995.
168. Dziuba, D.; Jurkiewicz, P.; Cebecauer, M.; Hof, M.; Hocek, M. A Rotational BODIPY nucleotide: an environment-sensitive fluorescence-lifetime probe for DNA interactions and applications in live-cell microscopy *Angew. Chem. Int. Ed.* **2016**, *55*, 174–178.
169. Allen, D. J.; Benkovic, S. J. Resonance energy transfer measurements between substrate binding sites within the large (Klenow) fragment of Escherichia coli DNA polymerase I *Biochemistry.* **1989**, *28*, 9586–9593.
170. Guest, C. R.; Hochstrasser, R. A.; Allen, D. J.; Benkovic, S. J.; Millar, D. P.; Dupuy, C. G. Interaction of DNA with the Klenow fragment of DNA polymerase I studied by time-resolved fluorescence spectroscopy. *Biochemistry.* **1991**, *30*, 8759–8770.

171. Riedl, J.; Pohl, R.; Ernsting, N. P.; Ors  g, P.; Fojta, M.; Hocek, M. Labelling of nucleosides and oligonucleotides by solvatochromic 4-aminophthalimide fluorophore for studying DNA–protein interactions. *Chem. Sci.* **2012**, 3, 2797.
172. Koo, C.-K.; Samain, F.; Dai, N.; Kool, E. T. DNA polyfluorophores as highly diverse chemosensors of toxic gases. *Chem. Sci.* **2011**, 2, 1910 – 1917.
173. Tan, S. S.; Kim, S. J.; Kool, E. T. Differentiating between fluorescence-quenching metal ions with polyfluorophore sensors built on a DNA Backbone. *J. Am. Chem. Soc.* **2011**, 133, 2664 – 2671.
174. Guo, J.; Wang, S.; Dai, N.; Teo, Y. N.; Kool, E. T. Multispectral labeling of antibodies with poly fluorophores on a DNA backbone and application in cellular imaging. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, 108, 3493 – 3498.
175. Wang, S.; Guo, J.; Ono, T.; Kool, E. T. DNA polyfluorophores for real-time multicolor tracking of dynamic biological systems. *Angew. Chem. Int. Ed.* **2012**, 51, 7176 –7180.
176. Beharry, A. A.; Lacoste, S.; O’Connor, T. R.; Kool, E. T. Fluorescence monitoring of the oxidative repair of DNA alkylation damage by ALKBH3, a prostate cancer marker. *J. Am. Chem. Soc.* **2016**, 138, 3647–3650.
177. Coleman, R. S.; Madaras, M. L. Synthesis of a novel coumarin C-riboside as a photophysical probe of oligonucleotide dynamics. *J. Org. Chem.* **1998**, 63, 5700–5703.
178. Sen, S.; Paraggio, N. A.; Gearheart, L. A. Connor, E. E.; Issa, A.; Coleman R. S.; Wilson, D. M.; Wyatt, M. D.; Berg, M. A. Effect of protein binding on ultrafast DNA dynamics: characterization of a DNA: APE1 complex *Biophys. J.* **2005**, 89 4129–4138.
179. Myers, J. C.; Moore, S. A.; Shamoo, Y.; Structure-based incorporation of 6-methyl-8-(2-deoxy- β -ribofuranosyl) isoxanthopteridine into the human telomeric repeat DNA as a probe for UP1 binding and destabilization of G-tetrad structures *J. Biol. Chem.* **2003**, 278 42300–42306.
180. Rist, M. J.; Marino, J. P. Fluorescent nucleotide base analogs as probes of nucleic acid structure, dynamics and interactions. *Curr. Org. Chem.* **2002**, 6, 775–793.
181. Sinkeldam, R. W.; Greco, N. J.; Tor, Y. Fluorescent Analogs of Biomolecular Building Blocks: Design, Properties, and Applications. *Chem. Rev.* **2010**, 110, 2579–2619.
182. Srivatsan, S. G.; Sawant, A. A. Fluorescent ribonucleoside analogues as probes for investigating RNA structure and function. *Pure Appl. Chem.* **2011**, 83, 213–232.
183. Arndt, J. W.; Gong, W.; Zhong, X.; Showalter, A. K.; Liu, J.; Dunlap, C. A.; Lin, Z.; Paxson, C.; Tsai, M. D.; Chan, M. K. Insight into the catalytic mechanism of DNA polymerase, structures of intermediate complexes. *Biochemistry*, **2001**, 40, 5368–5375.
184. Shah, A. M.; Li, S. X.; Anderson, K. S.; Sweasy, J. B. Y265H mutator mutant of DNA polymerase β . *J. Biol. Chem.* **2001**, 276, 10824–10831.
185. da Silva, E. F.; Mandal, S. S.; Reha-Krantz, L. J. Using 2-Aminopurine fluorescence to measure incorporation of incorrect nucleotides by wild type and mutant bacteriophage T4 DNA polymerases. *J. Biol. Chem.*, **2002**, 277, 40640–40649.
186. Purohit, V.; Grindley, N. D. F.; Joyce, C. M. Use of 2 amino purine fluorescence to examine conformational changes during nucleotide incorporation by DNA polymerase I (Klenowfragment). *Biochemistry*, **2003**, 42 10200–11.
187. Dunlap, C. A.; Tsai, M-D. Use of 2-aminopurine and tryptophan fluorescence as probes in kinetic analyses of DNA polymerase β *Biochemistry*, **2002**, 41, 11226–11235.
188. Frey, M. W.; Sowers, L. C.; Millar, D. P.; Benkovic, S. J. The nucleotide analog 2-aminopurine as a spectroscopic probe of nucleotide incorporation by the Klenow fragment of Escherichia coli polymerase I and bacteriophage T4 DNA polymerase. *Biochemistry*, **1995**, 349, 185–192.
189. Beechem, J. M.; Otto, M. R.; Bloom, L. B.; Eritja, R.; Reha-Krantz, L. J.; Goodman, M. F. Exonuclease–polymerase active site partitioning of primer–template DNA

- strands and equilibrium Mg^{2+} binding properties of bacteriophage T4 DNA polymerase. *Biochemistry* **1998**, 37, 10144–10155.
190. Neely, R. K. Time-resolved fluorescence of 2-amino purine as a probe of base flipping in M.HhaI-DNA complexes. *Nucleic Acids Res.* **2005**, 33, 6953–6960.
191. Lenz, T.; Bonnist, E. Y. M.; Pljevaljčić, G.; Neely, R. K.; Dryden, D. T. F.; Scheidig, A.J.; Jones, A.C.; Weinhold, E. 2-Aminopurine flipped into the active site of the adenine specific DNA methyl transferase M.TaqI: crystal structure and time-resolved fluorescence. *J. Am. Chem. Soc.* **2007**, 129, 6240–6248.
192. Ma, L.; Wu X.; Wilson G.; Jones, A. C.; Dryden, D.T. F. Time-resolved fluorescence of 2-amino purine in DNA duplexes in the presence of the EcoP15I Type III restriction modification enzyme. *Biochem. Biophys. Res. Commun.* **2014**, 449, 120–125.
193. Katilius, E.; Katilene, Z.; Woodbury, N. W. Signaling aptamers created using fluorescent nucleotide analogues. *Anal. Chem.* **2006**, 78, 6484–6489.
194. Stephens, O. M.; Yi-Brunozzi, H. Y.; Beal, P. A. Analysis of the RNA-editing reaction of ADAR2 with structural and fluorescent analogues of the GluR-B R/G editing site. *Biochemistry*, **2000**, 39, 12243–12251.
195. Yi-Brunozzi, H.Y.; Stephens, O. M.; Beal, P. A. Conformational Changes That Occur during an RNA-editing adenosine deamination reaction. *J. Biol. Chem.* **2001**, 276, 37827–37833.
196. Ward, D. C.; 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.
197. Kawai, M.; Lee, M. J.; Evans, K. O.; Nordlund, T. M. Temperature and base sequence dependence of 2-aminopurine fluorescence bands in single- and double-stranded oligodeoxynucleotides. *J. Fluoresc.* **2001**, 11, 23–32.
198. Rachofsky, E. L.; Osman, R.; Ross, J. B. A. Probing structure and dynamics of DNA with 2-Aminopurine: effects of local environment on fluorescence. *Biochemistry*. **2001**, 40, 946–956.
199. Dash, C.; Rausch, J. W.; Le Grice, S. F. Using pyrrolo-deoxycytosine to probe RNA/DNA hybrids containing the human immunodeficiency virus type-1 3'-polypurine tract. *Nucleic Acids Res.*, **2004**, 32, 1539–1547.
200. Liu, C.; Martin, C. T. Promoter clearance by T7 RNA polymerase. *J. Biol. Chem.*, **2002**, 277, 2725–2731.
201. 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.
202. Wahba, A.S.; Esmaeili, A.; Damha, M. J.; Hudson, R. H. E. A single-label phenyl pyrrolocytidine provides a molecular beacon-like response reporting HIV-1 RT RNase H activity. *Nucleic Acids Res.* **2010**, 38, 1048–1056.
203. Rovira, A. R.; Fin, A.; Tor, Y. Chemical mutagenesis of an emissive RNA alphabet. *J. Am. Chem. Soc.* **2015**, 137, 14602–14605.
204. Shin, D.; Sinkeldam, R. W.; Tor, Y. Emissive RNA alphabet. *J. Am. Chem. Soc.* **2011**, 133, 14912–14915.
205. Kuchlyan, J.; Martinez-Fernandez, L.; Mori, M.; Gavvala, K.; Ciaco, S.; Boudier, C.; Richert, Didier, L.; P.; Tor, Y.; Improta, R.; Mély, Y. What makes thienoguanosine an outstanding fluorescent DNA probe? *J. Am. Chem. Soc.* **2020**, 142, 16999–17014.
206. McCoy, L. S.; Shin, D.; Tor, Y. Isomorphic emissive GTP surrogate facilitates initiation and elongation of in vitro transcription reactions. *J. Am. Chem. Soc.*, **2014**, 136, 15176–15184.

207. Li, Y. Fin, A.; McCoy, L.; Tor, Y. Polymerase-mediated site-specific incorporation of a synthetic fluorescent isomorph G surrogate into RNA. *Angew. Chem., Int. Ed.*, **2017**, *56*, 1303–1307.
208. Otomo, H.; Park, S.; Yamamoto, S.; Sugiyama, H. Amplification of fluorescent DNA through enzymatic incorporation of a highly emissive deoxyguanosine analogue. *RSC Adv.*, **2014**, *4*, 31341–31344.
209. Sinkeldam, R. W.; McCoy, L. S.; Shin, D.; Tor, Y. Enzymatic interconversion of isomorph fluorescent nucleosides: adenosine deaminase transforms thA to thI, two distinct fluorophores. *Angew. Chem., Int. Ed.*, **2013**, *52*, 14026–14030.
210. Srivatsan, S. G.; Tor, Y. Fluorescent pyrimidine ribonucleotide: synthesis, enzymatic incorporation, and utilization. *J. Am. Chem. Soc.* **2007**, *129*, 2044 – 2053.
211. Srivatsan, S. G.; Tor, Y. Using an emissive uridine analogue for assembling fluorescent HIV-1 TAR constructs. *Tetrahedron* **2007**, *63*, 3601 – 3607.
212. Tanpure, A. A.; Srivatsan, S. G. A Microenvironment-sensitive fluorescent pyrimidine ribonucleoside analogue: synthesis, enzymatic incorporation, and fluorescence detection of a DNA Abasic site. *Chem. Eur. J.* **2011**, *17*, 12820–12827.
213. Pawar, M. G.; Srivatsan, S. G. Synthesis, photophysical characterization, and enzymatic incorporation of a microenvironment-sensitive fluorescent uridine analog. *Org. Lett.* **2011**, *13*, 1114–1117.
214. Manna, S.; Srivatsan, S. G. Synthesis and enzymatic incorporation of a responsive ribonucleoside probe that enables quantitative detection of metallo-base pairs. *Org. Lett.* **2019**, *21*, 4646–4650.
215. Nuthanakanti, A.; Boerneke, M. A.; Hermann, T.; Srivatsan S. G. Structure of the ribosomal decoding site RNA containing a Se-modified responsive fluorescent ribonucleoside probe. *Angew. Chem. Int. Ed.* **2017**, *56*, 2640–2644.
216. George, J. T.; Srivatsan, S. G. Responsive fluorescent nucleotides serve as efficient substrates to probe terminal uridylyl transferase. *Chem. Commun.* **2020**, *56*, 12307–12318.
217. Ghosh, P.; Kropp, H. M.; Betz, K.; Ludmann, S.; Diederichs, K.; Marx, A.; Srivatsan, S. G. Microenvironment-sensitive fluorescent nucleotide probes from benzofuran, benzothiophene and selenophene as substrates for DNA polymerases *J. Am. Chem. Soc.* **2022**, *144*, 10556 –10569.

Chapter 2

Microenvironment-sensitive fluorescent nucleotide probes from benzofuran, benzothiophene and selenophene as substrates for DNA polymerases*

2.1 Introduction

Enzymatic functionalization of nucleic acids with biophysical tools and probes is used in many biotechnological applications as well as in fundamental research of nucleic acid structure and function.¹⁻⁵ This process greatly relies on the ability of native and engineered DNA and RNA polymerases and transferases to incorporate functionalized nucleotides into oligonucleotides (ONs).⁶⁻¹⁷ Though unpredictable, base-modified nucleotide analogs, which do not affect the Watson-Crick base-pairing are reasonably well tolerated by these enzymes.^{18,19} For example, nucleotide analogs bearing fluorophores, redox-, NMR- and EPR-active labels,²⁰⁻²⁷ and nucleotides containing handles for postsynthetic modification are widely employed.²⁸⁻³² A notable number of analogs are also metabolically incorporated by endogenous enzymes, which has further enabled the profiling and visualization of endogenous nucleic acids in living cells.³³⁻³⁸ However, the mechanism by which polymerases accept unnatural nucleotides as substrates and incorporate them into the nascent nucleic acid chains is not yet fully understood, so prediction of substrate tolerance is often not possible. Nevertheless, “platforms” have been established by combining biochemical techniques and X-ray crystallography to gain functional and structural insights into the mechanism of acceptance and incorporation of certain modified nucleotides by DNA polymerases.^{18,39} Successful co-crystallization of a truncated version of *Thermus aquaticus* DNA polymerase, KlenTaq, in complex with primer•template duplex and 2'-deoxynucleoside triphosphate (dNTP) surrogates provided structural basis for the processivity of the DNA polymerase.¹⁸ Polar flexible modifications (e.g., aminopentynyl) and a nonpolar aromatic group (e.g., ethynylbenzene) attached to C5 of pyrimidines or C7 of 7-deazapurines *via* an alkyne bond tend to stabilize substrate-enzyme interaction and show good incorporation efficiency.^{40,41}

*The work in this chapter is published: Ghosh, P.; Kropp, H. M.; Betz, K.; Ludmann, S.; Diederichs, K.; Marx, A.; Srivatsan, S. G. Microenvironment-sensitive fluorescent nucleotide probes from benzofuran, benzothiophene, and selenophene as substrates for DNA polymerases. *J. Am. Chem. Soc.* **2022**, *144*(23), 10556–10569. doi:10.1021/jacs.2c03454 (American Chemical Society acknowledged).

In contrast, a rigid and nonpolar spin label connected to a nucleotide distorts the enzyme conformation and thus, its efficiency of incorporation.⁴² Though these studies are indicative, a comprehensive understanding of the chemical space that can be handled by nucleic acid processing enzymes is very important. In this context, as the next important step, we sought to investigate the ability of DNA polymerases to process heterocycle-modified environment-sensitive fluorescent nucleotide probes, because they represent an important class of biophysical tool box in the analysis of nucleic acid.⁴³⁻⁴⁶

Environment-sensitive fluorescent nucleoside analogs derived by conjugating or fusing heterocyclic groups to pyrimidine and purine bases are highly valuable in probing conformations, dynamics and recognition properties of nucleic acids.⁴⁷⁻⁵² A notable number of such responsive analogs are amicable to enzymatic incorporation by polymerases and are useful in studying nucleic acid lesions, nucleic acid-drug/protein binding and conformations of therapeutically relevant nucleic acid motifs including bacterial ribosomal decoding site and tetraplexes.⁵³⁻⁵⁷ While the incorporation efficiency by different polymerases vary, which is typically determined by gel electrophoresis, the structural basis by which the fluorescent nucleotide analogs are accepted and processed by different family of DNA polymerases has not been well studied. Here, our idea is to devise a platform by harnessing the inherent responsiveness of the fluorescent nucleotide substrates and X-ray crystallography technique to probe the mechanism of enzymatic incorporation. For this purpose, we chose three nucleobase modified 2'-deoxyuridine-5'-triphosphate substrates, each containing a benzofuran, benzothiophene or selenophene moiety, as the test system (Figure 1). Notably, these base-modified fluorescent nucleoside analogs are highly conformation sensitive and show great promise in probing the structure and ligand binding of therapeutic nucleic acid motifs.^{53,57-59} Characterizing the enzymatic incorporation of this set of fluorescent nucleotides has two advantages apart from understanding the role of expanded chemical space compared to the existing structural information—(i) the substrates itself could serve as probes providing real time fluorescence readout to study the events in the catalytic cycle under equilibrium conditions and (ii) the incorporation efficiency and binding data determined by biochemical and fluorescence assays could be directly correlated with the structure obtained by X-ray crystallography.

Here, we report the synthesis and incorporation of microenvironment-sensitive fluorescent nucleotide analogs, which are derived by conjugating benzofuran, benzothiophene and selenophene at the C5 position of 2'-deoxyuridine (Figure 1). The

nucleotide analogs are incorporated into ONs by KlenTaq (belonging to the sequence family A DNA polymerases), KOD and Vent(exo-) DNA polymerases (both belonging to the B family) with varying efficiencies depending on the size of the heterocycle modification. KOD and Vent(exo-) DNA polymerases tolerate these modifications well, and hence, are suited for synthesizing ONs with multiple labels. In contrast, KlenTaq exclusively adds a single nucleotide analog at the 3'-end, thereby enabling the development of a simple two-step site-specific labeling procedure by PEX (Figure 1). Further, we describe the mechanism of incorporation of heterocycle-modified nucleotides by different family of DNA polymerases, which was obtained by comparing the incorporation profile, complex formation monitored using the responsiveness of the nucleotide analog and X-ray structures of ternary complexes (enzyme-duplex-nucleotide) containing native and modified nucleotides.

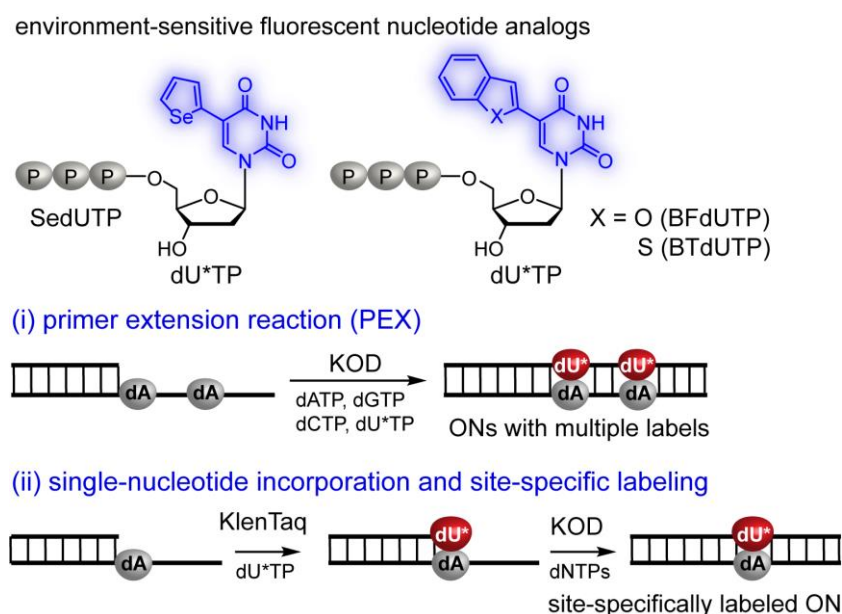


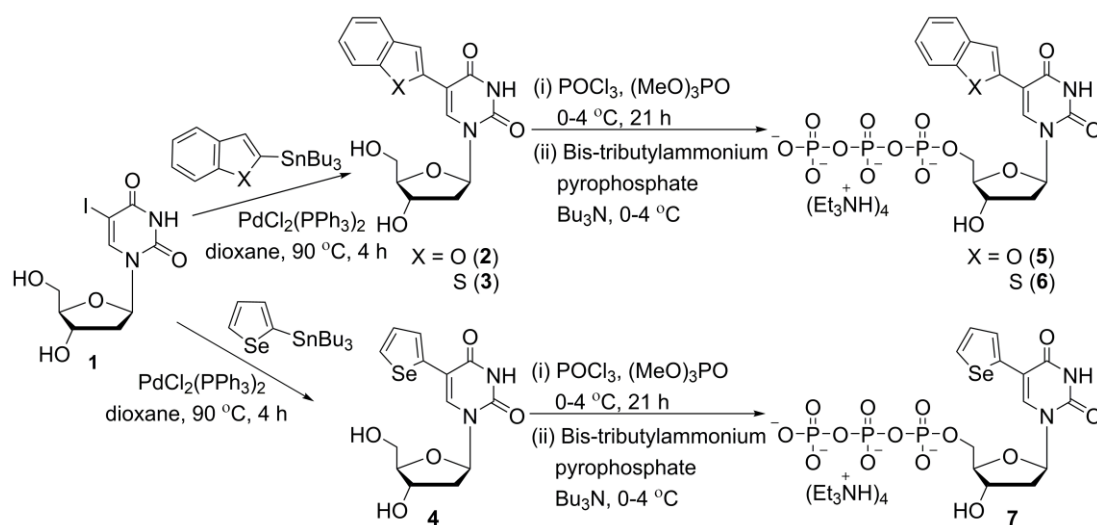
Figure 1. C5 heterocycle-conjugated fluorescent dU*TP analogs. The analogs are suitable for enzymatic functionalization of ONs by DNA polymerases. Depending on the enzyme choice and size of the heterocycle modification, ONs with (i) multiple labels and (ii) site-specific labels can be synthesized.

2.2 Result and discussions

2.2.1 Design and synthesis of fluorescent nucleotide analogs

We became interested in the C5 heterocycle-modified pyrimidine nucleotides 5–7 and their incorporation into ONs by DNA polymerases for the following reasons (Scheme 1). Conjugation of heterocycles at the C5 position of pyrimidines via a rotatable bond (molecular rotor element)^{60,61} generates conformation-sensitive fluorescent probes (Figure 1).^{25,43,46,62,63} Moreover, modifications at the C5 position of pyrimidines that can be accommodated in the

major groove of duplexes show good substrate properties in DNA polymerization reactions.^{3,18,64-72} Benzofuran (**2**), benzothiophene (**3**), and selenophene (**4**)-conjugated 2'-deoxyuridine analogs were synthesized by reacting 5-iodo-2'-deoxyuridine **1** with respective stannylated heterocycles under Stille cross-coupling reaction conditions using a palladium catalyst, $\text{Pd}(\text{PPh}_3)_2\text{Cl}_2$ (Scheme 1). These nucleoside analogs possess moderate quantum yields and their photophysical properties are highly sensitive to microenvironment polarity and viscosity changes.^{57,59,73} In particular, **2** and **4** incorporated into the human telomeric repeat ON by solid-phase synthesis enabled the conformational analysis and estimate the ligand binding ability of different G-quadruplex topologies by fluorescence and X-ray crystallography techniques.^{57,59} To probe the DNA polymerases activity, the analogs were further converted to nucleoside 5'-triphosphates **5**, **6** and **7** in a single-pot two-stage reaction using POCl_3 and bis-tributylammonium pyrophosphate (Scheme 1). The fluorescence profile, quantum yield and lifetime of triphosphates **5–7** were similar to respective nucleoside analogs (Figure 1 and Table 1).



Scheme 1. Synthesis of BFdUTP **5**, BTdUTP **6** and SedUTP **7**.

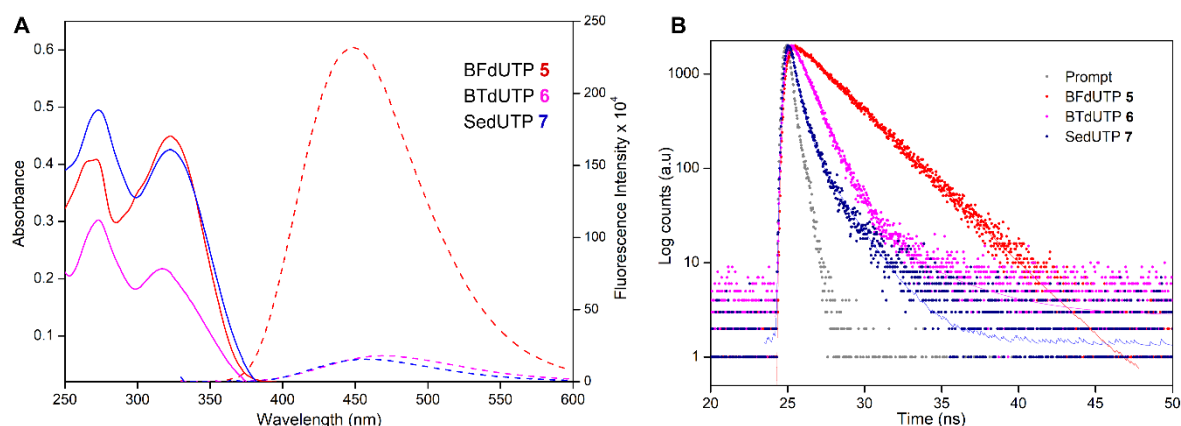


Figure 2. (A) Absorption (solid line, 25 μ M) and emission (dashed line, 5 μ M) spectra of nucleotides (5–7) in water. Emission spectrum was recorded by exciting the samples at respective absorption maximum (Table S1) with an excitation and emission slit width of 3 nm and 4 nm, respectively. (B) Excited-state decay profile of nucleotides (5–7, 5 μ M) in water. Laser profile (prompt) is shown in gray and curve fits are shown in solid lines. Fluorescence decay kinetics was performed with a time correlated single photon counting (TCSPC) set up (Horiba Delta Flex). Samples were excited using 340 nm delta diode (Horiba Scientific, DD-340). Fluorescence signal collected at the emission maximum was analyzed using EZ Time analysis software. The quality of curve fit was ascertained by χ^2 value, which was very close to unity for all fits.

Table 1. Photophysical properties of modified triphosphate 5–7 in water.

Nucleotides	λ_{max} (nm) ^a	λ_{em} (nm)	Φ ^b	τ_{av} ^b
BFdUTP 5	322	448	0.18	2.6021
BTdUTP 6	318	469	0.033	1.0751
SedUTP 7	324	458	0.013	0.6079

^aThe lowest energy absorption maximum is provided.

^bStandard deviations for quantum yield (Φ) and average lifetime (τ_{av}) are ≤ 0.01 and 0.04 ns, respectively.

2.2.2 DNA polymerases show varying incorporation efficiency

The potential of the three modified nucleotides to serve as a thymidine analog for DNA polymerases was tested in primer extension reactions (PEX) in the presence of commonly used A and B family DNA polymerases. A series of templates (**T1–T5**) was designed to guide the incorporation of the modified nucleotides at the PEX start site, distal from the start site, and at multiple positions (Table 2). A 5'-FAM-labeled primer **P1** was hybridized with the respective templates and the reactions were performed in the presence of dATP, dGTP, dCTP and dTTP/modified dU*TPs (5–7, Figure 3A). Reaction products were resolved by polyacrylamide gel electrophoresis (PAGE) under denaturing conditions and visualized using a gel fluorescence scanner. Reactions were first performed with **T1** and **T2** that would guide the incorporation of a single modified nucleotide analog at the start site and away from the

start site, respectively. Regardless of the size of modification, KlenTaq belonging to the A family efficiently incorporated the mononucleotides of 5–7 to produce full-length PEX products (Figure 3B, lanes 5–7 and lanes 10–12). The incorporation efficiency was as good as the reaction performed in the presence of natural dTTP (compare with lanes 4 and 9). When control reactions were performed in the absence of dTTP/modified dU*TPs but in the presence of other dNTPs, the enzyme did not produce full-length products (lanes 3 and 8). These results indicate that indeed the modified dU*TPs are incorporated specifically opposite to the templating dA and the full-length products are not formed due to misincorporation. However, with templates **T3** and **T4** that would introduce two modifications at alternate and consecutive sites, the incorporation efficiency was found to depend on the size of modification. While mononucleotides of benzofuran- and benzothiophene-modified dUTPs (**5** and **6**) were barely incorporated, selenophene-modified dUTP **7** was processed with an apparent efficiency as the natural nucleotide (Figure 3C, lanes 4–7 and lanes 9–12). When the stringency of incorporation was further increased by using **T5** (three modifications in close positions), KlenTaq did not produce full-length products with **5** and **6** (Figure 5, lanes 5 and 6). However, the enzyme tolerated the smaller selenophene modification and produced triply-modified ON product in very good yields (Figure 5, lane 7).

Table 2. ON sequences used in PEX, single nucleotide incorporation and crystallization studies, and fluorescence assays.^[a]

ONs	Sequence
P1	5' FAM-GTG GTG CGA AAT TTC TGA CAG ACA 3'
T1	3' CAC CAC GCT TTA AAG ACT GTC TGT ACT GTC TGC GTG 5'
T2	3' CAC CAC GCT TTA AAG ACT GTC TGT GCT GTC TAC GTG 5'
T3	3' CAC CAC GCT TTA AAG ACT GTC TGT GCT GT A TAC GTG 5'
T4	3' CAC CAC GCT TTA AAG ACT GTC TGT GCT GTC TAA GTG 5'
T5	3' CAC CAC GCT TTA AAG ACT GTC TGT GCT GT A TAA GTG 5'
P2	5' FAM-GGA GCT CAG CCT TCA CTG C 3'
P3	5' GGA GCT CAG CCT TCA CTG C 3'
T6	3' CCT CGA GTC GGA AGT GAC G ACG ATT TGA TAT AAC GAT CGC GTG 5'
P2*	5' FAM-GGA GCT CAG CCT TCA CTG C- BFdU 3'
P3*	5' GGA GCT CAG CCT TCA CTG C- BFdU 3'
T7	3' CCT CGA GTC GGA AGT GAC G ATG TCG CTC CGC GTC 5'
P4	5' GAC CAC GGC CA 3'
T8	3' CTG GTG CCG GT G ACAA 5'
P5	5' GGA GCT CAG CCT TCA CTG ddC 3'

^[a]ON primer **P1** and templates **T1–T5** were used in PEX. ONs **P2/P3** and **T6** were used in single-nucleotide incorporation reactions. **P2*** and **P3*** are 3'-BFdU-labeled primers **P2** and **P3**, respectively. **P3*** was used in fluorescence assays. Template **T7** was used in binding studies by fluorescence and EMSA. 2'-deoxyadenosines colored in red in template strands guide the incorporation of modified nucleotides analogs in the PEX and single

nucleotide incorporation assays. ONs **P4** and **T8** were used in crystallization trails. In **T8**, 2'-deoxyguanosine colored in blue is used for terminating the extension by using ddCTP in crystallization trails.

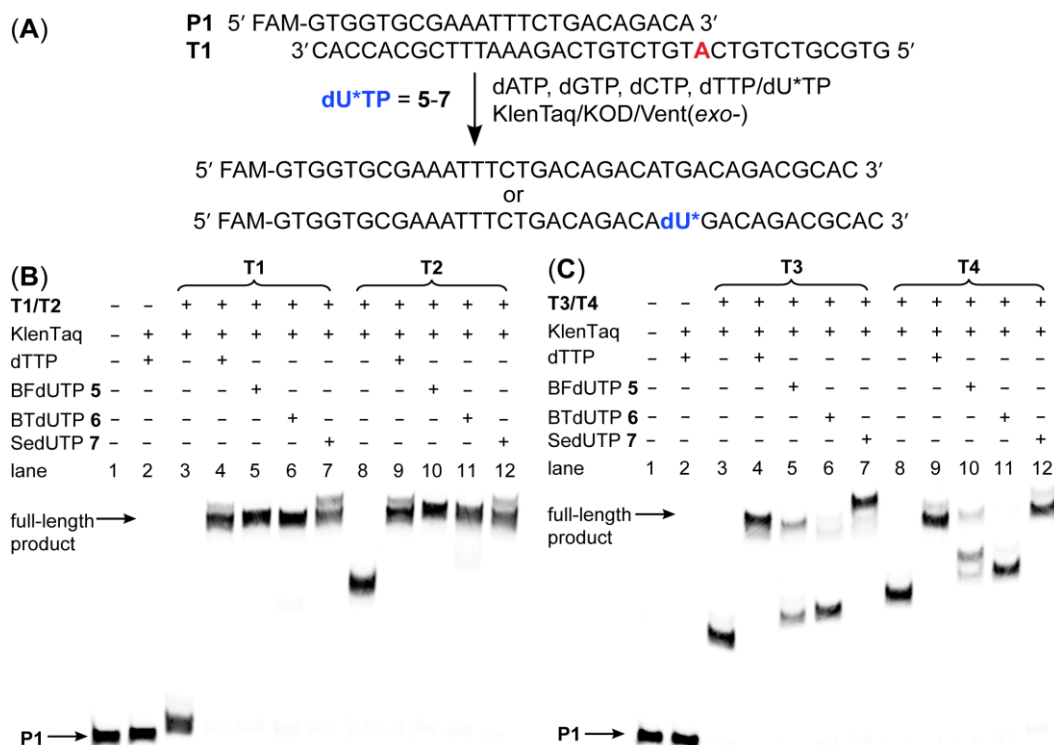


Figure 3. (A) PEX setup used to incorporate the mononucleotide analogs of 5–7 using different DNA polymerases. Reaction setup using **T1** is given as an example. (B and C) Incorporation of mononucleotides of BFdUTP 5, BTdUTP 6 and SedUTP 7 by KlenTaq using primer **P1** and templates **T1–T4**. All reactions were conducted in the presence of dATP, dGTP, dCTP.

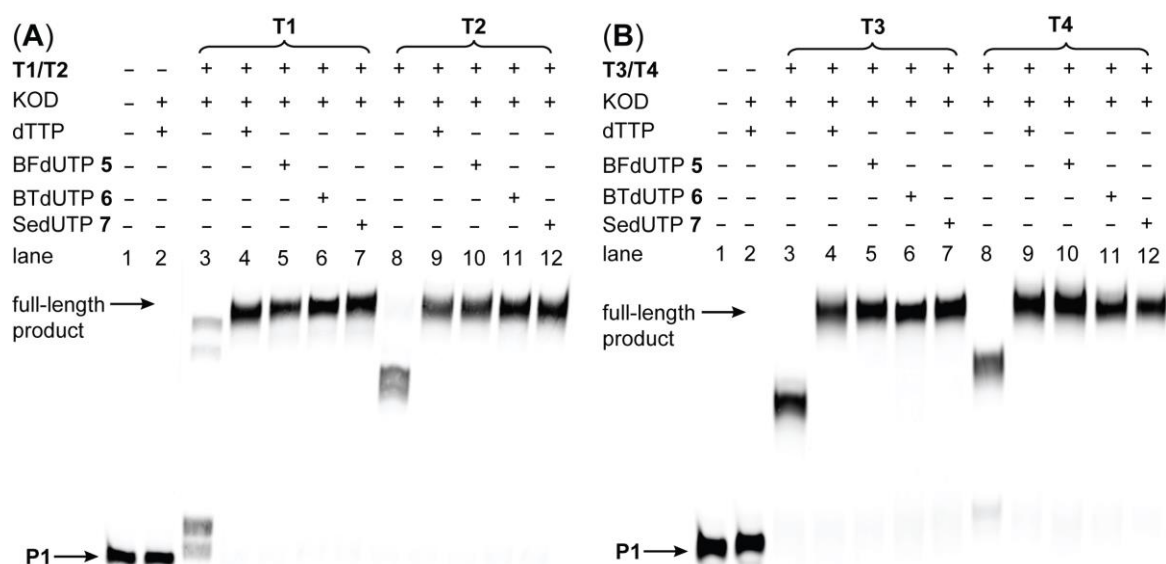


Figure 4. Incorporation of mononucleotides of BFdUTP 5, BTdUTP 6 and SedUTP 7 by KOD DNA polymerase using primer **P1** and templates **T1–T4**. All reactions were conducted in the presence of dATP, dGTP, dCTP.

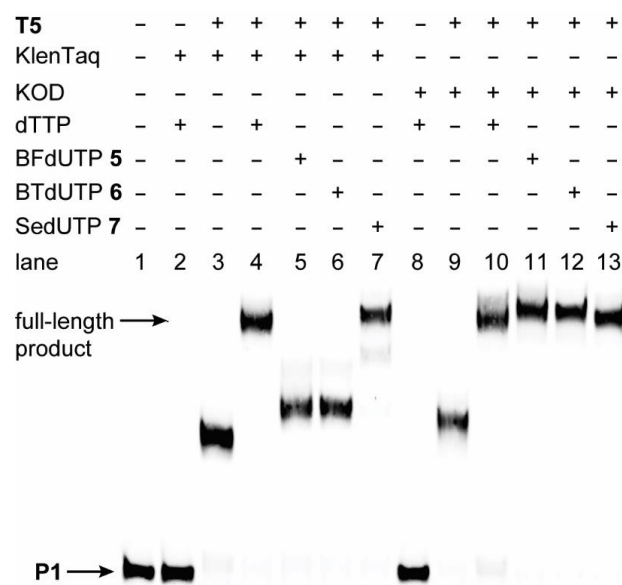


Figure 5. Incorporation of mononucleotides of BFdUTP 5, BTdUTP 6 and SedUTP 7 using KlenTaq and KOD DNA polymerases in the presence of duplex made of primer **P1** and template **T5**.

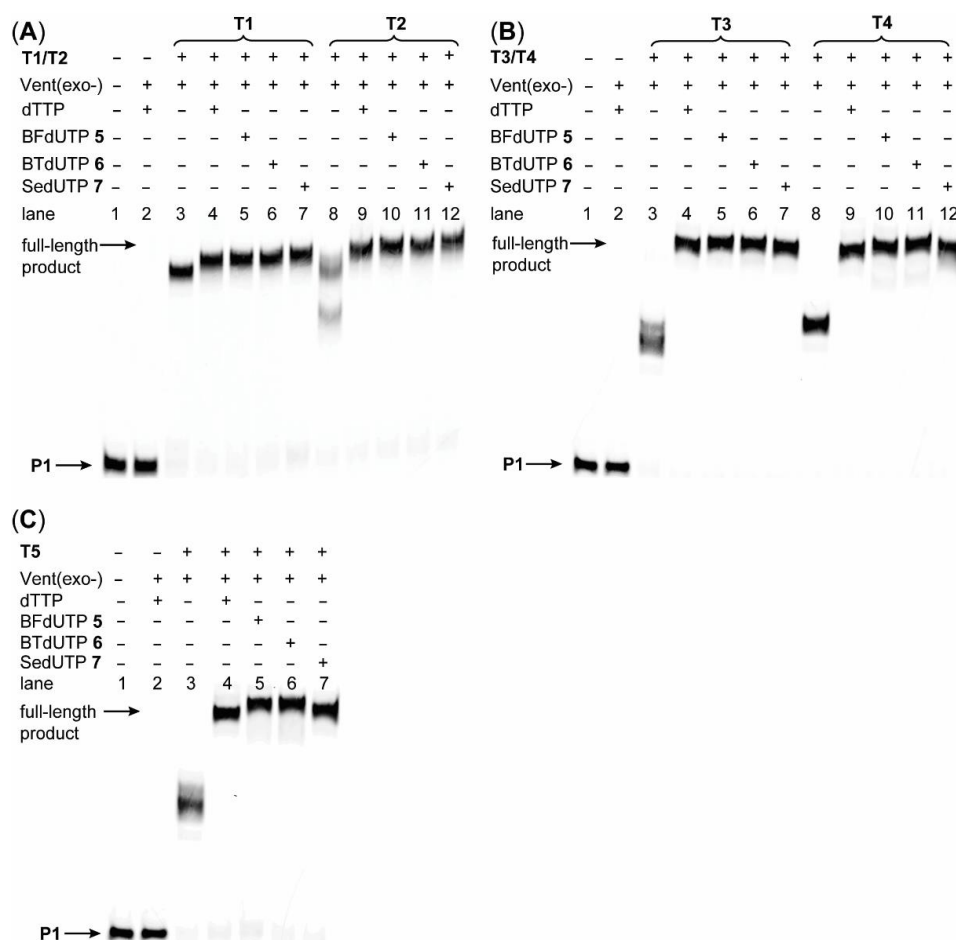


Figure 6. Incorporation of mononucleotides of BFdUTP 5, BTdUTP 6 and SedUTP 7 using Vent(exo-) in the presence of duplexes made of primer **P1** and templates **T1–T5**.

Next, the B family DNA polymerases were investigated. BFdUTP **5**, BTdUTP **6** and SedUTP **7** were processed proficiently by KOD using templates **T1–T5**, generating the full-length products with yields similar to the products formed in the presence of native dNTPs (Figure 4 and Figure 5, lanes 10–13). Similarly, all the heterocycle-modified nucleotides (**5–7**) served as good substrates for Vent(exo-), belonging to the same family of DNA polymerases (Figure 6). In the presence of **T1–T5** the enzyme produced fully extended primers containing one or more modifications near and away from the start site. Taken together, KOD and Vent(exo-) DNA polymerases generate full-length products with all the tested templates and nucleotides while KlenTaq has difficulties with more demanding templates (two templating adenosines in a row or one nucleotide apart) and can generate multiple labeled full-length products only with the smaller modified SedUTP.

2.2.3 Single nucleotide incorporation and site-specific labeling

Site-specific labeling of ONs with only one functional group is difficult using templates that contain the respective templating nucleotide at several positions (e.g., **T6**). Therefore, we chose to develop a two-step protocol by first incorporating one modified nucleotide and after a reaction clean-up (to remove the unreacted modified triphosphate), add the four natural nucleotides to get the full-length product by PEX (Figure 7A). Prerequisite for the used enzyme should be that it faithfully incorporates only one modified nucleotide at the 3'-end of the primer in a template-dependent manner while it also fully elongates the modified primer with natural substrates or the singly modified primer should be compatible for enzymatic ligation.⁷⁴⁻⁷⁹ In order to establish such a method, we tested single nucleotide incorporation properties of KOD and KlenTaq in the presence of modified nucleotides. A 5'-FAM-labeled primer **P2** was hybridized with a template DNA ON **T6** containing a dA nucleotide that would guide the incorporation of a modified mononucleotide at the 3'-end of the primer. Reactions in the presence of KOD DNA polymerase resulted in largely doubly modified ON products due to one correct incorporation of dU* opposite dA and one misincorporation of dU* opposite dC (Figure 7B). Attempts to obtain a singly modified primer with KOD DNA polymerase by varying the concentration of the enzyme and dU*TP were not successful (data not shown). Rewardingly, KlenTaq DNA polymerase incorporated all the three modified nucleotides with approximately same apparent efficiency as dTTP, yielding exclusively singly-modified primer ON products (Figure 7C). To unambiguously confirm the formation of singly-labeled products, large-scale reactions were performed and the identity of the

isolated products was confirmed by mass analysis (Table 3 and Figure 8A–D). Typically, reactions performed in a 1.25 nmol scale and purified by PAGE gave 60–70% of the labeled products.

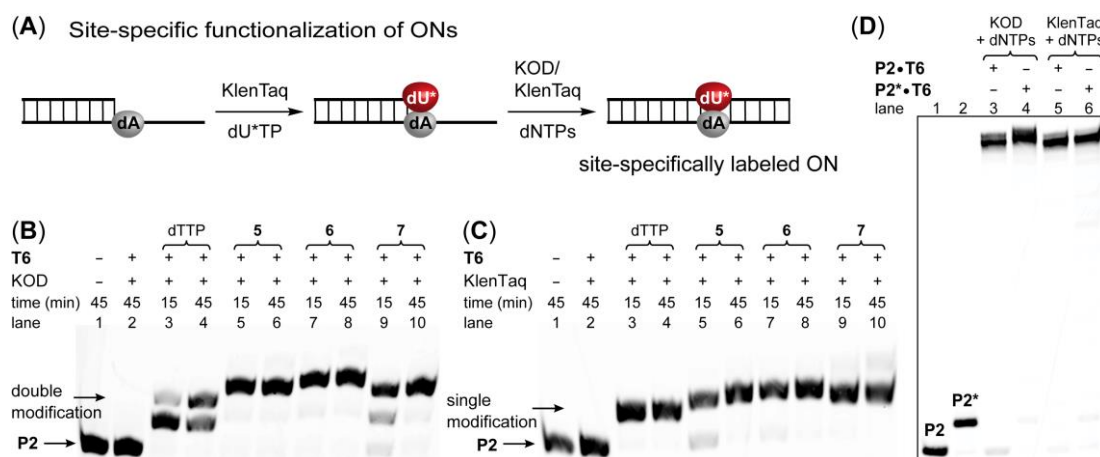


Figure 7. (A) Schematic diagram illustrating the site-specific functionalization approach. dU*TP represents a modified nucleotide analog probe. (B) In the presence of dTTP or modified nucleotides (5–7) KOD DNA polymerase added two consecutive nucleotides to the primer **P2**. (C) KlenTaq DNA polymerase produced only singly-modified ONs. (D) 3'-BFdU-labeled primer **P2*** was efficiently extended by KOD and KlenTaq DNA polymerase to give the full-length site-specifically functionalized ON product **P2****. See Table 2 and Table 3 for ON sequences.

Encouraged by these results, we devised a two-step template-dependent site-specific labeling technique by combining the stringent and flexible processivity of KlenTaq and KOD DNA polymerases, respectively. In the first step, **P2•T6** duplex was reacted with BFdUTP in the presence of KlenTaq, which resulted in the quantitative incorporation of a single BFdU label at the 3'-end to produce ON **P2*** (Figure 7D, lane 2). The reaction mixture was heated and the modified primer•template duplex was precipitated using ethanol and washed well. This step was performed to inactivate the enzyme and remove unreacted modified nucleotide. The precipitated **P2*•T6** duplex was annealed in the reaction buffer in the same tube and was subjected to PEX in the presence of dNTPs and KOD. The reaction resulted in almost quantitative production of the full-length site-specific labeled ON product **P2**** (Figure 7D lane 4, Table 3). A control PEX reaction performed in the presence of **P2•T6** duplex and dNTPs confirmed the formation of full-length product formed in the presence of modified primer•template duplex (compare lanes 3 and 4, products move with a similar mobility). The identity of the product **P2**** was further ascertained by mass analysis (Figure 8E and Table 3). The site-specific labeling could also be achieved using KlenTaq in the second step as it

efficiently processed the modified primer•template duplex to give the full-length product in the presence of all four natural nucleotides (Figure 7D, lane 6). If required, the template can be specifically digested using lambda exonuclease giving access to the labeled single-stranded ON.⁸⁰ Collective, these results indicate that the size of the heterocycle modification does not matter in the single nucleotide incorporation step using KlenTaq as it selectively incorporates all three modified dU*MPs opposite one templating adenosine. Furthermore, it does not show misincorporation of a second nucleotide as it is the case with KOD DNA polymerase rendering it superior for the first step in the site-specific labeling approach while for the elongation KlenTaq as well as KOD DNA polymerases can be used. This method could allow the site-specific installation of a variety of nucleoside probes, which can be used for nucleic acid conformation and function analysis.^{5,43-45,56}

Table 3. ESI-Mass and yield data for DNA ON products.

Product	Calculated Mass	Found Mass	Isolated yield (nmole)	Isolated yield (%)
5'-FAM,3'-BFdU labeled ON P2*	6708.48	6708.13	7.9 ^a	70
3'-BFdU labeled ON P3*	6170.94	6170.88	7.7 ^a	68
5'-FAM,3'-BTdU labeled ON P2[†]	6724.54	6724.25	2.5 ^b	65
5'-FAM,3'-SedU labeled ON P2[‡]	6721.39	6721.38	2.3 ^b	61
Site-specifically labeled ON P2**	13778.05	13778.0	nd ^c	nd

^aIsolated yields are given for a reaction scale of 11 nmol. ^bIsolated yields are given for a reaction scale of 3.8 nmol. See below Table 4 for ON sequences. For ESI mass spectra see Appendix I. ^cThe yield for the product was not determine as it was isolated as a duplex (**P2**•T6**).

Table 4. Sequences of modified ONs

Modified ONs	Sequence
P2*	5' FAM-GGA GCT CAG CCT TCA CTG C- BFdU 3'
P2[†]	5' FAM-GGA GCT CAG CCT TCA CTG C- BTdU 3'
P2[‡]	5' FAM-GGA GCT CAG CCT TCA CTG C- SedU 3'
P3*	5' GGA GCT CAG CCT TCA CTG C- BFdU 3'
P2**	5' FAM-GGA GCT CAG CCT TCA CTG C- BFdU -GC TAA ACT ATA TTG CTA GCG CAC 3'

2.2.4 Responsive fluorescent BFdU provides insights on the enzyme–primer•template duplex complex formation

To study the efficacy of the nucleotide substrates to serve as probes to monitor the binding processes involved in the polymerization reaction, we performed fluorescence assays. For this purpose, BFdU was chosen as it has been found to be more emissive and highly sensitive

to its local environment.^{53,59} The enzyme first forms a complex with primer•template duplex and then the cognate nucleotide binds to the active site to form the ternary complex. We first studied the effect of modification on the binary complex (enzyme–primer•template duplex) formation by incubating unmodified **P2•T7** and 3'-BFdU-modified **P2*•T7** duplexes with increasing concentrations of KlenTaq. Electrophoretic mobility shift assay (EMSA) performed under non-denaturing conditions clearly revealed the formation of binary complexes, wherein BFdU-modified ($K_d = 0.24 \pm 0.04 \mu\text{M}$) and native ($K_d = 0.23 \pm 0.03 \mu\text{M}$) duplexes exhibited similar binding affinities for the enzyme (Figure 8). Encouraged by these results, we designed an experimental setup to follow the formation of the binary complex by monitoring the changes in the fluorescence of duplex **P3*•T7** as a function of increasing enzyme concentration (Figure 9A). Here, a BFdU-modified primer sequence (**P3***) same as **P2*** but without a FAM label was used to avoid the interference from fluorescein. Upon excitation ($\lambda_{max} = 330 \text{ nm}$), a concentration-dependent increase in fluorescence intensity with negligible changes in emission maximum ($\lambda_{em} = 430 \text{ nm}$) as a result of formation of the binary complex was observed (Figure 9B). A plot of normalized fluorescence intensity *versus* enzyme concentration gave an apparent K_d of $0.76 \pm 0.10 \mu\text{M}$, which is comparable to the EMSA data (Figure 9C). BFdU is highly sensitive to micropolarity changes, and exhibits enhanced fluorescence intensity and significant red shift in emission maximum when it is located in a polar environment.⁵⁹ The free nucleoside analog exhibits a λ_{em} of 446 nm and 423 nm in water and methanol, respectively. When incorporated into ONs its fluorescence also depends on neighboring base environment. For example, stacking interaction and electron transfer processes between the modified nucleoside and adjacent bases, particularly with guanine base, lead to fluorescence quenching.⁵⁹ On the contrary, rigidification of the fluorophore due to the presence of a molecular rotor element in the probe (benzofuran is attached to the dU ring *via* a rotatable bond) can result in significant enhancement in fluorescence intensity. Therefore, a combination of these factors, in a given nucleic acid conformation, can influence the fluorescence outcome of the probe. The free duplex **P3*•T7** displays a very low fluorescence intensity due to stacking interaction and free rotation of the fluorophore. In case of the binary complex, an emission band at 430 nm suggests that BFdU placed in the active site experiences an environment, which is reasonably polar (near methanol). However, the complex exhibits significant enhancement in fluorescence intensity, which is most likely due to the rigidification of the fluorophore in the enzyme pocket.

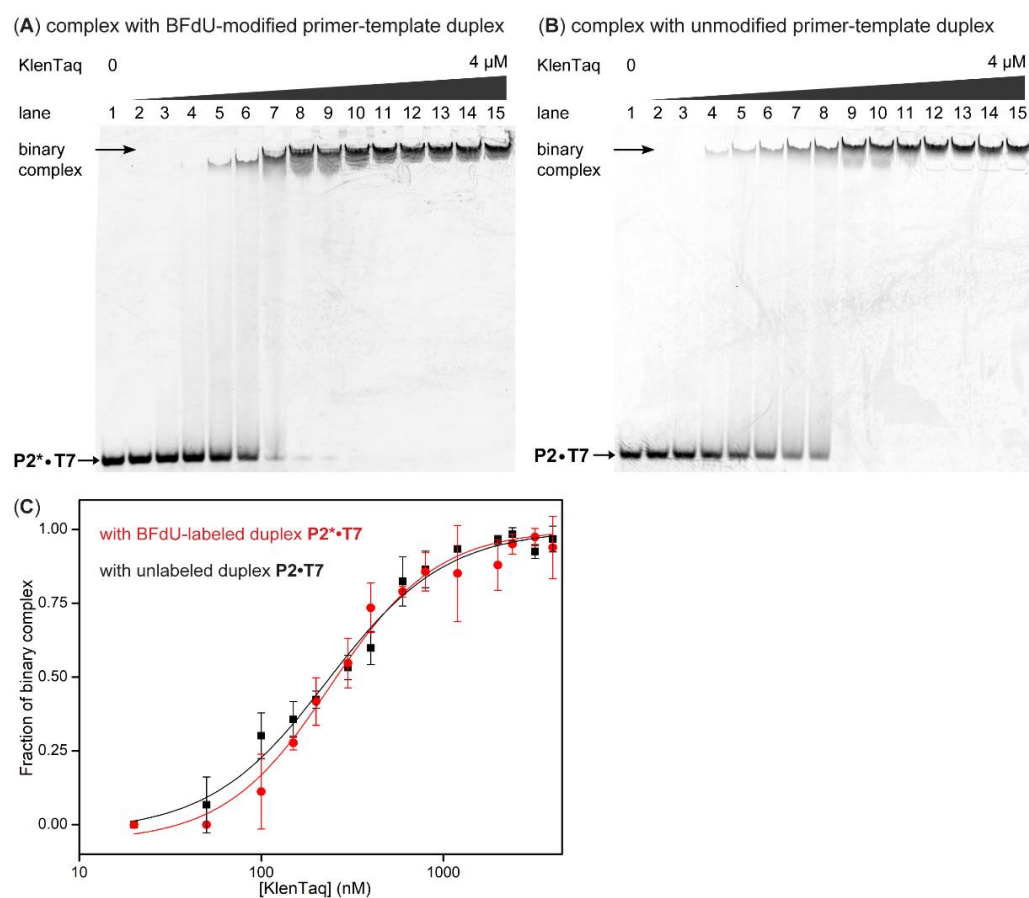


Figure 8. Binary complex formation using EMSA. **(A)** Complex formation with BFdU-modified primer template duplex ($P2^* \bullet T7$) with increasing concentrations of the enzyme. **(B)** Complex formation with unmodified primer template duplex ($P2 \bullet T7$) with increasing concentrations of the enzyme. **(C)** Curve fits for the formation of the same.

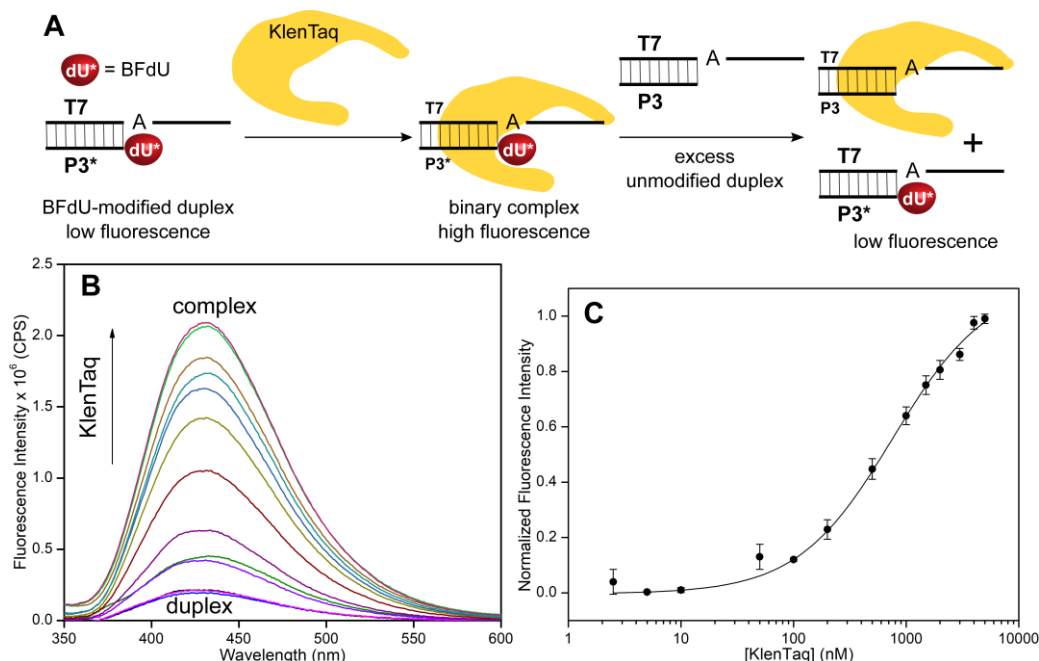


Figure 9. Fluorescence of BFdU reports the formation of a binary complex. **(A)** Scheme for fluorescence monitoring of the binary complex formation and competitive displacement assay. **(B)** A plot showing the increase in fluorescence intensity of duplex **P3*•T7** (50 nM) as a function of increasing KlenTaq concentration (0–5 μ M). Samples were excited at 330 nm with excitation and emission slit widths of 7 nm and 8 nm, respectively. **(C)** Curve fit for the formation of the binary complex. Normalized fluorescence intensity at λ_{em} (430 nm) plotted against concentration of the enzyme.

To prove that the increment in fluorescence intensity is a result of specific binding of BFdU-labeled DNA duplex in the active site, a competitive displacement binding assay was performed. A binary complex (duplex **P3*•T7** 50 nM and KlenTaq 1 μ M), which would produce near saturation fluorescence, was prepared and titrated with increasing concentrations of the unlabeled DNA duplex (**P3•T7**). Addition of unlabeled duplex resulted in a progressive decrease in fluorescence intensity due to the displacement of the less emissive modified duplex from the active site (Figure 10). It is to be noted that even 20-fold excess of the unmodified duplex did not completely displace the modified duplex from the enzyme. We confirmed this competitive displacement process by EMSA also. Much like a binary complex made of modified duplex, which was hard to displace using unmodified duplex, a complex of enzyme-unmodified duplex was strong and required excess of modified duplex to break the complex (Figure 11). Further, addition of enzyme alone to the nucleotide analog did not significantly affect its fluorescence indicating that the analog does not bind to the free enzyme (Figure 12).

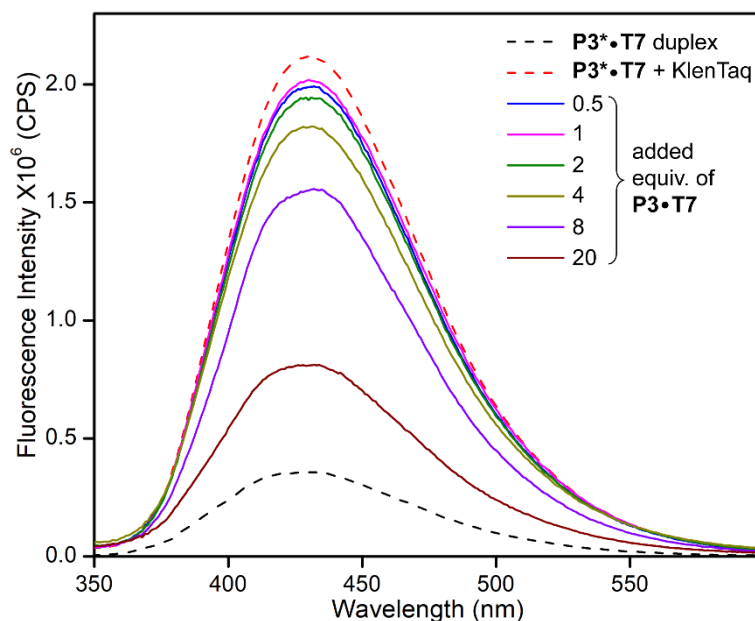


Figure 10. Competitive displacement assay by fluorescence. BFdU-modified primer-template duplex **P3*•T7** (50 nM) shows very low fluorescence (black dashed line). A binary complex of **P3*•T7** (50 nM) and KlenTaq (1 μ M) shows very high fluorescence (red dashed line). Addition of increasing equivalents of unlabeled duplex (**P3•T7**) displaces the complex resulting in progressive decrease in fluorescence (solid lines). See Section 5 for details.

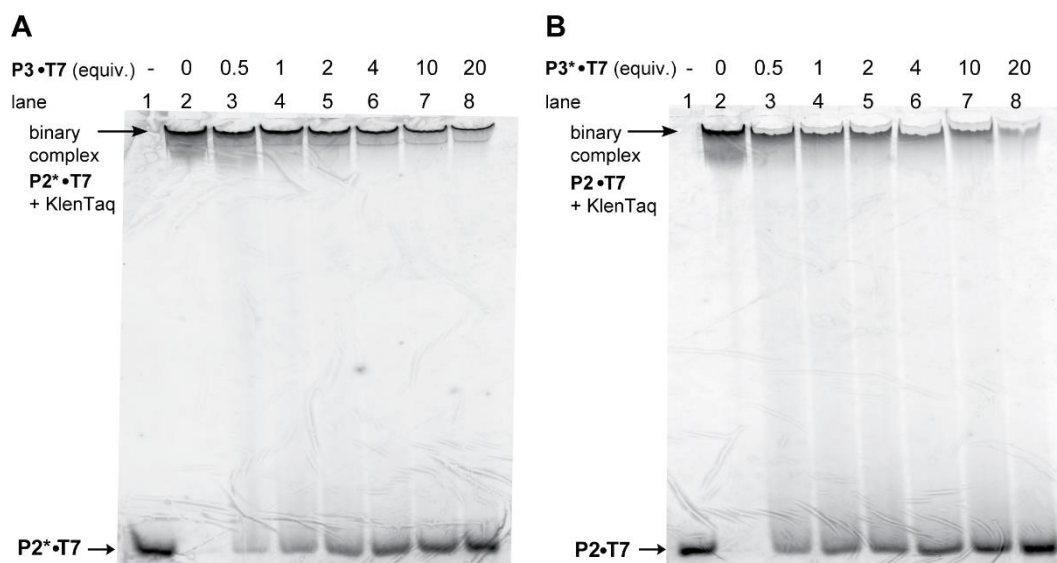


Figure 11. Competitive displacement assay by EMSA. (A) Addition of unmodified duplex **P3•T7** to a binary complex of **P2*•T7** (5'-FAM, 3'-BFdU-modified primer template duplex) and KlenTaq resulted in displacement of the duplex **P2*•T7** (lower band). **P3** is the same sequence as **P2**, but without 5'-FAM label. (B) Addition of 3'-BFdU-modified primer-template duplex **P3*•T7** to a binary complex of **P2•T7** and KlenTaq resulted in displacement of the duplex **P2•T7** (lower band). In both the cases, the binary complex with native and BFdU-modified duplexes were strong and even excess of the respective duplexes could not completely disrupt the complexes, which is also observed in fluorescence assay.

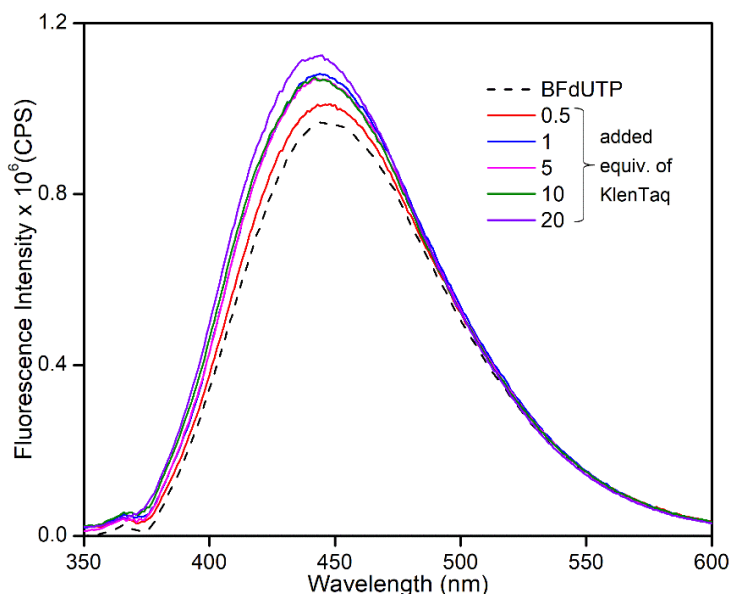


Figure 12. Fluorescence spectra of BFdUTP 5 (50 nM) as a function of increasing KlenTaq concentration. Samples were excited at 330 nm with excitation and emission slit widths of 6 nm and 7 nm, respectively. Each sample was incubated at 25 °C for 30 min.

Fluorescence studies using our probe provide further insights on the functioning of the KlenTaq DNA polymerase. The enzyme efficiently incorporated the nucleotide analogs at the 3'-end of the primer, which is consistent with the active site conformation and narrow channel volume observed in the crystal structures (*vide infra*). The enzyme was further able to efficiently extend the singly-labeled primer•template duplex in the presence of native dNTPs to produce the full-length site-specifically labeled product (Figure 7D). For this to happen, the enzyme needs to form a stable binary complex with the 3'-BFdU-labeled primer•template duplex. Competitive fluorescence assay complemented by EMSA indicate that KlenTaq forms a stable complex with the pre-annealed BFdU-labeled duplex (Figure 12).

2.2.5 Nucleotide substrate itself serves as a probe to report its incorporation by KlenTaq

The differences in fluorescence exhibited by the free nucleotide analog BFdUTP 5 and 3'-BFdU-labeled primer•template duplex prompted us to design an assay to detect the activity of KlenTaq (Figure 13A). For this purpose, **P3•T7** primer•template duplex, which would incorporate only one modified nucleotide analog at the 3'-end was used. Addition of BFdUTP to primer•template duplex or KlenTaq alone produced only a small change in fluorescence intensity (Figure 14B, red and blue curves). Rewardingly, when BFdUTP was added to samples of a binary complex (enzyme and primer•template duplex), each containing

increasing concentrations of the enzyme, a significant reduction in fluorescence intensity accompanied by a blue shift in emission maximum (~ 430 nm) as compared to the free nucleotide analog (dashed black curve, $\lambda_{em} = \sim 445$ nm) was observed (Figure 13B). This fluorescence outcome is due to the incorporation of a single BFdUMP at the 3'-end of the primer, which produces a weakly emissive 3'-BFdU-labeled primer•template duplex (**P3•T7**) as shown in the earlier experiment (see Figure 9B). In a control experiment, a binary complex made of KlenTaq and **P5•T7** duplex (**P5** contains a ddC at the 3'-end) was added to BFdUTP. Since this primer would not allow the incorporation of the fluorescently modified nucleotide analog, no decrease in fluorescence intensity was observed (magenta curve). Collectively, these results indicate that the nucleotide substrate, which is environment-sensitive, can itself be used as a probe to study the binary complex formation and also the activity of polymerase enzymes, here KlenTaq, by fluorescence spectroscopy.

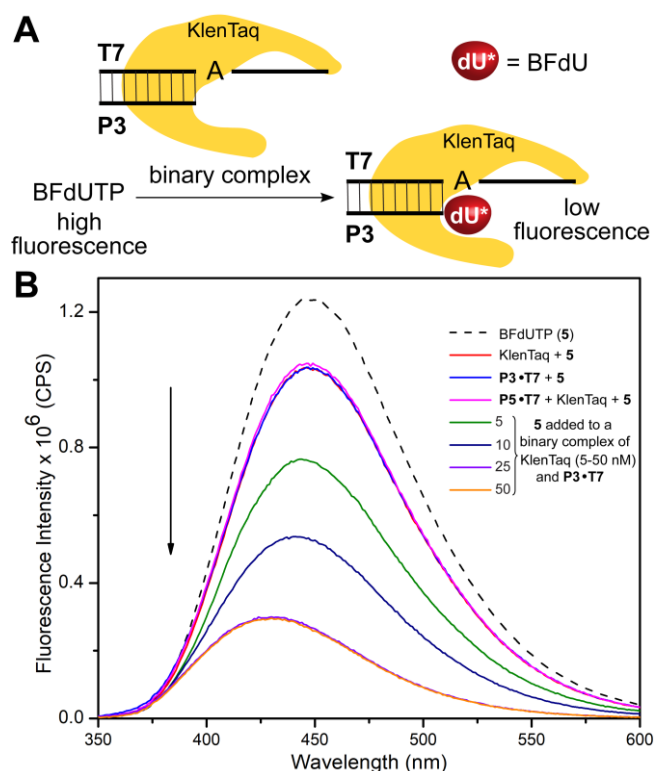


Figure 13. BFdU fluorescently reports nucleotide analog incorporation by KlenTaq. **(A)** Scheme showing the assay setup. **(B)** A plot showing decrease in fluorescence intensity of BFdUTP (50 nM) as a function of increasing KlenTaq concentration (0–50 nM) in the presence of **P3•T7** primer•template duplex (200 nM). Incorporation of BFdUMP at the 3'-end results in decrease in fluorescence intensity as the labeled duplex shows low fluorescence. In control experiments, addition of KlenTaq (200 nM) or **P3•T7** (200 nM) to BFdUTP (50 nM) showed only small changes in fluorescence intensity. Addition of a binary complex made of 1:1 molar solution (200 nM) of KlenTaq and **P5•T7** duplex (**P5** contains a ddC at the 3'-end) to BFdUTP (50 nM) did not show appreciable reduction in fluorescence intensity. All samples were excited at 330 nm with excitation and emission slit widths of 6 nm and 7 nm, respectively.

2.2.6 X-ray structure of the ternary complex sheds light on the incorporation behaviour

To gain further insights into the incorporation propensities observed in PEX reactions for KlenTaq and KOD DNA polymerases, we planned to solve crystal structures of these enzymes in the replicative state poised to incorporate the fluorescently-modified dU*MPs. A template **T8** was designed such that it codes for the incorporation of ddCMP and modified nucleotides as the first and second nucleotides, respectively (Figure 14A). Priming with ddCTP terminates the polymerization process due to the absence of the 3'-OH group, thereby enabling the capture of modified nucleotides in the nucleotide binding site prior to incorporation. Rewardingly, we obtained a high-resolution structure of the ternary closed complex composed of KlenTaq, primer•template duplex and BFdUTP (termed KlenTaq-BFdUTP, 1.9 Å, PDB: 7OWF, Figure 14B, Table 5). Crystallization attempts with KOD, however, were unsuccessful. KlenTaq-BFdUTP adopts a structure, which closely resembles the structure of KlenTaq bound to a native nucleoside triphosphate in the nucleotide binding site (PDB: 3RTV, Figure 14B and 14C).¹⁹ The RMSD value for C α atoms is 0.274 Å (467 atoms aligned). The nucleotide analog is positioned nicely to form complementary Watson-Crick pairing with the templating dA residue in the active site (Figure 14D). The benzofuran and uracil rings are slightly non-coplanar and partially stack with ddCMP incorporated at the 3' end of the primer (Figure 14D and Figure 16A).

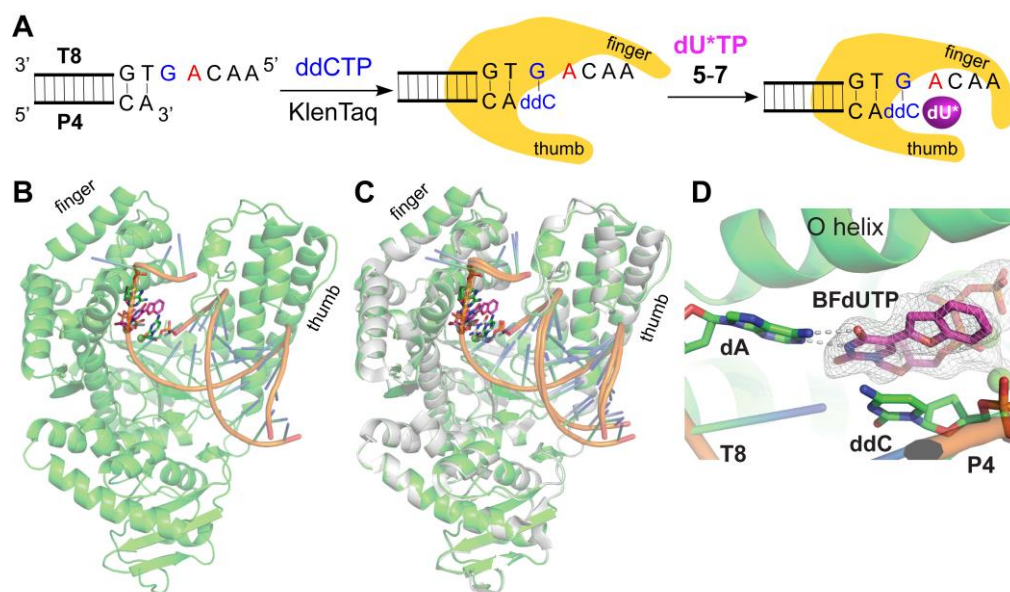


Figure 14. (A) Crystallization setup to trap the fluorescent nucleotide probes (5–7) in the ternary complex. (B) Overall structure of KlenTaq-BFdUTP ternary complex (green). (C) Superimposition of KlenTaq-BFdUTP (green) with a native KlenTaq ternary complex (gray; PDB: 3RTV). (D) Close-up

view of the active site showing the bound nucleotide analog. Depicted is the omit map of BFdUTP at 3σ . The Watson-Crick base pairing between the nucleotide analog and dA of the template strand is indicated with dashed lines. In all structures BFdUTP is highlighted in magenta.

Typical of an active ternary complex, the finger domain is in a closed state and the O helix is placed on top of BFdUTP bound in the active site (Figure 15A). The main difference in KlenTaq-BFdUTP compared to the natural complex is the orientation of the two important amino acid residues Arg660 and Arg587 near the modification (Figure 15A and 15B). Arg660 usually forms a hydrogen bond to the phosphate backbone of the primer 3'-terminus, which is suggested to stabilize the closed and active conformation in KlenTaq.^{81,82} Here, Arg660 flips away from the active site to accommodate the benzofuran modification. Simultaneously, Arg587 moves closer towards the benzofuran ring and is now placed between the benzofuran ring and the phosphate backbone of the 2nd primer nucleotide. This arrangement results in the formation of a crevice, which nicely encases the heterocycle modification in the nucleotide binding site (Figure 15C and Figure 16B). Relatively high B factors and poor electron density suggest that Arg660 and Arg587 are reasonably flexible. Due to the movement of Arg660, the interaction of the O-helix to the primer backbone is lost which might contribute to a higher flexibility of the whole finger domain. The distance between the 3'-end carbon atom of the primer and the α -phosphorus of the modified nucleotide (3.9 Å) is very similar to the distance observed in the presence of a natural substrate (3.8 Å, for 3RTV), which indicates that the formation of a phosphodiester bond should not be hindered by the presence of the modification (Figure 15D and 15E). These findings correlate with the efficient incorporation of all three nucleotides in primer extension experiments by KlenTaq. In *Geobacillus stearothermophilus* (Bst) DNA polymerase (another structurally well characterized A family DNA polymerase),⁸³⁻⁸⁵ R703 and R629 residues are analogous to the flexible R660 and R587 residues in KlenTaq.⁸⁴ We speculate that Bst could possibly accommodate BFdUTP in the active site in a similar fashion with the help of these two flexible residues, and hence, could be potentially used to incorporate base-modified nucleotides.

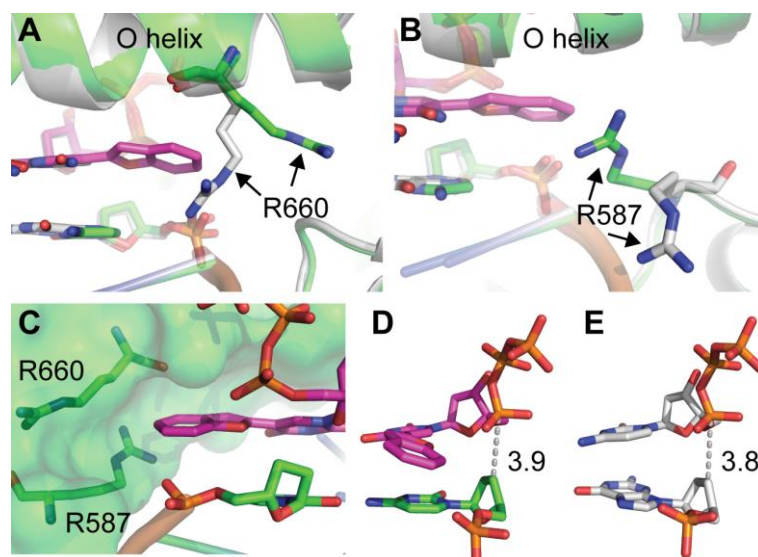


Figure 15. (A) and (B): Superimposition of KlenTaq-BFdUTP (green) and KlenTaq (gray; PDB: 3RTV) highlighting the different orientation of Arg660 and Arg587, respectively. (C) The cleft in which the benzofuran ring is located in the active site is shown. (D) and (E) Close-up view showing the distance between the α -phosphorus and 3'-end of the primer in the ternary complex structures of KlenTaq-BFdUTP and 3RTV, respectively. BFdUTP is highlighted in magenta.

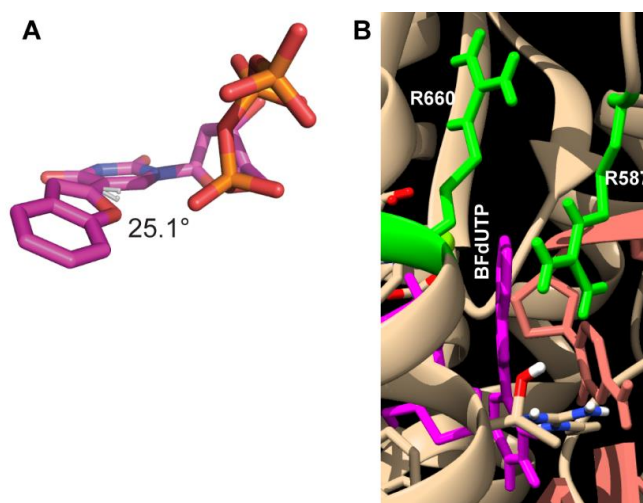


Figure 16. (A) Benzofuran and uracil rings are non-coplanar. Depicted is the dihedral angle of the non-planar BFdUTP 5. (B) Another view of the cleft in which the BFdU ring (magenta) is encased is shown (rendered using UCSF Chimera). Arg660 and Arg587 residues are colored green. Protein and DNA are shown in brown and salmon colors, respectively.

Table 5. Summary of data collection and refinement statistics.

	KlenTaq-BFdUTP
PDB ID	7OWF
Wavelength (Å)	1.0
Space group	P3 ₁ 2 ₁ 2 (152)
Cell dimensions	

a, b, c (Å)	110.18, 110.18, 91.30
α , β , γ (°)	90.00, 90.00, 120
Resolution (Å)*	47.71 – 1.90 (1.90 – 2.02)
Total no. of reflections	732774 (87716)
No. of unique reflections	96475 (15157)
R _{meas} (%)	13.4 (318.9)
I / σ	7.45 (0.46)
Completeness (%)	99.5 (97.0)
Redundancy	7.6 (5.8)
CC _{1/2} (%)	99.7 (24.7)
ISa	9.84
Refinement	
Resolution (Å)	47.71 – 1.91
No. of reflections	96258
R _{work} / R _{free}	19.53 / 24.12
Coordinate error	0.35
No. of atoms	
Protein	8622
DNA (p/t/dNTP)	373/509/37
Water	102
Average B-factors (Å)	
Protein	66.45
DNA (p/t/dNTP)	54.26 / 60.76 / 55.75
Water	50.91
R.m.s. deviations	
Bond lengths (Å)	0.013
Bond angles (°)	1.310
Ramachandran (%)	
Favored/ Allowed/ Outlier	95.17 / 4.46 / 0.37

* Values in parentheses correspond to those in the outer resolution shell.

2.2.7 Structural data correlates with KOD and KlenTaq incorporation and elongation behaviour

In general, B family polymerases (e.g., KOD) are more efficient at utilizing base-modified nucleotides compared to the A family polymerases (e.g., KlenTaq).^{18,66} A comparative analysis of the structures of ternary complexes of KlenTaq and KOD bound to a native nucleotide or our modified nucleotide analog provides a cogent explanation for the observed differences in the processivity of these families of enzymes (Figure 17).⁸⁶ The study identified two main differences between the enzymes to be responsible for the different acceptance of modified substrates (Figure 17A–17C). I: The KOD ternary complex exhibits a larger channel within the enzyme⁸⁷ that allows diffusion of the triphosphate, leaving of the pyrophosphate as well as accommodation of modifications attached to the nucleobases. II: The DNA conformation close to the active site, which is B-form in KOD and A-form in the first three primer nucleotides in KlenTaq. B-form DNA is characterized by larger internucleobase distances and a wider major groove compared to A-form DNA offering more space to accommodate major groove modifications as studied here. These two properties

could be the reason why KOD (and B-family DNA polymerases in general) is able to incorporate not only one modified nucleotide but also two in a row or one nucleotide apart in more complex template sequences while KlenTaq is unable to incorporate multiple nucleotides with bigger modifications. At the same time, this could explain why KOD (in contrast to KlenTaq) inserts an additional mismatched dU* in the single nucleotide incorporation experiments. An important point considering the elongation of a modified substrate is the interaction pattern of the protein with the DNA duplex further away from the active site. In KlenTaq as well as in KOD the primer-template duplex is encased in the groove formed by the thumb and palm domains (Figure 15B and 15C, Figure 17A).⁸² It was shown that the architecture of the tip of the thumb domain contacting the primer strand differs greatly in the two enzymes (Figure 17D–17F).⁸⁸ In KOD, the loop formed by residues 668–675 hovers over the minor groove only (Figure 17D), while in KlenTaq the loop residues 506–509 extends over the phosphate backbone towards the major groove (Figure 17E). This is also the case in KlenTaq-BFdUTP structure (Figure 17F). A recent structural analysis conducted to probe the movement of C5-ethynylbenzene-modified cytidine along the confines of KlenTaq enzyme shows the plasticity of the enzyme in accommodating the modification during the course of primer extension.⁸⁹ Despite the narrow channel and loop conformation in KlenTaq, primers containing a ethynylbenzene-modified cytidine placed at different positions from the 3' terminus are well accommodated. Based on this study and the KlenTaq-BFdUTP structure, we infer that the incorporation of a single fluorescent nucleotide analog (irrespective of the heterocycle modification) into the primer does not affect its movement along the channel to produce respective full-length PEX products. However, the narrow channel could potential attenuate the ability of KlenTaq to insert multiple bulky modifications in close proximity. In support of this notion, SedUTP **7** containing a smaller heterocycle (selenophene) is efficiently processed by KlenTaq to produce full-length products with all the templates, while the enzyme showed poor processivity with the bulkier BFdUTP **5** and BTdUTP **6** substrates, producing truncated products. Taken together, the existing data and present findings provide a cogent understanding on how to choose a polymerase depending on the modification when devising a labeling method. For example, B family polymerases are suited for high density incorporation of fluorescent nucleotide analogs, whereas A family polymerases are more suited for site-specific 3'-end and single internal labeling.

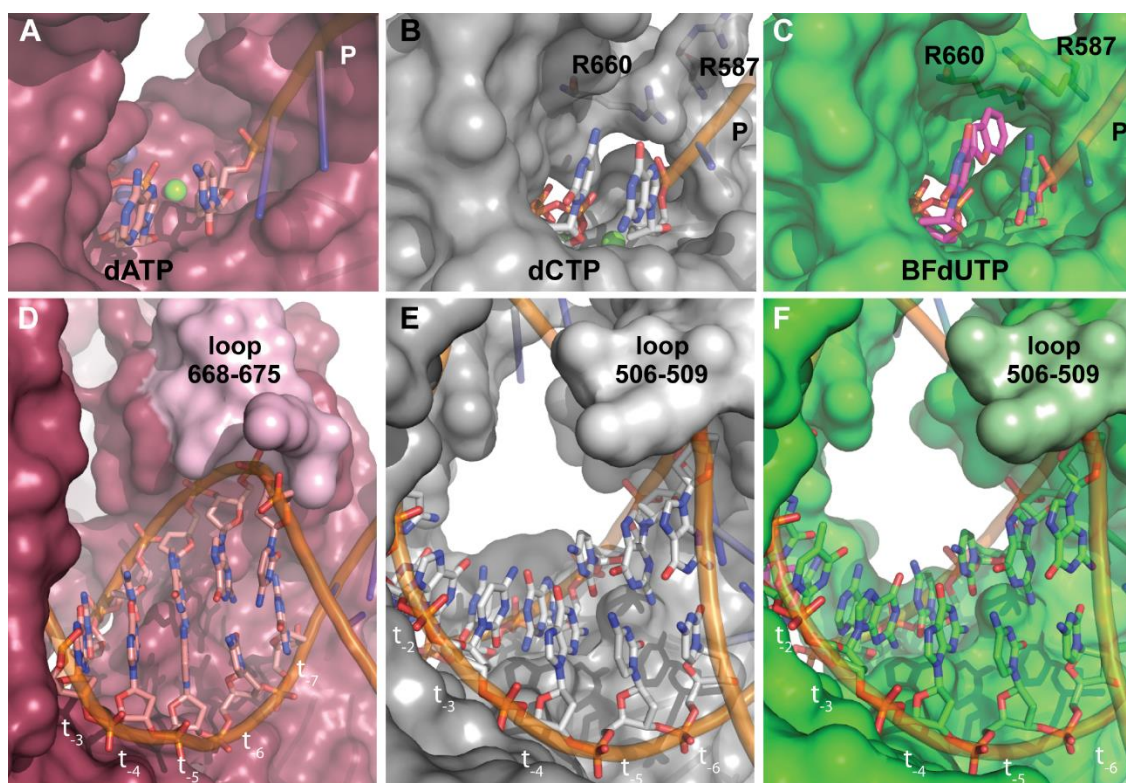


Figure 17. Comparison of KOD (raspberry; PDB: 5OMF), KlenTaq (gray; PDB: 3RTV) and KlenTaq-BFdUTP (green). The proteins are shown as surface, primer and template are depicted as ribbons and relevant nucleotides are shown in sticks using the respective color-codes. BFdUTP is highlighted in magenta. (A-C) A zoom-in of the active site showing the differences in the binding pocket space. (E-F) Loop residues (highlighted in brighter color) of the tip of the thumb domain of respective polymerases interacting with the primer strand. Primer strand is denoted as P.

2.3 Conclusion

A battery of biochemical and structural analyses using responsive fluorescent nucleotide analogs has provided valuable insights into the processing ability of A and B family DNA polymerases. Reactions with different templates that guide the insertion of a single and multiple nucleotide analogs in different phases of the polymerization process indicate differential incorporation behaviour of the enzymes, which depends on the size of the heterocycle modification. B family polymerases processed all the nucleotide analogs efficiently to produce ONs with labels at different positions, which was consistent with the situation in the active site (B-form DNA compared to A-form DNA) and the large channel volume available within the enzyme (e.g., KOD) to accommodate the modifications. On the other hand, KlenTaq showed considerable stringency in incorporating bulkier nucleotides particularly at consecutive positions in PEXs. In the single nucleotide incorporation setup, KlenTaq precisely and efficiently added only one modified nucleotide, irrespective of the size of modification in contrast to KOD, which misincorporated a second nucleotide. Therefore,

KlenTaq was preferably used in the first step to generate site-specifically labeled ON products. After a purification step, the modified primer can then be fully elongated by further primer extension with either KOD or KlenTaq.

Notably, BFdU label serves as an efficient fluorescent reporter by providing a real-time read-out for the binary complex formation and enzyme activity under equilibrium conditions. The fluorescence study provides useful insights on the affinity of KlenTaq to bind and process 3'-end-labeled primer to produce site-specifically labeled ON product. These results expand the utility of our environment-sensitive nucleotide analogs not only as probes for nucleic acid conformational analysis but also as probes for monitoring events in the polymerization process. Further, comparing the native and KlenTaq-BFdUTP structures, the following understanding on the processivity of the polymerase was gained. Three important features of the ternary complex, namely (i) the flexibility of residues Arg660 and Arg587 near the active site, (ii) the narrow channel within the enzyme and (iii) the geometry of the primer•template duplex near the active site, appear to be important factors in the processing of heterocycle-modified nucleotides by KlenTaq in this regard. The flexible Arg660 and Arg587 residues serving as “gatekeepers” adopt an altered conformation to accommodate the analogs in the nucleotide binding site, in a manner that allows catalysis to occur. Once the modification is incorporated, the efficiency of elongation process is now controlled by the plasticity of the channel and the loop that hovers over the major groove side of the primer. In our case, we observe that a smaller modification (SedU) is processed unhampered, whereas judicious template design is required for incorporating larger modifications (BFdU and BTdU). Taken together, these new findings will be highly valuable in designing DNA polymerase-compatible nucleotide probes and also in choosing the right enzyme for a given application, namely site-specific *versus* high density labeling. In future, we plan to use these nucleotide substrates as probes to study the incorporation mechanism of nucleic acid processing enzymes from human pathogens (e.g., viral polymerases) and develop a discovery platform to identify potential binders by competitive fluorescence-based ligand-binding assay described in this work.

2.4 Experimental Sections

2.4.1 Materials

5-Iodo-2'-deoxy uridine was purchased from TCI chemicals. *Bis* (triphenylphosphine)-palladium (II) chloride, benzofuran, benzothiophene, selenophene, butyllithium and

tributyltin chloride were purchased from Sigma Aldrich. 5'-FAM labeled DNA ON, primers and templates were purchased from Integrated DNA Technologies, Inc. DNA templates were purified by polyacrylamide gel electrophoresis (PAGE) under denaturing conditions, and desalted using Sep-Pak Classic C18 cartridges (Waters Corporation). dNTPs were purchased from Thermo Fischer Scientific. Vent(exo-) DNA polymerase was purchased from New England BioLabs. KlenTaq and KOD DNA polymerases were expressed and purified as reported in the literature.^{S1,S2}

2.4.2 Instrumentation

NMR spectra were recorded on a 400 MHz Jeol ECS-400 and Bruker Ultrashield 400 MHz WB Plus spectrometers. Mass measurement were performed using Water Synapt G2 High Definition mass spectrometer. HPLC analysis was done using Agilent Technologies 1260 Infinity HPLC. Reverse-phase flash chromatographic (C18 RediSepRf column) purifications of nucleotides were carried out using Teledyne ISCO, Combi Flash Rf. ESI-MS analysis of DNA ONs was performed using Water Synapt G-Si Mass spectrometry instrument in negative mode. DNA and nucleotides were quantified by measuring absorbance in Shimadzu UV-2600 or NanoDropTM 2000c spectrophotometer. Fluorescent DNA ONs were resolved by PAGE on an OWL S4S sequencing gel electrophoresis instrument and was imaged using Typhoon FLA 9500 Variable mode imager. Steady-state fluorescence spectra were recorded using micro-fluorescence cuvette (Hellma, path length 1.0 cm) on Horiba Fluoromax-4 spectrophotometer.

2.4.3 Synthesis of nucleotide analogs 5–7

5-Benzofuran-modified 2'-deoxyuridine **2** and 5-selenophene modified 2'-deoxyuridine **4** were synthesized using previously reported protocols.^{57,73} Both the compounds were characterized by NMR and mass spectrometry and matched well with the previously reported data.

BFdU 2. ¹H-NMR (400 MHz, *d*₆-DMSO): δ (ppm) 11.75 (s, 1H), 8.74 (s, 1H), 7.63 (d, *J* = 8.4 Hz, 1H), 7.56 (d, *J* = 8.4 Hz, 1H), 7.34 (s, 1H), 7.28 (t, *J* = 8 Hz, 1H), 7.22 (t, *J* = 7.6 Hz, 1H), 6.23 (t, *J* = 6.4 Hz, 1H), 5.30 (d, *J* = 4.4 Hz, 1H), 5.23 (t, *J* = 4.4 Hz, 1H), 4.34 (m, 1H), 3.87 (d, *J* = 3.2 Hz, 1H), 3.72–3.67 (m, 2H), 2.30–2.19 (m, 2H); ¹³C NMR (100 MHz, *d*₆-DMSO): δ (ppm) 160.8, 153.5, 149.9, 149.6, 137.6, 129.3, 124.8, 123.5, 121.5, 111.3, 105.2, 104.4, 88.1, 85.6, 70.6, 61.4, 41; HRMS: *m/z* calculated for C₁₇ H₁₅ N₂ O₆ [M-H]⁻ = 343.0936, found [M-H]⁻ = 343.0937.

BTdU 3. To a suspension of 5-iodo-2'-deoxyuridine **1** (500 mg, 1.41 mmol, 1.0 equiv.) in 15 mL of degassed anhydrous dioxane was added *bis* (triphenylphosphine)-palladium (II) chloride (49.6 mg, 0.07 mmol, 0.05 equiv.). To this, 2-(tri-*n*-butylstannyl) benzothiophene (1.44 g, 3.53 mmol, 2.5 equiv.) was added drop wise under N₂ atmosphere. The reaction mixture was heated at 90 °C for 4 h and cooled to RT after completion of the reaction. The reaction mixture was then filtered through Celite pad and washed with hot dioxane:methanol (1:1) solution (3 x 15 mL) and concentrated to get a yellowish solid residue. The residue was washed with hexane several times and then dissolved in 5 ml of dioxane:methanol (1:1) solution at 60 °C. The solution was cooled on an ice bath and product was precipitated by adding hexane dropwise. The solution was decanted and the product **3** was dried under high vacuum (white solid, 286 mg, 56%). TLC (CHCl₃:MeOH = 9:1) *R_f*: 0.42; ¹H-NMR (400 MHz, *d*₆-DMSO): δ (ppm) 11.76 (s, 1H), 8.73 (s, 1H), 7.90 (d, *J* = 7.6 Hz, 1H), 7.80–7.79 (br, 2H), 7.35–7.30 (m, 2H), 6.23 (t, *J* = 6 Hz, 1H), 5.39 (t, *J* = 4.4 Hz, 1H), 5.30 (d, *J* = 4.4 Hz, 1H), 4.34 (br, 1H), 3.87 (d, *J* = 2.8 Hz, 1H), 3.74–3.70 (m, 2H), 2.30–2.19 (m, 2H); ¹³C NMR (100 MHz, *d*₆-DMSO): δ (ppm) 161.8, 149.8, 139.6, 139.3, 138.1, 135.7, 124.9, 124.6, 123.6, 122.5, 119.7, 108.5, 88.1, 85.5, 70.4, 61.3, 41; HRMS: *m/z* calculated for C₁₇H₁₅N₂O₅S [M-H][−] = 359.0702, found [M-H][−] = 359.0695.

SedU 4. ¹H-NMR (400 MHz, *d*₆-DMSO): δ (ppm) 11.76 (s, 1H), 8.72 (s, 1H), 8.09 (d, *J* = 6.4 Hz, 1H), 7.60 (d, *J* = 2.8 Hz, 1H), 7.30 (m, 1H), 6.23 (t, *J* = 6.4 Hz, 1H), 5.35 (t, *J* = 4.8 Hz, 1H), 5.28 (d, *J* = 4.4 Hz, 1H), 4.35 (m, 1H), 3.87 (d, *J* = 3.2 Hz, 1H), 3.75–3.63 (m, 2H), 2.30–2.16 (m, 2H); ¹³C NMR (100 MHz, *d*₆-DMSO): δ (ppm) 162.3, 149.7, 138.2, 135, 132.2, 129.1, 123.3, 110.4, 88.1, 85.6, 70.4, 61.3, 41; HRMS: *m/z* calculated for C₁₃H₁₄N₂O₅Se [M-H][−] = 356.9995, found [M-H][−] = 357.0009.

BFdUTP 5. 5-Benzofuran modified 2'-deoxyuridine BFdU **2** (70 mg, 0.20 mmol, 1 equiv.) was dissolved in trimethyl phosphate (1 mL) in N₂ atmosphere. Freshly distilled POCl₃ (48 μL, 0.51 mmol, 2.6 equiv.) was added to the reaction mixture at 0–4 °C and stirred for 21 h at ~4 °C. After that, in an ice cold condition a solution of *bis*-tributylammonium pyrophosphate (0.5 M in DMF, 2.03 ml, 5 equiv.) and tributylamine (0.53 ml, 2.23 mmol, 11 equiv.) was added simultaneously to the reaction mixture. After 2 h, 1 M triethylammonium bicarbonate buffer (TEAB, pH 7.5, 10 mL) was added and washed with ethyl acetate (2 x 20 mL). Aqueous layer was collected and evaporated to dryness. Residue was first purified using DEAE sephadex-A25 anion exchange column (10 mM to 1 M TEAB buffer, pH 7.5) and

further purified by reverse-phase flash column chromatography (C18 RediSepRf, 0–50% acetonitrile in 100 mM triethylammonium acetate buffer, pH 7.4, 60 min, flow rate: 7 mL/min). Appropriate fractions were lyophilized to get the desired triphosphate product **5** as a triethylammonium salt (29 mg, 15%). ¹H-NMR (400 MHz, D₂O): δ (ppm) 8.02 (s, 1H), 7.57–7.52 (m, 2H), 7.30 (t, *J* = 7.6 Hz, 1H), 7.21 (t, *J* = 7.2 Hz, 1H), 7.14 (s, 1H), 6.12 (t, *J* = 6.8 Hz, 1H), 4.63 (s, 1H), 4.22 (br, 3H), 2.40 (br, 2H); ¹³C NMR (100 MHz, D₂O): δ (ppm) 161.5, 153.3, 150, 147.7, 136.6, 128.6, 125, 123.2, 121.1, 111.3, 106, 105, 85.9, 85.6 (d, *J* = 8.1 Hz), 70.9, 65.6, 38.3; ³¹P NMR (162 MHz, D₂O): δ (ppm) -9.90 (br, P_γ), -11.56 (br, P_α), -23.12 (br, P_β); HRMS: *m/z* calculated for C₁₇H₁₈N₂O₁₅P₃ [M-H]⁻ = 582.9926, found [M-H]⁻ = 582.9931.

BTdUTP 6. The nucleotide analog **6** was synthesized as mentioned above using BTdU **3** (70 mg, 0.19 mmol). Yield (14 mg, 8%). ¹H-NMR (400 MHz, D₂O): δ (ppm) 8.06 (s, 1H), 7.90–7.85 (m, 2H), 7.68 (s, 1H), 7.41–7.35 (m, 2H), 6.16 (t, *J* = 6.8 Hz, 1H), 4.66 (s, 1H), 4.26 (br, 3H), 2.43–2.40 (m, 2H); ¹³C NMR (100 MHz, D₂O): δ (ppm) 162.8, 150.1, 139.1, 139, 137.1, 133.4, 124.7, 124.6, 123.7, 121.9, 121, 109.3, 86.1, 85.8 (d, *J* = 8.5 Hz), 71, 65.7, 38.7; ³¹P NMR (162 MHz, D₂O): δ (ppm) -10.78 (br, P_γ), -11.62 (d, *J* = 20.08 Hz, P_α), -23.23 (br, P_β); HRMS: *m/z* calculated for C₁₇H₁₈N₂O₁₄P₃S [M-H]⁻ = 598.9697, found [M-H]⁻ = 598.9687.

SedUTP 7. The nucleotide analog **7** was synthesized as mentioned above using SedU **4** (60 mg, 0.16 mmol). Yield (15 mg, 9%). ¹H-NMR (400 MHz, D₂O): δ (ppm) 8.08–8.05 (m, 2H), 7.63 (s, 1H), 7.31 (s, 1H), 6.21 (br, 1H), 4.60 (s, 1H), 4.17 (br, 3H), 2.36 (br, 2H); ¹³C NMR (100 MHz, D₂O): δ (ppm) 156, 150.2, 136, 134.5, 132.2, 129.6, 125.4, 111.6, 85.8, 85.7, 70.8, 65.6, 38.9; ³¹P NMR (162 MHz, D₂O): δ (ppm) -10.15 (br, P_γ), -11.61 (d, *J* = 19.4 Hz, P_α), -23.24 (br, P_β); HRMS: *m/z* calculated for C₁₃H₁₆N₂O₁₄P₃Se [M-H]⁻ = 596.8985, found [M-H]⁻ = 596.8993.

2.4.4 PEX with nucleotide analogs 5–7

Reaction using KlenTaq and KOD DNA polymerases

Primer•template duplexes were prepared by annealing 5'-FAM-labeled primer **P1** (50 μM) and templates **T1–T5** (55 μM) in a reaction buffer (50 mM Tris-HCl, 16 mM (NH₄)₂SO₄, 2.5 mM MgCl₂, 0.1% (v/v) Tween 20, pH 8.0) at 95 °C for 3 min. The samples were left to come to room temperature and were placed on an ice bath for 30 min. PEXs were performed at 55 °C in a reaction buffer (50 mM Tris-HCl, 16 mM (NH₄)₂SO₄, 2.5 mM MgCl₂, 0.1% (v/v) Tween 20, pH 8.0) containing 1.50 μM of primer•template duplex, 150 μM of dATP, dGTP, dCTP, dTTP or modified dU*TP (**5–7**) and 3 nM of KlenTaq/KOD DNA polymerase in a 20

μL reaction volume. After 45 min, the reaction was quenched by adding 140 μL of denaturing loading buffer (8.3 M urea in 10 mM Tris-HCl, 100 mM EDTA, 0.05% bromophenol blue, pH 8). The samples were heated at 95 °C for 5 min and then cooled on an ice bath. 5 μL (0.94 pmol) of the sample was loaded on an analytical 18% denaturing polyacrylamide gel and was electrophoresed at 30 W for 2.5 h. The gel was imaged using Typhoon gel scanner at FAM wavelength. Control reactions to study misincorporation were performed in the presence of primer•template duplex, enzyme, dATP, dGTP and dCTP but without dTTP or modified dU*TPs.

Reaction using Vent(exo-) DNA polymerase

A reaction mixture containing 2 μM primer **P1** and 2.5 μM templates **T1–T5**, 300 μM of dATP, dGTP, dCTP, dTTP or modified dU*TP (**5–7**), 1U/ μL of Vent(exo-) in 1x Thermopol reaction buffer (20 mM Tris-HCl, 10 mM $(\text{NH}_4)_2\text{SO}_4$, 10 mM KCl, 2 mM MgSO_4 , 0.1% Triton X-100, pH 8.8) was incubated at 60 °C. After 40 minutes, 180 μL of denaturing loading buffer (8.3 M urea in 10 mM Tris-HCl, 100 mM EDTA, 0.05% bromophenol blue, pH 8) was added to the reaction mix. The samples were heated at 95 °C for 5 min and cooled on an ice bath. 5 μL of the sample (1 pmol) was loaded on an analytical 18% denaturing polyacrylamide gel and was electrophoresed at 30 W for 2.5 h. The gel was imaged using Typhoon gel scanner at FAM wavelength. Control reactions to study misincorporation were performed in the presence of primer•template duplex, enzyme, dATP, dGTP and dCTP but without dTTP or modified dU*TPs.

2.4.5 Single nucleotide incorporation assay

5'-FAM-labeled primer **P2** (50 μM) and template **T6** (55 μM) were annealed in a reaction buffer (50 mM Tris-HCl, 16 mM $(\text{NH}_4)_2\text{SO}_4$, 2.5 mM MgCl_2 , 0.1% (v/v) Tween 20, pH 8.0) at 95 °C for 3 min. The sample was cooled to room temperature and kept on an ice bath for 30 min. Reactions were carried out at 55 °C in a reaction buffer (50 mM Tris-HCl, 16 mM $(\text{NH}_4)_2\text{SO}_4$, 2.5 mM MgCl_2 , 0.1 % (v/v) Tween 20, pH 8.0) containing 1.5 μM of the annealed primer•template duplex, 150 μM of dTTP or modified dU*TP (**5–7**), 3 nM KlenTaq/KOD DNA polymerase in a 20 μL reaction volume. Reaction aliquots (10 μL) after 15 and 45 min were mixed with 35 μL of denaturing loading buffer (8.3 M urea in 10 mM Tris-HCl, 100 mM EDTA, 0.05% bromophenol blue, pH 8) and heated at 95 °C for 5 min and cooled on an ice bath. 5 μL (1.6 pmol) of the sample was loaded on an analytical 20% denaturing polyacrylamide gel and electrophoresed at 30 W for 2.5 h and imaged using Typhoon gel scanner at FAM wavelength.

Large-scale single nucleotide incorporation of BFdUTP 5 using KlenTaq DNA polymerase

Single nucleotide incorporation reaction was performed in 840 μ L reaction volume using 150 μ M BFdUTP 5, 1.50 μ M annealed primer•template duplex (**P2•T6** or **P3•T6**) and 3 nM KlenTaq in 50 mM Tris-HCl buffer (pH 8.0, 16 mM (NH₄)₂SO₄, 2.5 mM MgCl₂, 0.1 % (v/v) Tween 20). The reaction solution was incubated at 55 °C for 45 min. The reaction was heated at 95 °C for 5 min and cooled on an ice bath. The solution was completely concentrated using SpeedVac and 300 μ L of 5 M ammonium acetate and 700 μ L of 100 % ethanol were added. The sample was kept at -40 °C for 12 h. The sample was centrifuged at 4 °C at 5000 rpm for 30 min. DNA pellet was collected and washed 3 times with 100% ethanol and dried at room temperature. Nine such reactions were performed using primers **P2** and **P3** to produce BFdU-modified ON **P2*** and **P3***. DNA pellets were pooled together and dissolved in 30 μ L of denaturing loading buffer (8.3 M urea in 10 mM Tris-HCl, 100 mM EDTA, 0.05% bromophenol blue, pH 8) and loaded onto a preparative 20% polyacrylamide gel and electrophoresed under denaturing conditions. The gel was UV shadowed and bands corresponding to the products were excised. The ON products were extracted from the gel using 0.5 M ammonium acetate followed by desalting using Sep-Pak C18 cartridge. Under these conditions approximately 70% of pure **P2*** and **P3*** was isolated. The mass of ONs was confirmed by ESI mass analysis in negative mode by direct injection of 200 pmol of DNA ONs in 50% ACN (LCMS grade) in 10 mM triethylamine and 100 mM hexafluoro-2-propanol.

Template dependent site-specific labeling

In a 1.5 mL Eppendorf tube 150 μ M BFdUTP 5 and 1.50 μ M annealed primer•template duplex (**P2•T6**) in 50 mM Tris-HCl (pH 8.0), 16 mM (NH₄)₂SO₄, 2.5 mM MgCl₂, 0.1 % (v/v) Tween 20 was incubated with 3 nM KlenTaq DNA polymerase at 55 °C for 45 min. The total reaction volume was 840 μ L (containing 1.26 nmol of primer•template). The reaction was heated at 95 °C for 5 min and cooled on an ice bath. The solution was concentrated nearly to dryness using SpeedVac and 300 μ L of 5 M ammonium acetate and 700 μ L of 100 % ethanol were added. The sample was kept at -40 °C for 12 h and was centrifuged at 4 °C at 5000 rpm for 30 min. The pellet was washed 3 times with 100% ethanol to remove unreacted nucleotide analog and dried at room temperature. The pellet was re-dissolved in 663.6 μ L of reaction buffer (50 mM Tris-HCl, 16 mM (NH₄)₂SO₄, 2.5 mM MgCl₂, 0.1 % (v/v) Tween 20, pH 8.0) and annealed at 95 °C for 3 min. Reaction mixture

was cooled to room temperature and kept on an ice bath for 30 min. To the same reaction tube 126 μL of 1 mM dNTPs (dATP, dGTP, dCTP, dTTP) and 50.4 μL 50 nM KOD/KlenTaq DNA polymerase were added and incubated at 55 $^{\circ}\text{C}$ for 45 min. Final concentration of dNTPs was 150 μM and enzyme was 3 nM. The reaction was heated at 95 $^{\circ}\text{C}$ for 5 min and cooled on an ice bath. In order to monitor the progress of the reaction, after each step 10 μL of the reaction mixture was taken and mixed with 35 μL of denaturing loading buffer (8.3 M urea in 10 mM Tris-HCl, 100 mM EDTA, 0.05% bromophenol blue, pH 8). 5 μL samples were loaded onto a sequencing 15% denaturing polyacrylamide gel and electrophoresed at 30 W for 2.5 hours. The gel was imaged using Typhoon gel scanner at FAM wavelength.

Remaining solution was concentrated using SpeedVac and 300 μL of 5 M ammonium acetate and 700 μL of 100 % ethanol were added. The sample was kept at -40 $^{\circ}\text{C}$ for 12 h and was centrifuged at 4 $^{\circ}\text{C}$ at 5000 rpm for 30 min. DNA pellet was collected and washed 3 times with 100% ethanol and dried at room temperature. Four such reactions were performed and combined DNA pellets were dissolved in 30 μL of denaturing loading buffer (8.3 M urea in 10 mM Tris-HCl, 100 mM EDTA, 0.05% bromophenol blue, pH 8) and loaded onto a preparative 15% polyacrylamide gel and electrophoresed under denaturing conditions. The gel was UV shadowed and the band corresponding to the full length product was excised. The ON product was extracted from the gel using 0.5 M ammonium acetate followed by desalting using Sep-Pak C18 cartridge. The identity of the labeled product was confirmed by ESI mass analysis.

2.4.6 Fluorescence binding assay to determine binary complex formation

BFdU-modified ON **P3*** (10 μM) was annealed to a template **T7** (10.5 μM) at 95 $^{\circ}\text{C}$ for 3 min in a binding buffer (10 mM Tris-HCl, 50 mM KCl, 5 mM MgCl_2 , pH 7.9). The stock solution was cooled to room temperature and kept on an ice bath for 30 min. A series of primer•template duplex (**P3*•T7**, 50 nM) samples containing increasing concentrations of KlenTaq (0–5 μM) in the binding buffer was incubated at 25 $^{\circ}\text{C}$ for 45 min. The total volume of each sample was 200 μL . Fluorescence spectrum of individual samples was recorded at 25 $^{\circ}\text{C}$ by exciting the samples at 330 nm with excitation and emission slit widths of 7 nm and 8 nm, respectively. The amount of glycerol in samples was low (0–1.5%), which should not affect the fluorescence of the BFdU probe. Beyond 1.5% of glycerol in the sample results in slight increase in fluorescence intensity of the duplex itself, possibly due to rigidification of the probe in a more viscous medium (data not shown). Blank spectrum of a sample without

BFdU-modified primer•template duplex (**P3*•T7**) was subtracted from each spectrum. Fluorescence experiments were performed in triplicate. Increasing in fluorescence intensity at emission maximum (430 nm) was fit to Hill equation (Origin 8.5) by plotting normalized fluorescence intensity (F_N) versus KlenTaq concentration. A Hill coefficient (n) of nearly 1 was observed for binding of BFDU-labeled primer•template duplex (**P3*•T7**) to KlenTaq DNA polymerase. R^2 was close to unity.^{17,59}

$$F_N = \frac{F_i - F_0}{F_s - F_0}$$

F_i is the fluorescence intensity measured at each protein concentration. F_0 and F_s are the fluorescence intensity measured in the absence of protein and at saturation point. n is the Hill coefficient measuring the degree of cooperativity in binding. K_d is the apparent dissociation constant.

$$F_N = F_0 + (F_s - F_0) \left(\frac{[\text{KlenTaq}]^n}{[K_d]^n + [\text{KlenTaq}]^n} \right)$$

2.4.7 Fluorescence detection of incorporation of the nucleotide analog by KlenTaq

Primer **P3** (10 μ M) or 3'-ddCMP-modified primer **P5** (10 μ M) was annealed to a template **T7** (10.5 μ M) at 95 °C for 3 min in a buffer (10 mM Tris-HCl, 50 mM KCl, 5 mM MgCl₂, pH 7.9). The solution was cooled to room temperature and kept on an ice bath for 30 min. A series of primer-template duplex (**P3•T7**, 200 nM) samples containing increasing concentrations of KlenTaq (0–200 nM) in the above buffer was incubated at 25 °C for 45 min. To each sample, 5 μ L of BFDUTP 5 (final conc. 50 nM) was added and incubated at 55 °C for 15 min. The total volume was maintained at 200 μ L and glycerol concentration was less than 0.05%. Samples were cooled to 25 °C and fluorescence was measured at 25 °C. As a control experiment, **P5•T7** duplex (200 nM) was incubated with KlenTaq (200 nM) at 25 °C for 45 min and BFDUTP (final conc. 50 nM) was added, and kept at 55 °C for 15 minutes. Another control experiment was performed using the same procedure with sample containing only KlenTaq (200 nM) or primer•template duplex. All samples were excited at 330 nm with excitation and emission slit widths of 6 nm and 7 nm, respectively. Measurements were performed in triplicate.

2.4.8 EMSA to determine binary complex formation

5'-FAM labeled primer **P2** (10 μ M) or 5'-FAM, 3'-BFdU-labeled primer **P2*** (10 μ M) was annealed to a template **T7** (10.5 μ M) at 95 °C for 3 min in a binding buffer (10 mM Tris-HCl, 50 mM KCl, 5 mM MgCl₂, pH 7.9). The solution was cooled to room temperature and kept on an ice bath for 30 min. A series of primer-template duplex (**P2*•T7/P2•T7**, 200 nM) samples containing increasing concentrations of KlenTaq (0–4 μ M) in the binding buffer was incubated at 25 °C for 45 min. The total volume was 20 μ L and contained 0–3 % glycerol. To each sample, 10 μ L of native loading buffer (50% sucrose in 10 mM Tris-HCl, 100 mM EDTA, 0.05% bromophenol blue, pH 8) was added and 8 μ L was loaded on a native 12% polyacrylamide gel. The electrophoresis was carried out in 1x TBE buffer containing 5 mM MgCl₂ at 4 °C. The gel was imaged using Typhoon gel scanner at FAM wavelength. The binding assay was performed in triplicate. The binary complex of DNA duplex and enzyme and the duplex had marked different motilities on the gel. The band intensity of the binary complex at different KlenTaq concentration was obtained by using Fiji Image analysis software. A plot of fraction of binary complex versus KlenTaq concentration was fit to Hill equation (Origin 8.5) to obtain K_d . Hill coefficient n was found to be 1.3 to 1.5. R^2 was close to unity.

2.4.9 Competitive displacement assay

By Fluorescence

Duplex **P3*•T7** stock (10 μ M) was prepared as above. Samples of binary complex (**P3*•T7**, 50 nM and KlenTaq, 1 μ M) were incubated with increasing concentrations of the unlabeled primer-template duplex (**P3•T7**, 0–1 μ M) in a binding buffer (10 mM Tris-HCl, 50 mM KCl, 5 mM MgCl₂, pH 7.9). The total volume of each sample was maintained at 200 μ L. The samples were incubated for 30 minutes at 25 °C and were subjected to fluorescence measurement. Samples were excited at 330 nm with excitation and emission slit widths of 7 nm and 8 nm, respectively.

By EMSA

Stocks (10 μ M) of primer-template duplexes **P2*•T7** and **P2•T7** were prepared as mentioned above. A set of binary complex samples of **P2*•T7** duplex (0.2 μ M) and KlenTaq (1.2 μ M) was incubated with increasing concentration of unlabeled primer-template duplex **P3•T7** (0–4 μ M) in the binding buffer (10 mM Tris-HCl, 50 mM KCl, 5 mM MgCl₂, pH 7.9) in a final volume of 20 μ L. The samples were incubated for 45 minutes at 25 °C and 10 μ L of native

loading buffer (50% sucrose in 10 mM Tris-HCl, 100 mM EDTA, 0.05% bromophenol blue, pH 8) was added. The samples (8 μ L) were loaded on a native 12% polyacrylamide gel and run in 1x TBE supplemented with 5 mM MgCl₂. The gel was electrophoresed at 4 °C and imaged using Typhoon gel scanner at FAM wavelength. Similarly, a set of binary complex samples of **P2•T7** duplex (0.2 μ M) and KlenTaq (1.2 μ M) was incubated with increasing concentration of the BFdU-labeled primer-template duplex **P3*•T7** (0–4 μ M) in the binding buffer. EMSA was performed as mentioned above.

2.4.10 Crystallization, data collection and refinement

The crystallization trials to obtain the ternary complex of KlenTaq with primer•template duplex and modified dNTPs were setup as illustrated in Figure 5A. A 1:1 solution of primer **P4** and template **T8** (3 mM) was annealed at 95 °C for 5 min and slowly cooled to 4 °C over 30 min. The sample was diluted with a storage buffer (20 mM Tris-HCl pH 7.5, 1 mM EDTA, 1 mM β -mercaptoethanol, 150 mM NaCl), and a 1:1 solution (1 M) of MgCl₂ and MnCl₂, KlenTaq (20.9 mg/ml) and ddCTP (10 mM) were added. The reaction solution was incubated at 55 °C for 45 min, after which the modified nucleotide (e.g. BFdUTP, 10 mM) was added and was further incubated at 30 °C for 45 min. The final concentration of KlenTaq, primer•template duplex, ddCTP and modified nucleotide was 6.3 mg/ml, 0.12 mM, 0.31 mM and 1.03 mM, respectively (molar ratio 1:1.2:3:10). The final concentration of MgCl₂ in the above sample was 20 mM. The sample was filtered using a 0.1 μ M sterile filter and crystals were grown by hanging drop vapor diffusion method by mixing with the reservoir solution (commercial screens) in different ratios. Diffraction quality crystals with BFdUTP were obtained in a drop composed of a 1:2 ratio of protein to reservoir. Reservoir composition was 16% PEG 6000, 0.1 M Tris pH 8.0 and 0.2 M magnesium formate. The crystals were cryoprotected in the reservoir solution containing 20% ethylene glycol and frozen in liquid nitrogen.

Data was collected at the Swiss Light Source (SLS) of the Paul Scherrer Institute (PSI) in Villigen, Switzerland. Data collection and refinement statistics are summarized in Table S4. Data was processed using the XDS package⁹⁰ and the XDGUI graphical interface (<https://strucbio.biologie.uni-konstanz.de/xdswiki/index.php/XDGUI>). The structure was solved by rigid body refinement of the native KlenTaq ternary structure (PDB: 3RTV). Structure refinement was done with phenix.⁹¹ Model building was done in Coot^{92,93} and model quality was evaluated by the MolProbity⁹⁴ web server (<http://molprobity.biochem.duke.edu/>)

and the PDB validation server (<https://validate-rcsb-1.wwpdb.org>). Figures were created with PyMOL. For superimposition in figures and RMSD value calculation (after outlier rejection), structures were superimposed in PyMOL using the “align” command. The atomic coordinates and structure factors have been deposited on the Protein Data Bank with the PDB code 7OWF.

2.5 References

1. Kranaster, R.; Marx, A. Engineered DNA polymerases in biotechnology. *ChemBioChem* **2010**, *11*, 2077–2084.
2. Chen, T.; Hongdilokkul, N.; Liu, Z.; Thirunavukarasu, D.; Romesberg, F. E. The expanding world of DNA and RNA. *Curr Opin Chem Biol* **2016**, *34*, 80–87.
3. Hocek, M. Enzymatic synthesis of base-functionalized nucleic acids for sensing, cross-linking, and modulation of protein–DNA binding and transcription. *Acc. Chem. Res.* **2019**, *52*, 1730–1737.
4. Klöcker, N.; Weissenboeck, F. P.; Rentmeister, A. Covalent labeling of nucleic acids. *Chem. Soc. Rev.* **2020**, *49*, 8749–8773.
5. McKenzie, L. K.; El-Khoury, R.; Thorpe, J. D.; Damha, M. J.; Hollenstein, M. Recent progress in non-native nucleic acid modifications. *Chem. Soc. Rev.* **2021**, *50*, 5126–5164.
6. Jäger, S.; Rasched, G.; Kornreich-Leshem, H.; Engeser, M.; Thum, O.; Famulok, M. A versatile toolbox for variable DNA functionalization at high density. *J. Am. Chem. Soc.* **2005**, *127*, 15071–15082.
7. Burley, G. A.; Gierlich, J.; Mofid, M. R.; Nir, H.; Tal, S.; Eichen, Y.; Carell, T. Directed DNA metallization. *J. Am. Chem. Soc.* **2006**, *128*, 1398–1399.
8. Srivatsan, S. G.; Tor, Y. Fluorescent pyrimidine ribonucleotide: synthesis, enzymatic incorporation, and utilization. *J. Am. Chem. Soc.* **2007**, *129*, 2044–2053.
9. Bentley, D. R.; Balasubramanian, S.; Swerdlow, H. P.; Smith, G. P.; Milton, J.; Brown, C. G.; Hall, K. P.; Evers, D. J.; Barnes, C. L.; Bignell, H. R.; et al. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **2008**, *456*, 53–59.
10. Chen, F.; Gaucher, E. A.; Leal, N. A.; Hutter, D.; Havemann, S. A.; Govindarajan, S.; Ortlund, E. A.; Benner, S. A. Reconstructed evolutionary adaptive paths give polymerases accepting reversible terminators for sequencing and SNP detection. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 1948–1953.
11. Ramsay, N.; Jemth, A.-S.; Brown, A.; Crampton, N.; Dear, P.; Holliger, P. CyDNA: synthesis and replication of highly Cy-dye substituted DNA by an evolved polymerase. *J. Am. Chem. Soc.* **2010**, *132*, 5096–5104.
12. Winz, M. L.; Linder, E. C.; André, T.; Becker, J.; Jäschke, A. Nucleotidyl transferase assisted DNA labeling with different click chemistries. *Nucleic Acids Res.* **2015**, *43*, e110.
13. Matyašovský, J.; Perlíková, P.; Malnuit, V.; Pohl, R.; Hocek, M. 2-Substituted dATP derivatives as building blocks for polymerase-catalyzed synthesis of DNA modified in the minor groove. *Angew. Chem. Int. Ed.* **2016**, *55*, 15856–15859.
14. Muttach, F.; Rentmeister, A. A biocatalytic cascade for versatile one-pot modification of mRNA starting from methionine analogues. *Angew. Chem. Int. Ed.* **2016**, *55*, 1917–1920.

15. Zhou, C. Y.; Alexander, S. C.; Devaraj, N. K. Fluorescent turn-on probes for wash-free mRNA imaging via covalent site-specific enzymatic labeling. *Chem. Sci.* **2017**, *8*, 7169–7173.
16. Sarac, I.; Hollenstein, M. Terminal deoxynucleotidyl transferase in the synthesis and modification of nucleic acids. *ChemBioChem* **2019**, *20*, 860–871.
17. George J. T.; Srivatsan, S. G. Responsive fluorescent nucleotides serve as efficient substrates to probe terminal uridylyl transferase. *Chem. Commun.* **2020**, *56*, 12319–12322.
18. Hottin, A.; Marx, A. Structural insights into the processing of nucleobase-modified nucleotides by DNA polymerases. *Acc. Chem. Res.* **2016**, *49*, 418–427.
19. Betz, K.; Malyshev, D. A.; Lavergne, T.; Welte, W.; Diederichs, K.; Dwyer, T. J.; Ordoukhanian, P.; Romesberg, F. E.; Marx, A. KlenTaq polymerase replicates unnatural base pairs by inducing a Watson-Crick geometry. *Nat. Chem. Biol.* **2012**, *8*, 612–614.
20. Ju, J.; Kim, D. H.; Bi, L.; Meng, Q.; Bai, X.; Li, Z.; Li, X.; Marma, M. S.; Shi, S.; Wu, J.; Edwards, J. R.; Romu, A.; Turro, N. J. Four-color DNA sequencing by synthesis using cleavable fluorescent nucleotide reversible terminators. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 19635–19640.
21. Obeid, S.; Yulikov, M.; Jeschke, G.; Marx, A. Enzymatic synthesis of multiple spin-labeled DNA. *Angew. Chem. Int. Ed.* **2008**, *47*, 6782–6785.
22. Jarchow-Choy, S. K.; Krueger, A. T.; Liu, H.; Gao, J.; Kool, E. T. Fluorescent xDNA nucleotides as efficient substrates for a template-independent polymerase. *Nucleic Acids Res.* **2011**, *39*, 1586–1594.
23. Riedl, J.; Pohl, R.; Rulišek, L.; Hocek, M. Synthesis and photophysical properties of biaryl-substituted nucleos(t)ides. Polymerase synthesis of DNA probes bearing solvatochromic and pH-sensitive dual fluorescent and ^{19}F NMR labels. *J. Org. Chem.* **2012**, *77*, 1026–1044.
24. McCoy, L. S.; Shin, D.; Tor, Y. Isomorphic emissive GTP surrogate facilitates initiation and elongation of in vitro transcription reactions. *J. Am. Chem. Soc.* **2014**, *136*, 15176–15184.
25. Dziuba, D.; Jurkiewicz, P.; Cebecauer, M.; Hof, M.; Hocek, M. A rotational BODIPY nucleotide: an environment-sensitive fluorescence-lifetime probe for DNA interactions and applications in live-cell microscopy. *Angew. Chem. Int. Ed.* **2016**, *55*, 174–178.
26. Manna, S.; Srivatsan, S. G. Synthesis and enzymatic incorporation of a responsive ribonucleoside probe that enables quantitative detection of metallo-base pairs. *Org. Lett.* **2019**, *21*, 4646–4650.
27. Magriñá, I.; Toldrà, A.; Campàs, M.; Ortiz, M.; Simonova, A.; Katakis, I.; Hocek, M.; O'Sullivan, C. K. Electrochemical genosensor for the direct detection of tailed PCR amplicons incorporating ferrocene labelled dATP. *Biosens. Bioelectron.* **2019**, *134*, 76–82.
28. Gramlich, P. M. E.; Wirges, C. T.; Manetto, A.; Carell, T. Postsynthetic DNA modification through the copper-catalyzed azide-alkyne cycloaddition reaction. *Angew. Chem. Int. Ed.* **2008**, *47*, 8350–8358.
29. Samanta, A.; Krause, A.; Jäschke, A. A modified dinucleotide for site-specific RNA-labelling by transcription priming and click chemistry. *Chem. Commun.* **2014**, *50*, 1313–1316.
30. Someya, T.; Ando, A.; Kimoto, M.; Hirao, I. Site-specific labeling of RNA by combining genetic alphabet expansion transcription and copper-free click chemistry. *Nucleic Acids Res.* **2015**, *43*, 6665–6676.

31. Reisacher, U.; Ploschik, D.; Rönicke, F.; Cserép, G. B.; Kele, P.; Wagenknecht, H.-A. Copper-free dual labeling of DNA by triazines and cyclopropenes as minimal orthogonal and bioorthogonal functions, *Chem. Sci.* **2019**, *10*, 4032–4037.
32. George, J. T.; Azhar, M.; Aich, M.; Sinha, D.; Ambi, U. B.; Maiti, S.; Chakraborty, D.; Srivatsan, S. G. Terminal uridylyl transferase mediated site-directed access to clickable chromatin employing CRISPR-dCas9. *J. Am. Chem. Soc.* **2020**, *142*, 13954–13965.
33. Salic, A.; Mitchison, T. J. A chemical method for fast and sensitive detection of DNA synthesis in vivo. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 2415–2420.
34. Curanovic, D.; Cohen, M.; Singh, I.; Slagle, C. E.; Leslie, C. S.; Jaffrey, S. R. Global profiling of stimulus-induced polyadenylation in cells using a poly(A) trap. *Nat. Chem. Biol.* **2013**, *9*, 671–673.
35. Rieder, U.; Luedtke, N. W. Alkene–tetrazine ligation for imaging cellular DNA. *Angew. Chem. Int. Ed.* **2014**, *53*, 9168–9172.
36. Sawant, A. A.; Tanpure, A. A.; Mukherjee, P. P.; Athavale, S.; Kelkar, A.; Galande, S.; Srivatsan, S. G. A versatile toolbox for posttranscriptional chemical labeling and imaging of RNA. *Nucleic Acids Res.* **2016**, *44*, e16.
37. Zhang, Y.; Kleiner, R. E. A metabolic engineering approach to incorporate modified pyrimidine nucleosides into cellular RNA. *J. Am. Chem. Soc.* **2019**, *141*, 3347–3351.
38. Nainar, S.; Cuthbert, B. J.; Lim, N. M.; England, W. E.; Ke, K.; Sophal, K.; Quechol, R.; Mobley, D. L.; Goulding, C. W.; Spitale, R. C. An optimized chemical-genetic method for cell-specific metabolic labeling of RNA. *Nat. Methods* **2020**, *17*, 311–318.
39. Vyas, R.; Zahurancik, W. J.; Suo, Z. Structural basis for the binding and incorporation of nucleotide analogs with L-stereochemistry by human DNA polymerase λ . *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, E3033–E3042.
40. Bergen, K.; Steck, A. L.; Strütt, S.; Baccaro, A.; Welte, W.; Diederichs, K.; Marx, A. Structures of *KlenTaq* DNA polymerase caught while incorporating C5-modified pyrimidine and C7-modified 7-deazapurine nucleoside triphosphates. *J. Am. Chem. Soc.* **2012**, *134*, 11840–11843.
41. Obeid, S.; Bußkamp, H.; Welte, W.; Diederichs, K.; Marx, A. Interactions of non-polar and “Click-able” nucleotides in the confines of a DNA polymerase active site. *Chem. Commun.* **2012**, *48*, 8320–8322.
42. Obeid, S.; Baccaro, A.; Welte, W.; Diederichs, K.; Marx, A. Structural basis for the synthesis of nucleobase modified DNA by *Thermus aquaticus* DNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 21327–21331.
43. Sinkeldam, R. W.; Greco, N. J.; Tor, Y. Fluorescent analogs of biomolecular building blocks: design, properties, and applications. *Chem. Rev.* **2010**, *110*, 2579–2619.
44. Wachowius, F.; Höbartner, C. Chemical RNA modifications for studies of RNA structure and dynamics. *ChemBioChem* **2010**, *11*, 469–480.
45. Xu, W.; Chan, K. M.; Kool, E. T. Fluorescent nucleobases as tools for studying DNA and RNA. *Nat. Chem.* **2017**, *9*, 1043–1055.
46. Benoît, Y. M.; Dziuba, D.; Benhida, R.; Demchenko, A. P.; Burger, A. Probing of nucleic acid structures, dynamics, and interactions with environment-sensitive fluorescent labels. *Front. Chem.* **2020**, *8*, 112.
47. Sholokh, M.; Sharma, R.; Shin, D.; Das, R.; Zaporozhets, O. A.; Tor, Y.; Mély, Y. Conquering 2-aminopurine's deficiencies: highly emissive isomorphous guanosine surrogate faithfully monitors guanosine conformation and dynamics in DNA. *J. Am. Chem. Soc.* **2015**, *137*, 3185–3188.

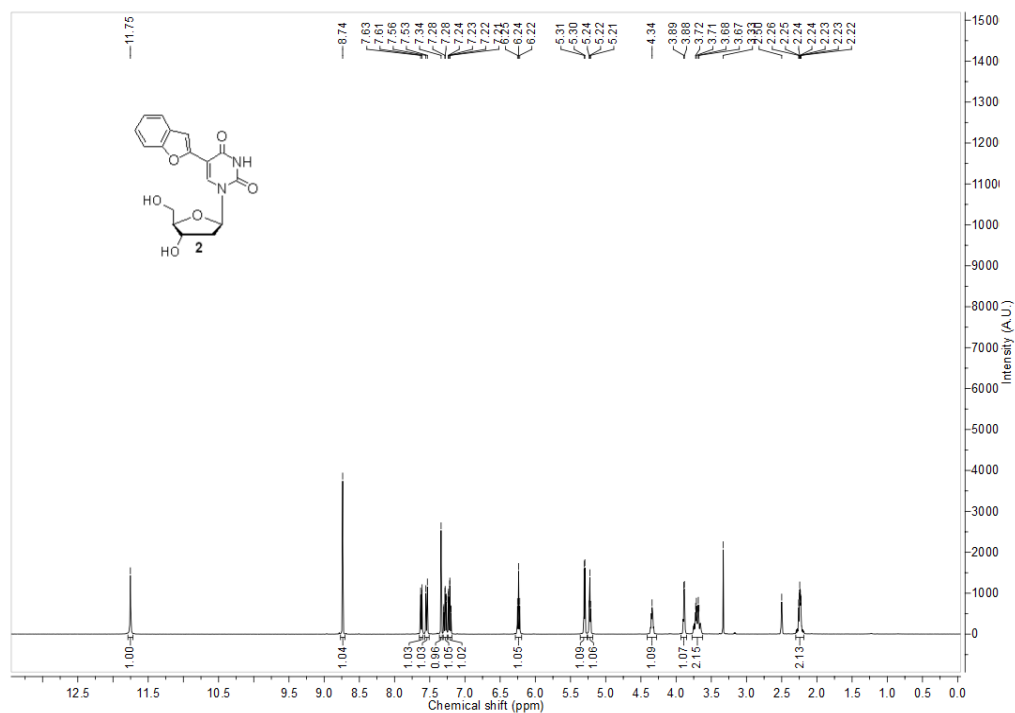
48. Sproviero, M.; Fadock, K. L.; Witham, A. A.; Manderville, R. A. Positional impact of fluorescently modified G-tetrads within polymorphic human telomeric G-quadruplex structures. *ACS Chem. Biol.* **2015**, *10*, 1311–1318.
49. Wranne, M. S.; Füchtbauer, A. F.; Dumat, B.; Bood, M.; El-Sagheer, A. H.; Brown, T.; Gradén, H.; Grøtli, M.; Wilhelmsson, L. M. Toward complete sequence flexibility of nucleic acid base analogue FRET. *J. Am. Chem. Soc.* **2017**, *139*, 9271–9280.
50. Eubanks, C. S.; Forte, J. E.; Kapral, G. J.; Hargrove, A. E. Small molecule-based pattern recognition to classify RNA structure. *J. Am. Chem. Soc.* **2017**, *139*, 409–416.
51. Manna, S.; Sarkar, D.; Srivatsan, S. G. A dual-app nucleoside probe provides structural insights into the human telomeric overhang in live cells. *J. Am. Chem. Soc.* **2018**, *140*, 12622–12633.
52. Karimi, A.; Börner, R.; Mata, G.; Luedtke, N. W. A highly fluorescent nucleobase molecular rotor. *J. Am. Chem. Soc.* **2020**, *142*, 14422–14426.
53. Tanpure, A. A.; Srivatsan, S. G. A microenvironment-sensitive fluorescent pyrimidine ribonucleoside analogue: synthesis, enzymatic incorporation, and fluorescence detection of a DNA abasic site. *Chem. Eur. J.* **2011**, *17*, 12820–12827.
54. Dziuba, D.; Pospíšil, P.; Matyašovský, J.; Brynda, J.; Nachtigallová, D.; Rulíšek, L.; Pohl, R.; Hof, M.; Hocek, M. Solvatochromic fluorene-linked nucleoside and DNA as color-changing fluorescent probes for sensing interactions. *Chem. Sci.* **2016**, *7*, 5775–5785.
55. Li, Y.; Fin, A.; McCoy, L.; Tor, Y. Polymerase-mediated site-specific incorporation of a synthetic fluorescent isomorphous G surrogate into RNA. *Angew. Chem. Int. Ed.* **2017**, *56*, 1303–1307.
56. Sandin, P.; Stengel, G.; Ljungdahl, T.; Börjesson, K.; Macao, B.; Wilhelmsson, L. M. Highly efficient incorporation of the fluorescent nucleotide analogs tC and tC⁰ by Klenow fragment. *Nucleic Acids Res.* **2009**, *37*, 3924–3933.
57. Nuthanakanti, A.; Ahmed, I.; Khatik, S. Y.; Saikrishnan, K.; Srivatsan S. G. Probing G-quadruplex topologies and recognition concurrently in real time and 3D using a dual-app nucleoside probe. *Nucleic Acids Res.* **2019**, *47*, 6059–6072.
58. Nuthanakanti, A.; Boerneke, M. A.; Hermann, T.; Srivatsan S. G. Structure of the ribosomal decoding site RNA containing a selenium-modified responsive fluorescent ribonucleoside probe. *Angew. Chem. Int. Ed.* **2017**, *56*, 2640–2644.
59. Tanpure, A. A.; Srivatsan, S. G. Conformation-sensitive nucleoside analogues as topology-specific fluorescence turn-on probes for DNA and RNA G-quadruplexes. *Nucleic Acids Res.* **2015**, *43*, e149.
60. Haidekker, M. A.; Theodorakis, E. A. Molecular rotors—fluorescent biosensors for viscosity and flow. *Org. Biomol. Chem.* **2007**, *5*, 1669–1678.
61. Förster, T.; Hoffmann, G. Die viskositätsabhängigkeit der fluoreszenzquantenausbeuten einiger farbstoffsysteme. *Z. Phys. Chem.* **1971**, *75*, 63–76.
62. Sinkeldam, R. W.; Wheat, A. J.; Boyaci, H.; Tor, Y. Emissive nucleosides as molecular rotors. *ChemPhysChem* **2011**, *12*, 567–570.
63. Noé, M. S.; Sinkeldam, R. W.; Tor, Y. Oligodeoxynucleotides containing multiple thiophene-modified isomorphous fluorescent nucleosides. *J. Org. Chem.* **2013**, *78*, 8123–8128.
64. Sawai, H.; Ozaki-Nakamura, A.; Mine, M.; Ozaki, H. Synthesis of new modified DNAs by hyperthermophilic DNA polymerase: substrate and template specificity of functionalized thymidine analogues bearing an sp³-hybridized carbon at the C5 α -position for several DNA polymerases. *Bioconjugate Chem.* **2002**, *13*, 309–316.

65. Ghadessy, F. J.; Ramsay, N.; Boudsocq, F.; Loakes, D.; Brown, A.; Iwai, S.; Vaisman, A.; Woodgate, R.; Holliger, P. Generic expansion of the substrate spectrum of a DNA polymerase by directed evolution. *Nat. Biotechnol.* **2004**, *22*, 755–759.
66. Kuwahara, M.; Nagashima, J.; Hasegawa, M.; Tamura, T.; Kitagata, R.; Hanawa, K.; Hososhima, S.; Kasamatsu, T.; Ozaki, H.; Sawai, H. Systematic characterization of 2'-deoxynucleoside-5'- triphosphate analogs as substrates for DNA polymerases by polymerase chain reaction and kinetic studies on enzymatic production of modified DNA. *Nucleic Acids Res.* **2006**, *34*, 5383–5394.
67. Gierlich, J.; Gutsmedl, K.; Gramlich, P. M. E.; Schmidt, A.; Burley, G. A.; Carell, T. Synthesis of highly modified DNA by a combination of PCR with alkyne-bearing triphosphates and click chemistry. *Chem. Eur. J.* **2007**, *13*, 9486–9494.
68. Brudno, Y.; Liu, D. R. Recent progress toward the templated synthesis and directed evolution of sequence-defined synthetic polymers. *Chem. Biol.* **2009**, *16*, 265–276.
69. Cherkasov, D.; Biet, T.; Bäuml, E.; Traut, W.; Lohoff, M. New nucleotide analogues with enhanced signal properties. *Bioconjugate Chem.* **2010**, *21*, 122–129.
70. Hollenstein, M. Synthesis of deoxynucleoside triphosphates that include proline, urea, or sulfonamide groups and their polymerase incorporation into DNA. *Chem. Eur. J.* **2012**, *18*, 13320–13330.
71. Goubet, A.; Chardon, A.; Kumar, P.; Sharma, P. K.; Veedu, R. N. Synthesis of DNA oligonucleotides containing C5-ethynylbenzenesulfonamidemodified nucleotides (EBNA) by polymerases towards the construction of base functionalized nucleic acids. *Bioorganic Med. Chem. Lett.* **2013**, *23*, 761–763.
72. Dutson, C.; Allen, E.; Thompson, M. J.; Hedley, J. H.; Murton, H. E.; Williams, D. M. Synthesis of polyanionic C5-modified 2'-deoxyuridine and 2'-deoxycytidine-5'-triphosphates and their properties as substrates for DNA polymerases. *Molecules* **2021**, *26*, 2250.
73. Tanpure, A. A.; Srivatsan, S. G. Synthesis and photophysical characterization of a fluorescent nucleoside analogue that signals the presence of an abasic site in RNA. *ChemBioChem* **2012**, *13*, 2392–2399.
74. Ménová, P.; Cahová, H.; Plucnara, M.; Havran, L.; Fojta, M.; Hocek, M. Polymerase synthesis of oligonucleotides containing a single chemically modified nucleobase for site-specific redox labelling. *Chem. Commun.* **2013**, *49*, 4652–4654.
75. Takezawa, Y.; Nakama, T.; Shionoya, M. Enzymatic synthesis of Cu(II)-responsive deoxyribozymes through polymerase incorporation of artificial ligand-type nucleotides. *J. Am. Chem. Soc.* **2019**, *141*, 19342–19350.
76. Kuba, M.; Kraus, T.; Pohl, R.; Hocek, M. Nucleotide-bearing benzylidene-tetrahydroxanthylum near-IR fluorophore for sensing DNA replication, secondary structures and interactions. *Chem. Eur. J.* **2020**, *26*, 11950–11954.
77. Flamme, M.; Röthlisberger, P.; Levi-Acobas, F.; Chawla, M.; Oliva, R.; Cavallo, L.; Gasser, G.; Marlière, P.; Herdewijn, P.; Hollenstein, M. Enzymatic formation of an artificial base pair using a modified purine nucleoside triphosphate. *ACS Chem. Biol.* **2020**, *15*, 2872–2884.
78. Röthlisberger, P.; Levi-Acobas, F.; Leumann, C. J.; Hollenstein, M. Enzymatic synthesis of biphenyl-DNA oligonucleotides. *Bioorg. Med. Chem.* **2020**, *28*, 115487.
79. Nakama, T.; Takezawa, Y.; Sasaki, D.; Shionoya, M. Allosteric regulation of DNAzyme activities through intrastrand transformation induced by Cu(II)-mediated artificial base pairing. *J. Am. Chem. Soc.* **2020**, *142*, 10153–10162.
80. Le, B. H.; Nguyena, V. T.; Seo, Y. J. Site-specific incorporation of multiple units of functional nucleotides into DNA using a step-wise approach with polymerase and its

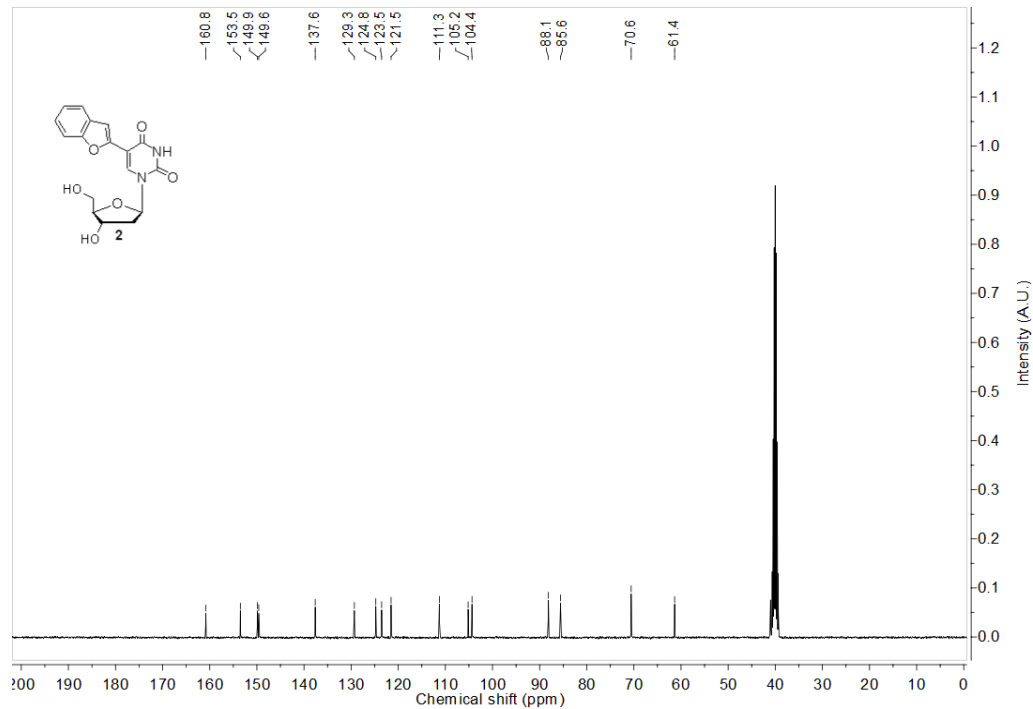
- application to monitoring DNA structural changes. *Chem. Commun.* **2019**, 55, 2158–2161.
81. Li, Y.; Kong, Y.; Korolev, S.; Waksman, G. Crystal structures of the Klenow fragment of *Thermus aquaticus* DNA polymerase I complexed with deoxyribonucleoside triphosphates. *Protein Sci.* **1998**, 7, 1116–1123.
 82. Li, Y.; Waksman, G. Crystal structures of a ddATP-, ddTTP-, ddCTP, and ddGTP-trapped ternary complex of KlenTaq1: insights into nucleotide incorporation and selectivity. *Protein Sci.* **2001**, 10, 1225–1233.
 83. Jackson, L. N.; Chim, N.; Shi, C.; Chaput, J. C. Crystal structures of a natural DNA polymerase that functions as an XNA reverse transcriptase. *Nucleic Acids Res.* **2019**, 47, 6973–6983.
 84. Johnson, S. J.; Taylor, J. S.; Beese, L. S. Processive DNA synthesis observed in a polymerase crystal suggests a mechanism for the prevention of frameshift mutations. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, 100, 3895–3900.
 85. Chim, N.; Meza, R. A.; Trinh, A. M.; Yang, K.; Chaput, J. C. Following replicative DNA synthesis by time resolved X-ray crystallography. *Nat. Commun.* **2021**, 12, 2641.
 86. Kropp, H. M.; Diederichs, K.; Marx, A. The structure of an archaeal B-family DNA polymerase in complex with a chemically modified nucleotide. *Angew. Chem. Int. Ed.* **2019**, 58, 5457–5461.
 87. Kropp, H. M.; Betz, K.; Wirth, J.; Diederichs, K.; Marx, A. Crystal structures of ternary complexes of archaeal B-family DNA polymerases. *PLoS One* **2017**, 12, e0188005.
 88. Bergen, K.; Betz, K.; Welte, W.; Diederichs, K.; Marx, A. Structures of KOD and 9°N DNA polymerases complexed with primer template duplex. *ChemBioChem* **2013**, 14, 1058–62.
 89. Kropp, H. M.; Dürr, S. L.; Peter, C.; Diederichs, K.; Marx, A. Snapshots of a modified nucleotide moving through the confines of a DNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, 115, 9992–9997.
 90. Kabsch, W. XDS. *Acta Crystallogr. D.* **2010**, 66, 125–132.
 91. Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.; Kapral, G. J.; GrosseKunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilligere, T. C.; Zwart, P. H. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D.* **2010**, 66, 213–221.
 92. Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D.* **2004**, 60, 2126–2132.
 93. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and development of Coot. *Acta Crystallogr. D.* **2010**, 66, 486–501.
 94. Chen, V. B.; Arendall III, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D.* **2010**, 66, 12–21.

2.6 Appendix-I NMR and Mass spectra of synthesized compounds

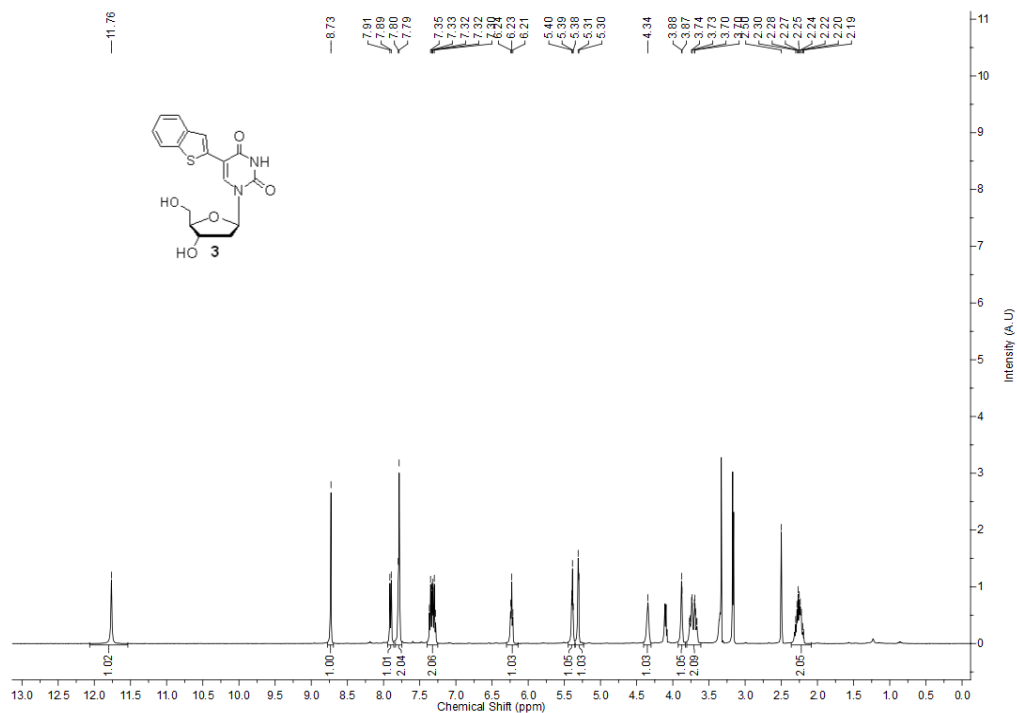
^1H -NMR of BFdU **2** (400 MHz, d_6 -DMSO).



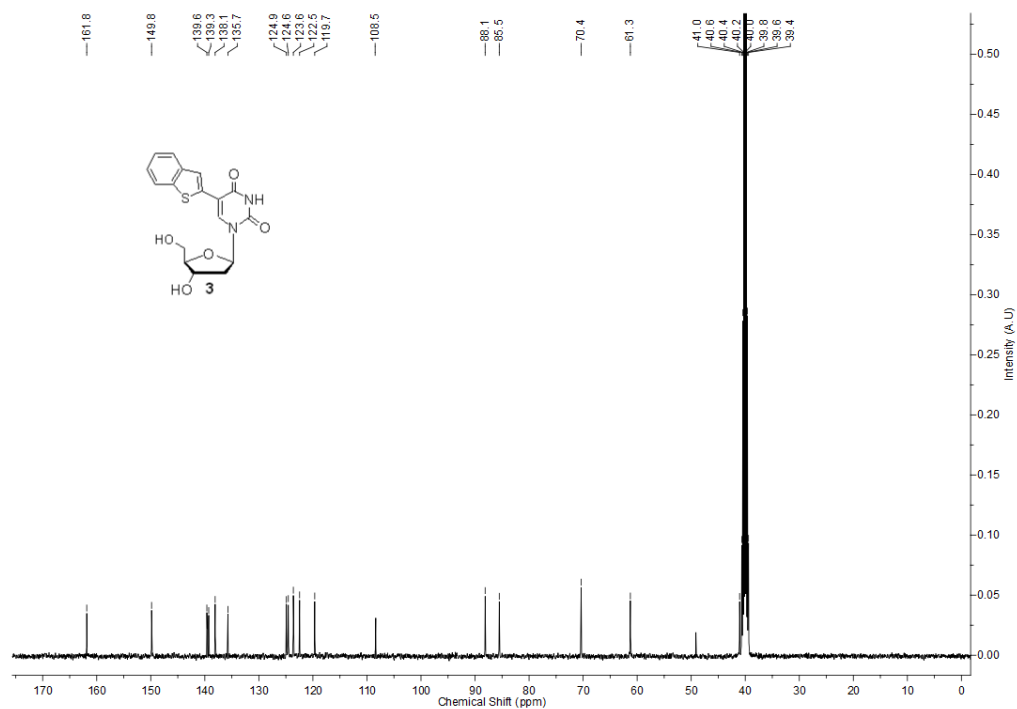
^{13}C -NMR of BFdU **2** (100 MHz, d_6 -DMSO).



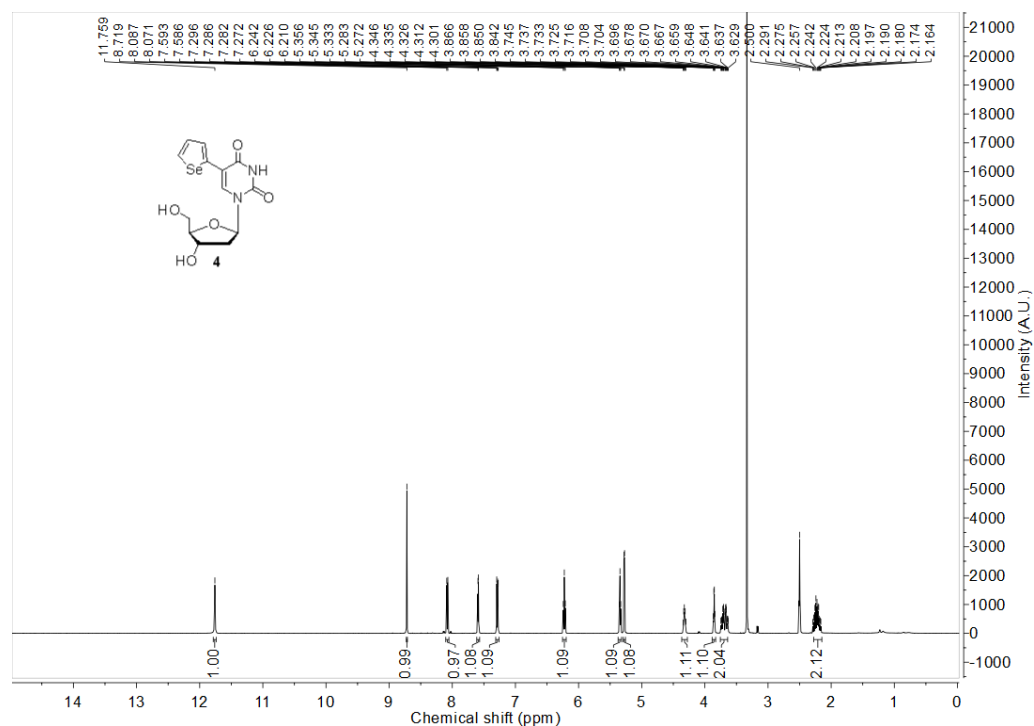
^1H -NMR of BTdU **3** (400 MHz, d_6 -DMSO), peaks for a small amount of methanol was present even after multiple cycles of vacuum drying.



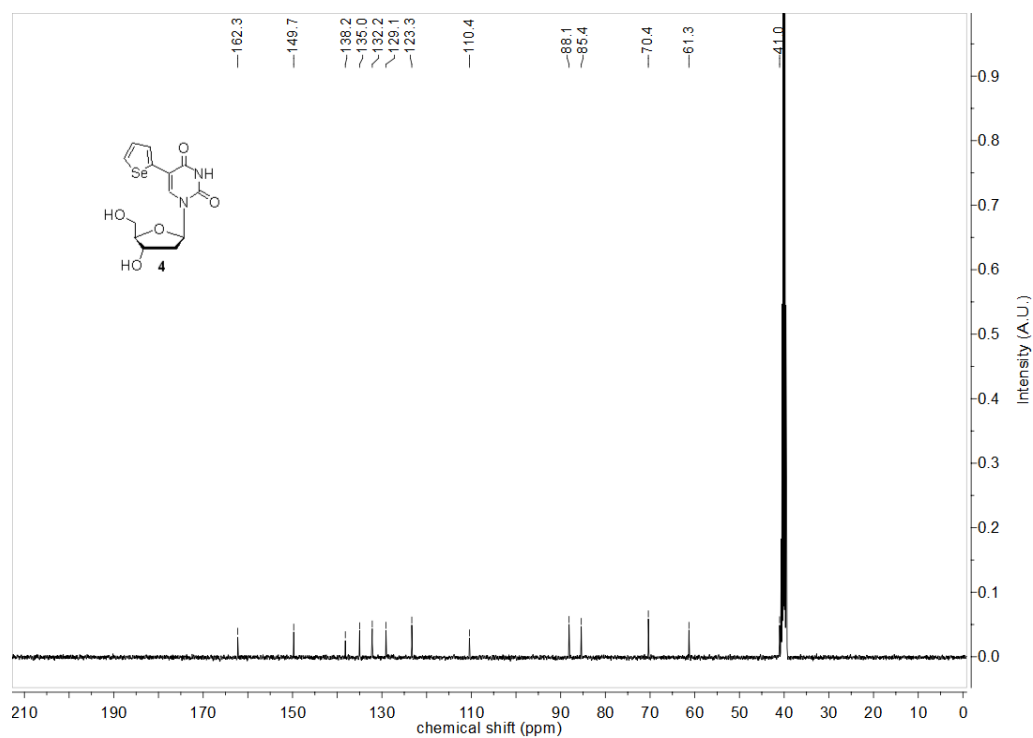
^{13}C -NMR of **3** (100 MHz, d_6 -DMSO), a peak for a small amount of methanol was present



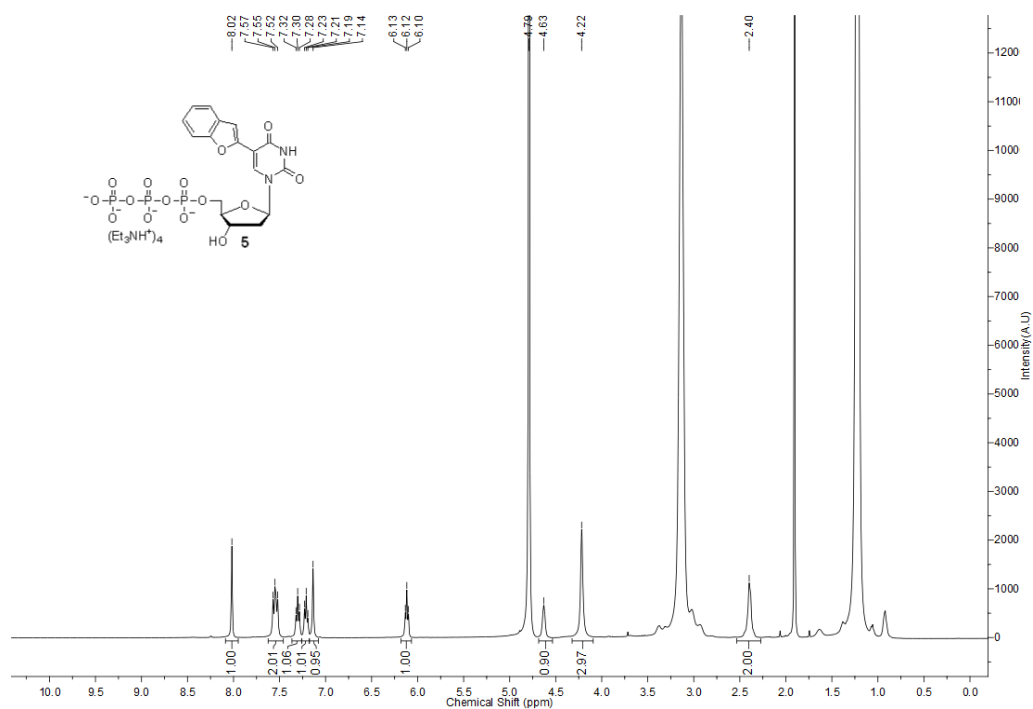
^1H -NMR of SedU **4** (400 MHz, d_6 -DMSO).



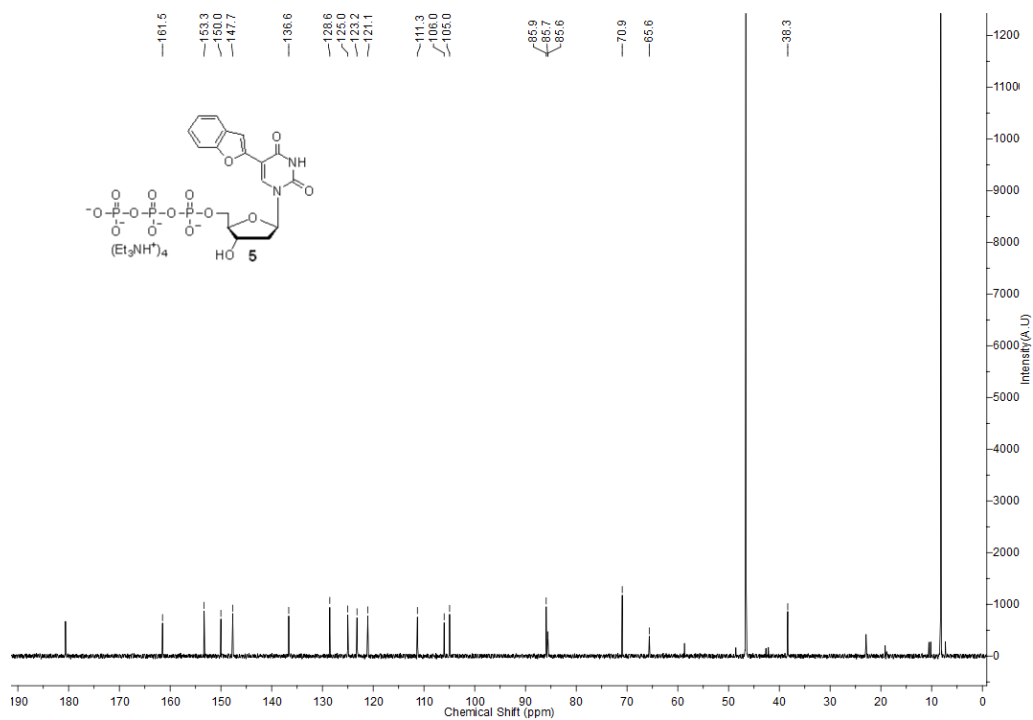
^{13}C -NMR of **4** (100 MHz, d_6 -DMSO).



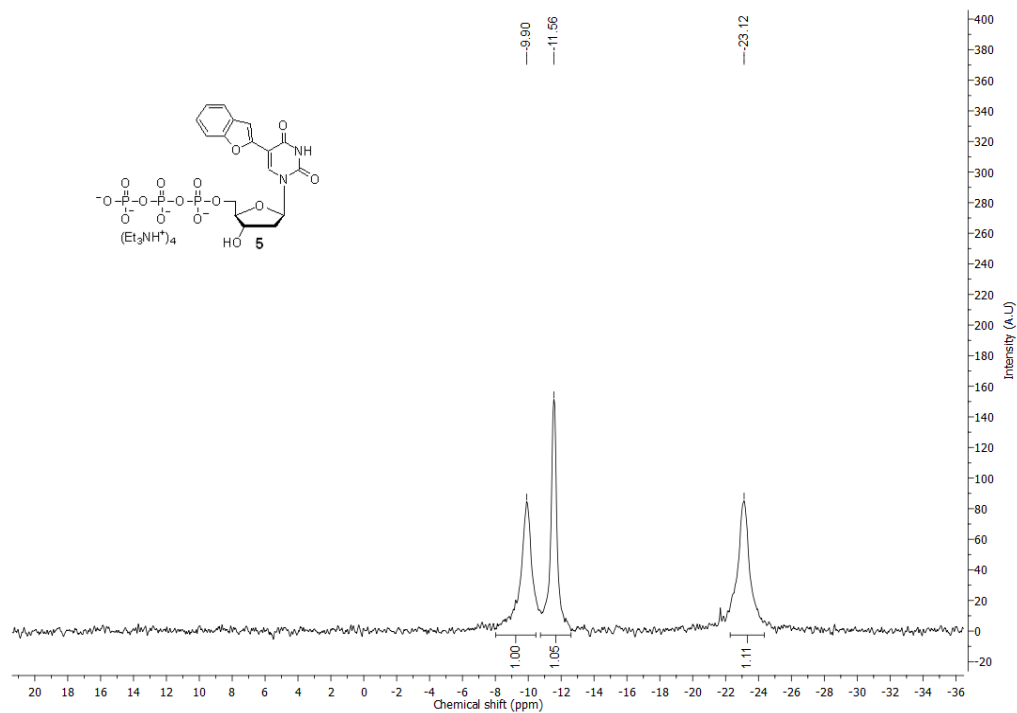
^1H -NMR of BFdUTP **5** (400 MHz, D_2O), trace amount of triethylammonium acetate buffer is present.



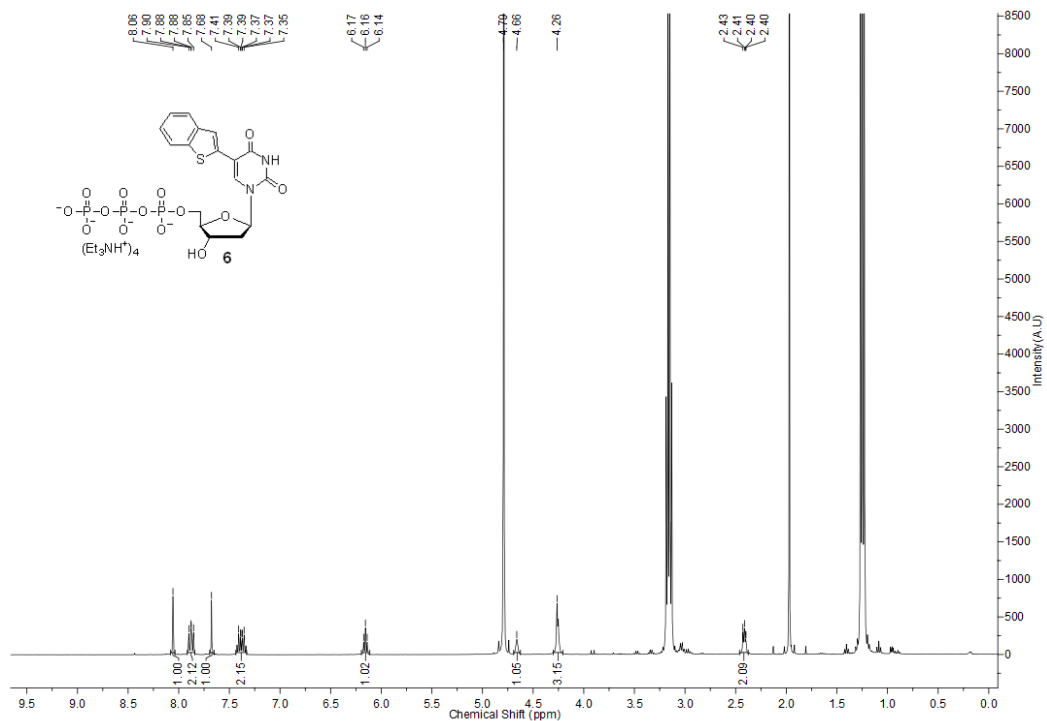
^{13}C -NMR of BFdUTP **5** (100 MHz, D_2O), trace amount of triethylammonium acetate buffer is present.



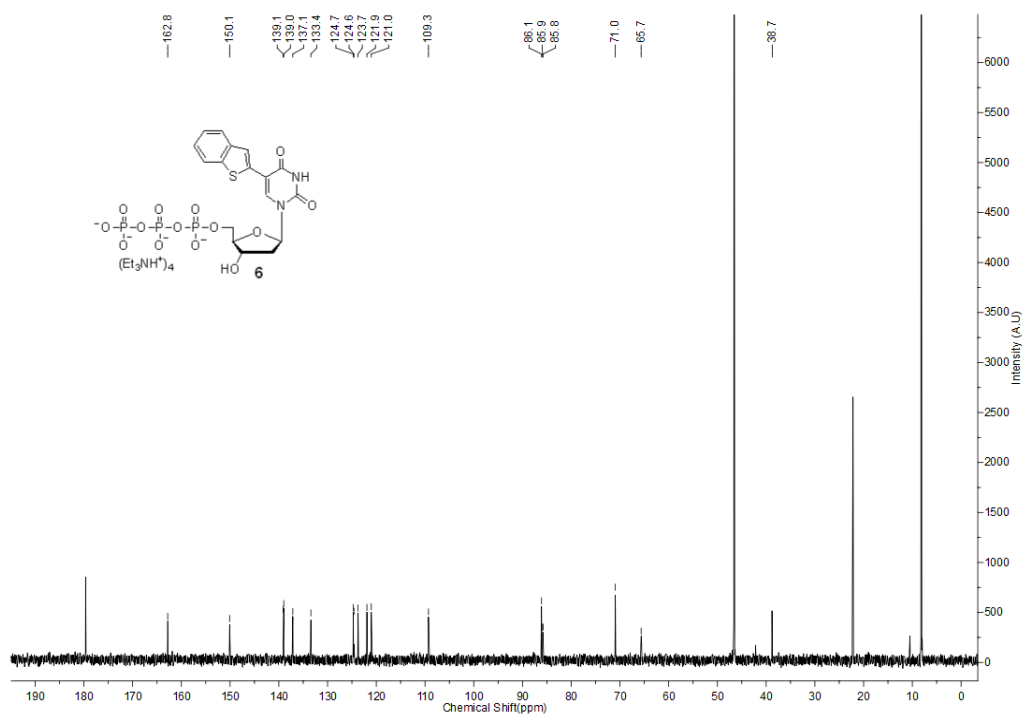
^{31}P -NMR of BFdUTP **5** (162 MHz, D_2O)



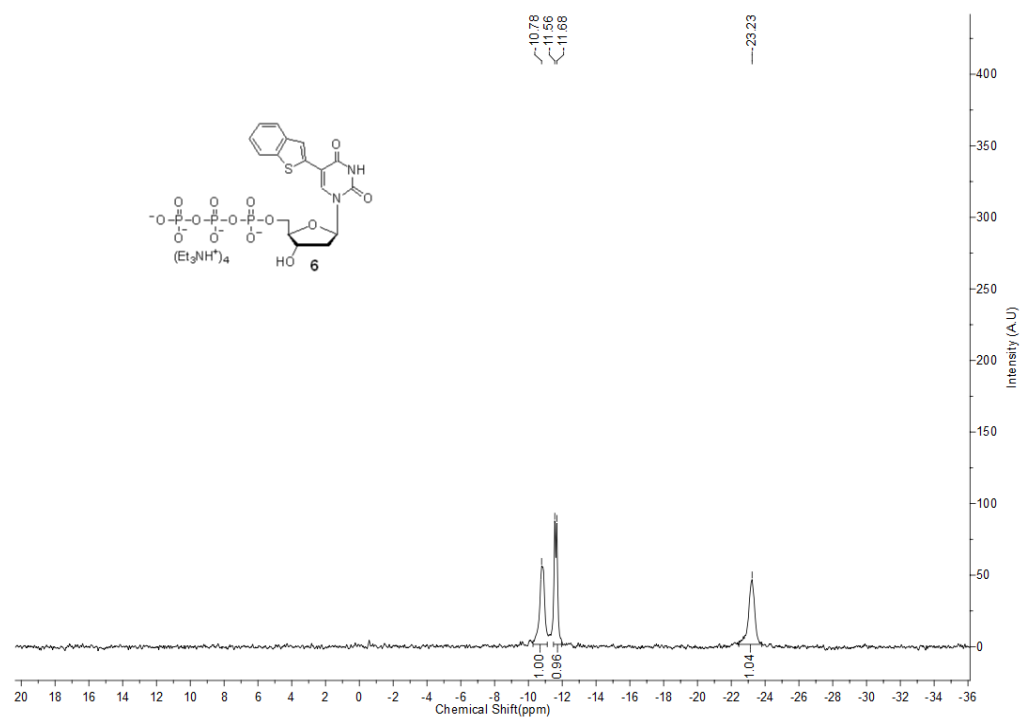
^1H -NMR of BTdUTP **6** (400 MHz, D_2O), trace amount of triethylammonium acetate buffer is present.



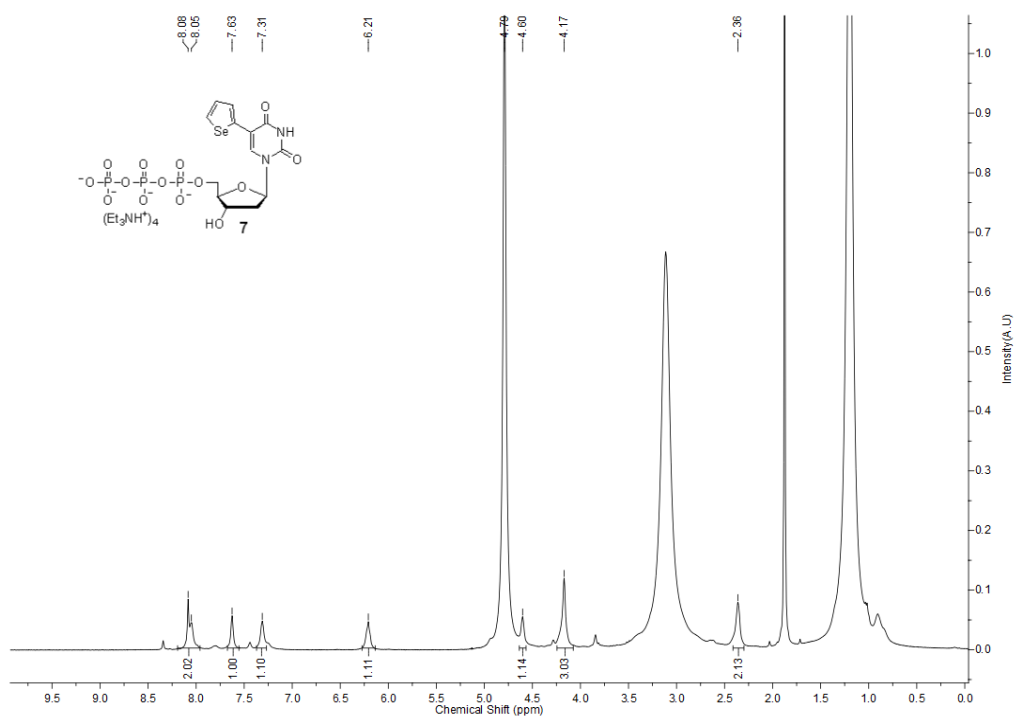
^{13}C -NMR of BTdUTP **6** (100 MHz, D_2O), trace amount of triethylammonium acetate buffer is present.



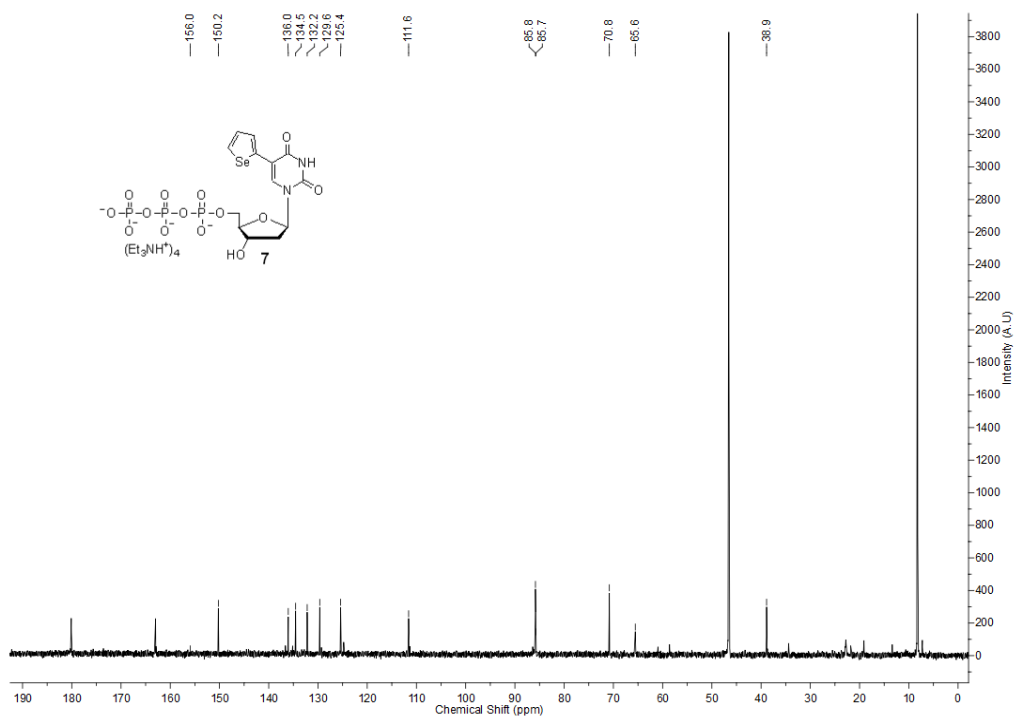
^{31}P -NMR of BTdUTP **6** (162 MHz, D_2O)



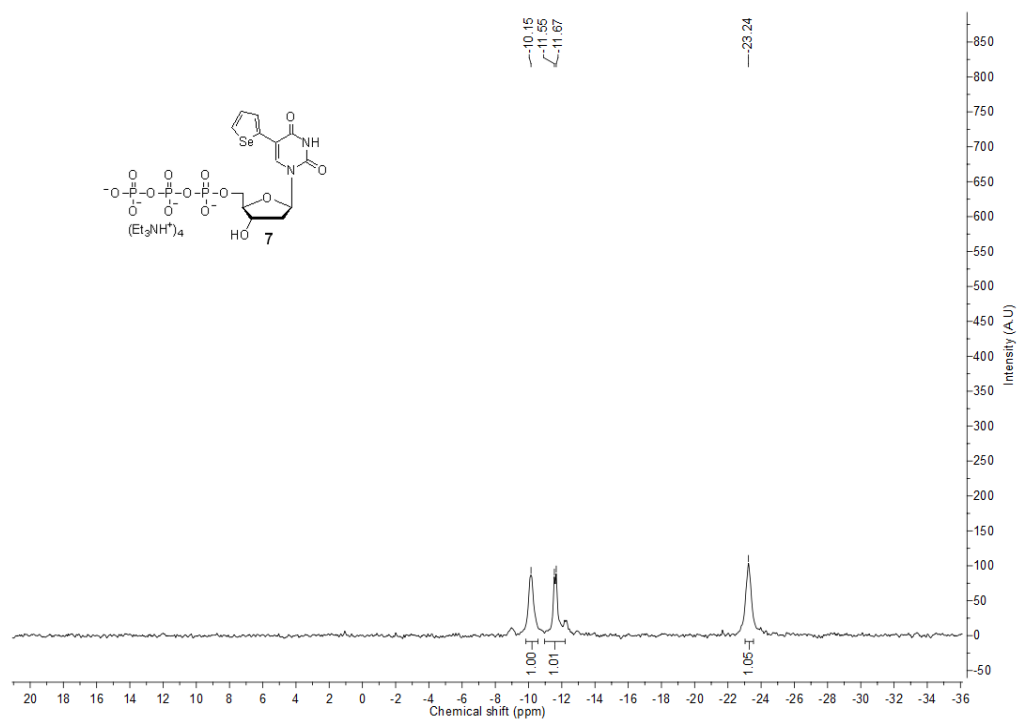
^1H -NMR of SedUTP **7** (400 MHz, D_2O), trace amount of triethylammonium acetate buffer is present.



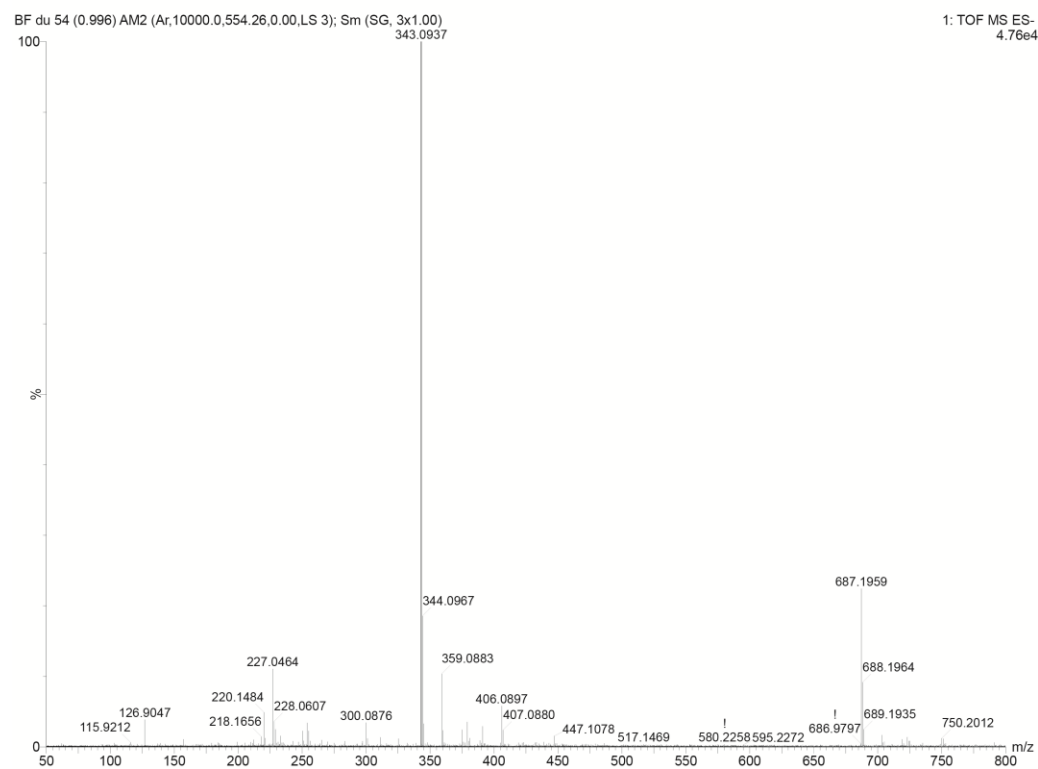
^{13}C -NMR of SedUTP **7** (100 MHz, D_2O), trace amount of triethylammonium acetate buffer is present.



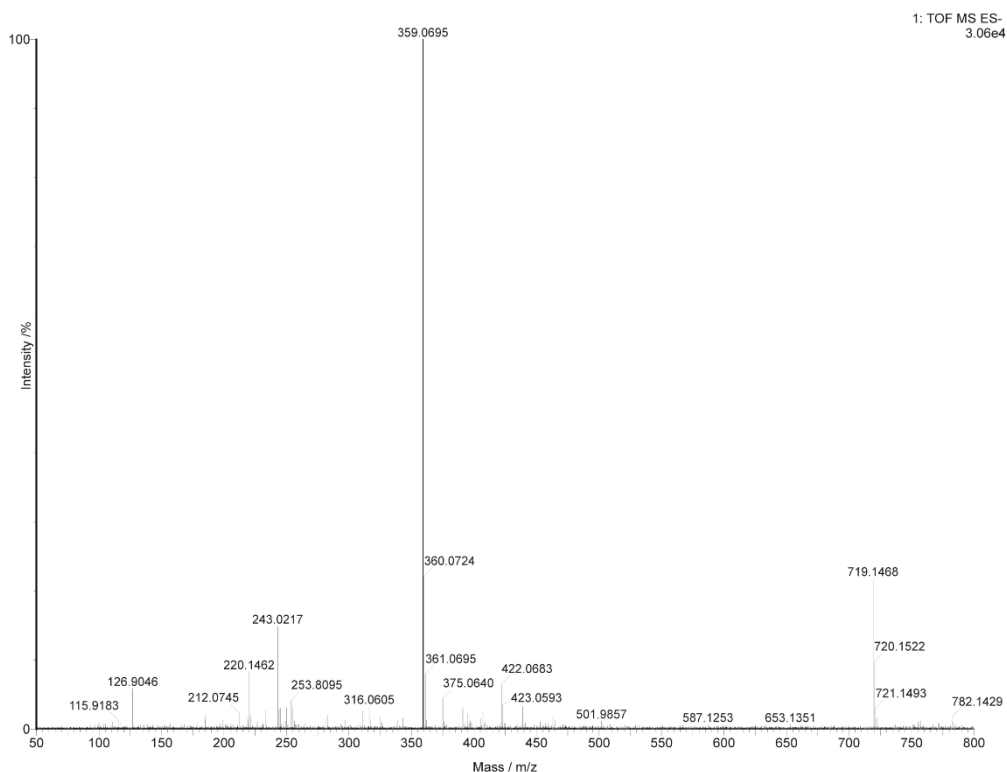
^{31}P -NMR of SedUTP **7** (162 MHz, D_2O)



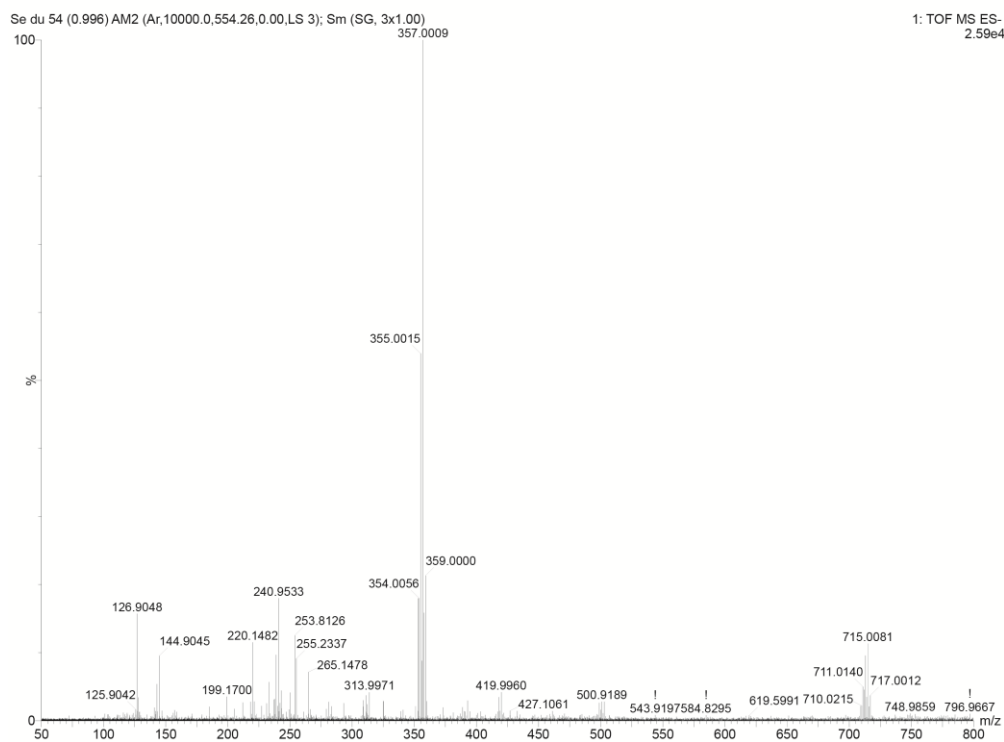
HRMS spectrum of BFdU **2** in negative mode. Calculated for $\text{C}_{17}\text{H}_{15}\text{N}_2\text{O}_5\text{S}$ $[\text{M}-\text{H}]^- = 343.0936$, found $[\text{M}-\text{H}]^- = 343.0937$.



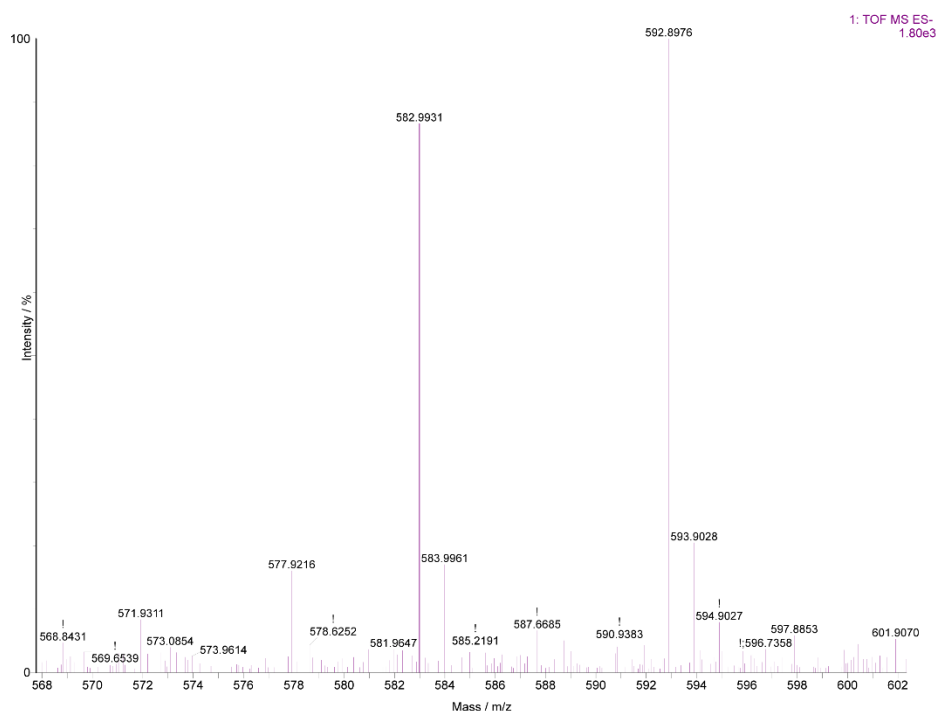
HRMS spectrum of BTdU **3** in negative mode. Calculated for $C_{17}H_{15}N_2O_5S$ $[M-H]^- = 359.0702$, found $[M-H]^- = 359.0695$.



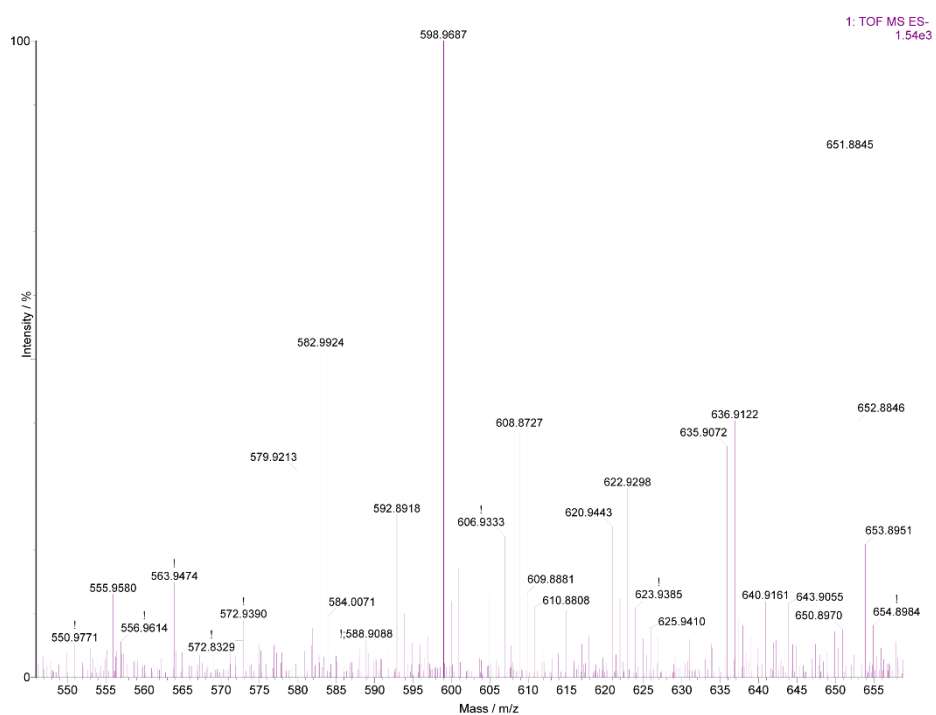
HRMS spectrum of SedU **4** in negative mode. Calculated for $C_{13}H_{14}N_2O_5Se$ $[M-H]^- = 356.9995$, found $[M-H]^- = 357.0007$.



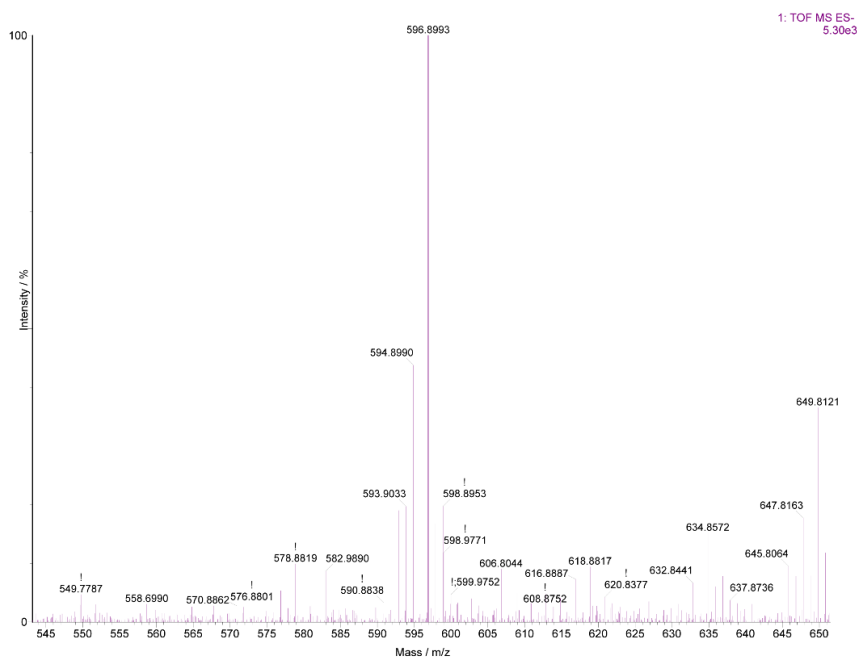
HRMS spectrum of BFdUTP **5** in negative mode. Calculated for $C_{17}H_{18}N_2O_{15}P_3$ $[M-H]^- = 582.9926$, found $[M-H]^- = 582.9931$.



HRMS spectrum of BTdUTP **6** in negative mode. Calculated for $C_{17}H_{18}N_2O_{14}P_3S$ $[M-H]^- = 598.9697$, found $[M-H]^- = 598.9687$.



HRMS spectrum of SedUTP **7** in negative mode. Calculated for $\text{C}_{13}\text{H}_{16}\text{N}_2\text{O}_{14}\text{P}_3\text{Se}$ $[\text{M}-\text{H}]^- = 596.8985$, found $[\text{M}-\text{H}]^- = 596.8993$.



Chapter 3

Probing the processing of environment-sensitive nucleotide analogs by DNA polymerases using fluorescence, ^{19}F NMR and X-ray crystallography*

3.1 Introduction

Enzymatic labeling of nucleic acids with modified nucleotides has gained paramount significance for numerous core biotechnological applications and investigating nucleic acid structure and functions.¹⁻⁸ Different classes of native and engineered DNA polymerases have been used to process several structurally different nucleotide analogs as illustrated in Chapter 1. The A-family of DNA polymerase KlenTaq has been employed as a model system to obtain structural insights into the processing of different C5-modified pyrimidine and C7-modified purine analogs and their interactions with the polymerases.⁹⁻¹³ Modifications involving polar flexible groups, such as aminopentynyl, and nonpolar aromatic moieties like ethynylbenzene, linked to either C5 position of pyrimidines or C7 position of 7-deazapurines through an alkyne bond have been observed to enhance the stability of the substrate-enzyme interaction, resulting in notable efficiency in incorporation.^{10,11} On the contrary, when a nucleotide is linked to a rigid, nonpolar spin label, it induces distortion in the enzyme conformation.¹² Consequently, this distortion negatively impacts the efficiency of incorporation. We recently demonstrated the usefulness of environment-responsive heterocycle-conjugated fluorescent nucleotides in probing the incorporation mechanism of KlenTaq DNA polymerase in solution and at the atomic level (Chapter 2). Further, the processivity of the KOD polymerases was also evaluated and a comparative analysis provided the following understanding. Incorporating efficiency mainly depends on the size of the heterocyclic moiety and the type of DNA polymerase used. The KlenTaq active site and channel are more compact than in KOD, and hence, KlenTaq is less tolerant to bulky modifications (e.g., benzofuran and benzothiophene). KlenTaq allows the incorporation of a single modified nucleotide in PEX reactions irrespective of the heterocycle size. KOD, on the contrary, tolerates all the nucleotides tested and efficiently allows single as

*Part of the work in this chapter is published: Ghosh, P.; Kropp, H. M.; Betz, K.; Ludmann, S.; Diederichs, K.; Marx, A.; Srivatsan, S. G. Microenvironment-sensitive fluorescent nucleotide probes from benzofuran, benzothiophene, and selenophene as substrates for DNA polymerases. *J. Am. Chem. Soc.* **2022**, *144*(23), 10556–10569. doi:10.1021/jacs.2c03454 (American Chemical Society acknowledged).

well as multiple insertions. Notably, benzofuran-conjugated nucleotide (BFdUTP) acts as a very good probe, providing direct fluorescence readouts to monitor enzyme activity and binding processes in the catalytic cycle.

Taking forward, we sought to evaluate the compatibility of a derivate of BFdUTP namely, 5-fluorobenzofuran-modified 2'-deoxyuridine triphosphate (FBFdUTP **3**) in PEX reactions catalyzed by KlenTaq and KOD enzymes (Figure 1). FBFdU and FBFU serve as excellent dual-app sensors in probing the structure and dynamics of non-canonical structures such G-quadruplexes and i-motifs by fluorescence and ^{19}F NMR techniques.¹⁴ Due to the high-gyromagnetic coefficient, 100 % natural abundance, sensitivity towards changes in the local chemical environment and absence of fluorine atoms in biomolecules, ^{19}F NMR spectroscopy has become a powerful biophysical tool for analysis of nucleic acid structures and its recognition in cellular environment.¹⁵⁻¹⁹ Various fluorine-modified nucleotides have been incorporated into oligonucleotide strands to study their structure and dynamics in solution or cellular conditions, along with their recognition properties with small molecules.²⁰⁻²⁷ However, only a few examples of such fluorine-labeled probes have been employed to investigate the interactions of nucleic acid binding protein and nucleic acid processing enzymes.²⁸⁻³³ Recently, 2'-fluoro substituted deoxy nucleotide has been used to study the nucleotide flipping in nucleic acids by thymine DNA glycosylase (TDG) enzymes using ^{19}F NMR.³¹ Alternatively, a KlenTaq DNA polymerase modified with multiple fluorine atoms has been employed for the investigation of various stages in enzymatic incorporation. However, the multiple modifications hampered the stability of the enzyme itself and caused difficulty in assigning specific interactions.³³ To the best of our knowledge, no current method has been developed to study the DNA polymerization process using ^{19}F NMR spectroscopy using environment-sensitive nucleotide analogs (Figure 1). The advantages of monitoring the enzymatic incorporation of FBFdUTP **3** are the following. The nucleoside comprises a fluorophore and ^{19}F NMR label, which opens up a two-channel spectroscopic platform to monitor the catalytic events under equilibrium conditions. The fluorophore and the ^{19}F NMR labels are integrated into the same electronic environment, and so, the data obtained from fluorescence and NMR experiments can be efficiently correlated with each other. Further, the efficiency of incorporation and binding data determined by biochemical assay, fluorescence, and NMR can be correlated with the structure obtained by X-ray crystallography to get deeper insights into the intricate steps underlying the DNA polymerization process.

In this chapter, the synthesis and the incorporation of FBFdUTP **3** by DNA polymerases are described. The modified nucleotide analog is incorporated into ONs by KlenTaq and KOD

DNA polymerases with varying efficiency. KOD DNA polymerase tolerates multiple modifications and is well-suited for synthesizing high-density labeled ONs. In contrast, KlenTaq exclusively accepts a single modification at the 3'-end like in the case of BFdUTP, and enables the synthesis of site-specifically labeled DNA (Figure 1). Further, we monitored the binary complex formation and the incorporation efficiency of KlenTaq and KOD DNA polymerases using fluorescence and ^{19}F NMR by utilizing the responsiveness of the dual-app probe. We also obtained high-resolution X-ray crystal structures of the ternary complexes of KlenTaq-duplex-FBFdUTP/SedUTP. Through biochemical, biophysical and X-ray structure analyses using BFdUTP, FBFdUTP and SedUTP elucidated the differences in the incorporation efficiency exhibited by different polymerases.

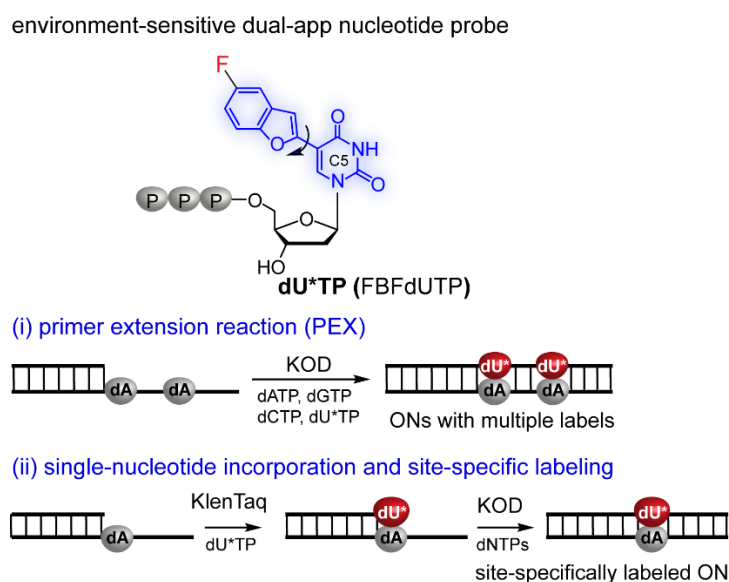


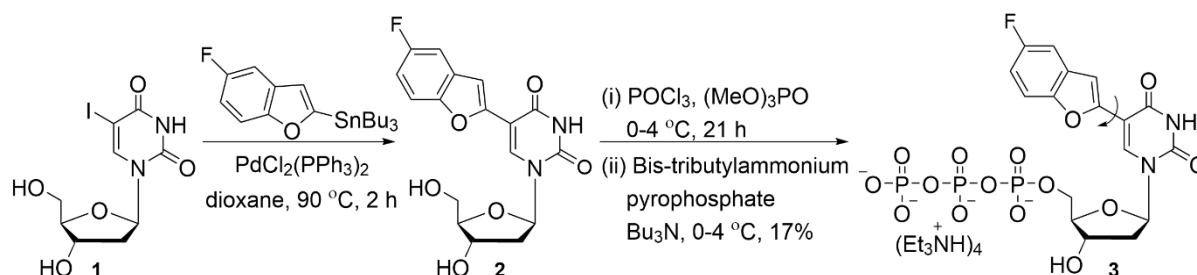
Figure 1. C5-Fluorobenzofuran-modified dual-app dU*TP analog is well-suited for enzymatic functionalization of oligonucleotides (ONs) by DNA polymerases. This allows for the synthesis of ONs with both (i) multiple and (ii) site-specific labels.

3.2 Results and Discussion

3.2.1 Synthesis of FBFdUTP (3)

FBFdU nucleoside probe (**2**) was synthesized by attaching a fluorobenzofuran moiety at the C5 position of 2'-deoxyuridine to study the acceptance of this analog by DNA polymerases for the following reasons (Scheme1). The modification at the C5 position of the pyrimidine ring minimally perturbs the DNA structure and exhibits favorable substrate properties in enzymatic incorporation.³⁴⁻⁴¹ Introducing the heterocyclic moiety via a rotatable aryl-aryl bond with the uridine ring makes it a conformation-sensitive probe.⁴²⁻⁴⁶ The analog demonstrates high sensitivity to surrounding polarity and viscosity, leading to changes in its photophysical properties such as emission maximum, lifetime, and quantum yield. Similarly, the nucleoside

exhibits distinct ^{19}F chemical shifts in solvents of different polarity and viscosity.⁴⁴ The synthesis involved initial stannylation of 5-fluorobenzofuran, followed by cross-coupling with 5-iodo-2'-deoxyuridine **1** in the presence of a palladium catalyst, yielding 5-fluorobenzofuran-2'-deoxyuridine (FBFdU) **2**. The nucleotide substrate (FBFdUTP) **3** required to study the DNA polymerase activity was synthesized by reacting FBFdU first with POCl_3 and then with bis-tributylammonium pyrophosphate (Scheme 1). The nucleotide was purified by anion exchange chromatography followed by RP-HPLC to afford the pure product in 17%.



Scheme 1. Synthesis of FBFdUTP **3**.

Table 1. ON sequences used in PEX, single nucleotide incorporation, competition assay and crystallization studies, and fluorescence assays

ONs	Sequence
P1	5' FAM-GTG GTG CGA AAT TTC TGA CAG ACA 3'
T1	3' CAC CAC GCT TTA AAG ACT GTC TGT ACT GTC TGC GTG 5'
T2	3' CAC CAC GCT TTA AAG ACT GTC TGT GCT GTC TAC GTG 5'
T3	3' CAC CAC GCT TTA AAG ACT GTC TGT GCT GT A TAC GTG 5'
T4	3' CAC CAC GCT TTA AAG ACT GTC TGT GCT GTC TAA GTG 5'
T5	3' CAC CAC GCT TTA AAG ACT GTC TGT GCT GT A TAA GTG 5'
P2	5' FAM-GGA GCT CAG CCT TCA CTG C 3'
P3	5' GGA GCT CAG CCT TCA CTG C 3'
T6	3' CCT CGA GTC GGA AGT GAC G ACG ATT TGA TAT AAC GAT CGC GTG 5'
P2*	5' FAM-GGA GCT CAG CCT TCA CTG C- FBFdU 3'
P3*	5' GGA GCT CAG CCT TCA CTG C- FBFdU 3'
P4	5' GAC CAC GGC CA 3'
T7	3' CCT CGA GTC GGA AGT GAC G ATG TCG CTC CGC GTC 5'
T8	3' CTG GTG CCG GT G ACAA 5'
P5	5' GGA GCT CAG CCT TCA CTG ddC 3'
P6	5' FAM-GAC CAC GGC CAC 3'

[a]In PEX reactions, the ON primer **P1** and templates **T1–T5** were employed. For single-nucleotide incorporation reactions, ONs **P2/P3** and **T6** were utilized. The primers **P2*** and **P3*** refer to 3'-FBFdU-labeled of **P2** and **P3**, respectively. **P3*** was specifically employed in fluorescence assays. Crystallization trials involved ONs **P4** and **T8**. The primer **P6** was used for the competition assay. The template ON **T7** was utilized in binding studies through fluorescence and EMSA. The incorporation of modified nucleotide analogs in PEX and single nucleotide incorporation assays is guided by 2'-deoxyadenosines, highlighted in red in template strands. In ON **T8**, during crystallization trials the termination of extension was achieved by incorporating ddCTP after 2'-deoxyguanosine, marked in blue.

3.2.2 DNA polymerases incorporate the nucleotide analog with varying efficiency

The ability of FBFdUTP (**3**) to act as a thymidine analog was tested by performing primer extension (PEX) reactions using KlenTaq and KOD DNA polymerases. A 5'-FAM labeled primer **P1** was hybridized with a set of templates (**T1–T5**). These templates code for incorporating FBFdUTP (**3**) at various positions in the PEX reaction, including the start site, a location distant from the start site, and at multiple sites along the growing ON strand. The reactions were initiated by adding dGTP, dCTP, dATP and either dTTP or FBFdUTP **3**. The PEX reaction products were analyzed by PAGE and visualized at the FAM wavelength using a fluorescence gel scanner. Notably, when PEX reactions were conducted with templates **T1** and **T2**, guiding a single modification, KlenTaq efficiently accepted the modified mononucleotide, resulting in full-length product formation (Figure 2B, lanes 4, 5 and 7, 8). As a control experiment, we performed a reaction without dTTP/BBFdUTP **3** but with other dNTPs, which led to the absence of full-length product formation (lanes 6 and 9). This result confirms that the full-length product obtained with FBFdUTP is not due to a misincorporation. However, with templates **T3** and **T4**, which would introduce two modifications at consecutive or alternative positions, KlenTaq failed to yield full-length products with modification (Figure 2B, lanes 10–14). In contrast, KOD DNA polymerase effectively processed FBFdUTP in the presence of all the templates (**T1–T5**), generating full-length products comparable to those formed in the presence of native dNTPs (Figure 2C, lanes 4–14 and Figure 3). These findings align with the results obtained in our previous study using BFdUTP and SedUTP (Chapter 2, section 2.2.2). We observed that introducing a fluorine label at the five position of the benzofuran ring does not impair its incorporation efficiency by DNA polymerases. While KOD DNA polymerase is suitable for multiple labeling, KlenTaq shows difficulties in incorporating FBFdUTP **3** with high demanding templates.

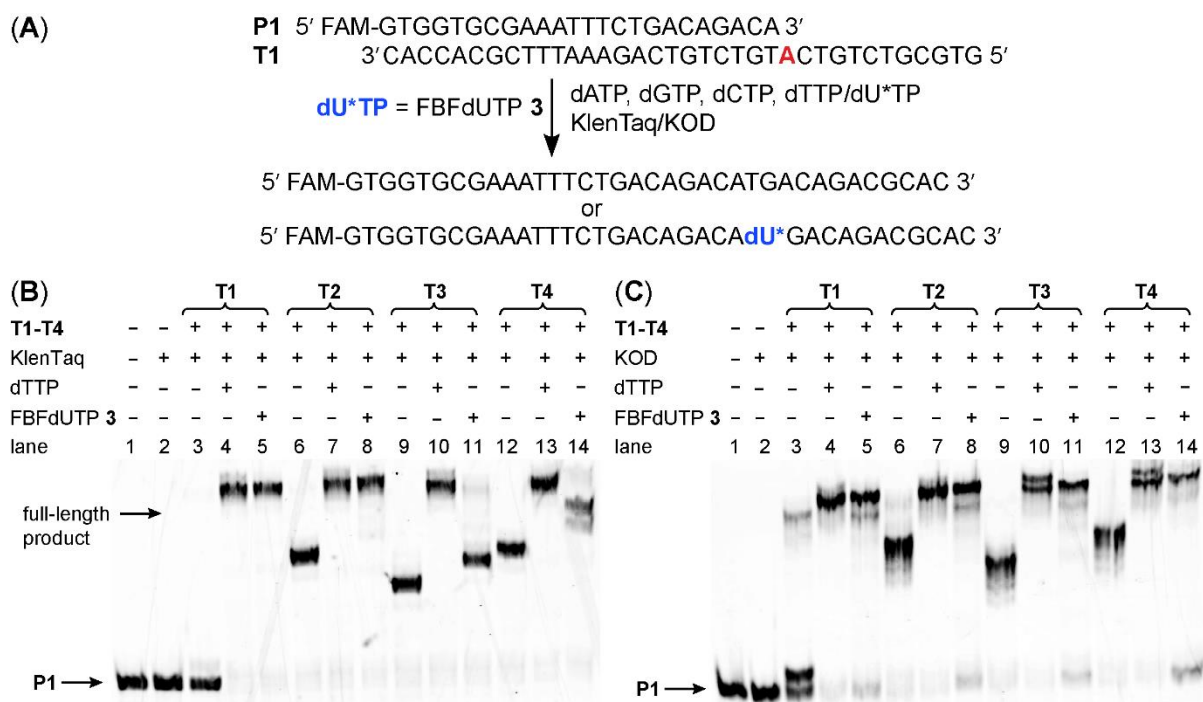


Figure 2. (A) The PEX reaction setup was employed for the integration of the mononucleotide analog of 3 with various DNA polymerases. (B and C) Incorporation of FBFdUTP 3 using KlenTaq and KOD DNA polymerases, respectively, in the presence of primer P1 and templates T1–T4.

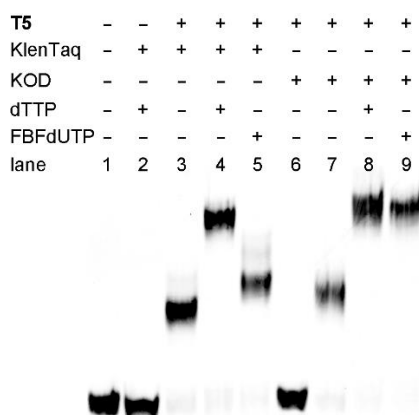


Figure 3. In the presence of a duplex composed of primer P1 and template T5, KlenTaq and KOD DNA polymerases were used to incorporate mononucleotides of FBFdUTP 3.

3.2.3 Site-specific labeling of FBFdUTP 3

ONs labeled site-specifically with ^{19}F -containing nucleosides have been used in studying nucleic acid-drug interactions or screening drug molecules for therapeutically relevant nucleic acid targets.^{47,48} The ONs are typically obtained by the solid-phase ON synthesis method. In our previous work, we reported a template-dependent two-step site-specific labeling method of DNA ONs using DNA polymerases (Chapter 2, section 2.2.3). The enzyme should insert a single modification at the 3'-end of the primer with exceptional fidelity, which then should be

amicable to extension in the presence of natural dNTPs to produce the full-length DNA with the modification at the desired position (Figure 4A). A 5'-FAM labeled primer-template duplex (**P2•T6**) was employed for this purpose, where the template **T6** contained a dA that would code the insertion of a single modified nucleotide at the 3'-end of the primer. We performed single nucleotide incorporation of the FBFdUTP **3** using KlenTaq and KOD DNA polymerase. KOD DNA polymerase produces largely double-modified products, due to the misincorporation of the FBFdUMP opposite to the dC residue in the template strand (Figure 4B). Whereas KlenTaq efficiently accepts the modification and provides a single-labeled product. Large-scale single-nucleotide incorporations were conducted, and the formation of 3'-end labeled products was confirmed by mass analysis.

The next objective was to check the feasibility of the two-step site-specific labeling technique for the incorporation of FBFdUTP into the DNA strand (Figure 4A). First, a single nucleotide reaction was conducted using KlenTaq, and the reactions were quenched by heating (Figure 4C, lane 2). The singly-labeled duplex DNA was precipitated using 70% ethanol and was washed multiple times to remove the excess unreacted modified nucleotide. The duplex was again dissolved in the reaction buffer and the PEX reaction was initiated by adding natural dNTPs and KlenTaq/KOD. Reactions in the presence of both the polymerases produced the full-length product (Figure 4C, lanes 4 and 6).

It is worth mentioning that the integration of the fluorine atom in the heterocyclic ring does not hamper the single nucleotide incorporation efficiency of the KlenTaq DNA polymerase (see section 2.2.3). The KOD polymerase exhibits misincorporation of the FBFdUTP and fails to provide a single modified product, so it is not suitable for single nucleotide incorporation. However, both enzymes accommodate the modification and successfully elongate the modified duplex to yield a full-length product. Collectively, this data guides a suitable position to introduce small modifications in the heterocycle ring which will not impair this enzymatic labeling strategy.

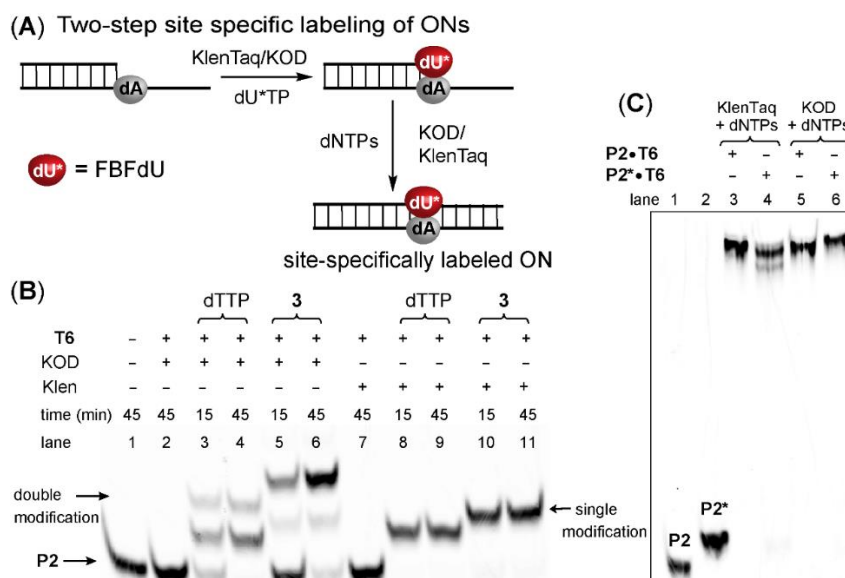


Figure 4. (A) A schematic diagram for the site-specific functionalization strategy of DNA with dU*TP, which represents FBFdUTP **3**. (B) KOD DNA polymerase added two consecutive nucleotides to the 3'-end of primer **P2** in the presence of dTTP or FBFdUTP **3**. KlenTaq DNA polymerase provided exclusively singly-modified ONs. (C) KlenTaq and KOD DNA polymerase successfully extended the 3'-FBFdU-labeled primer **P2*** to yield the full-length site-specifically functionalized ON product **P2****. See Table 1 and Table 2 for ON sequences.

Table 2. ESI-Mass and yield data for DNA ON products.

Product	Calculated Mass	Found Mass	Isolated yield (nmole)	Isolated yield (%)
5'-FAM,3'-FBFdU labeled ON P2*	6726.47	6727	3.4 ^{a,b}	65
3'-FBFdU labeled ON P3*	6188.93	6189.87	nd ^b	68

^aIsolated yields are given for a reaction scale of 5 nmol. ^b oligonucleotides were further synthesized by solid-phase synthesis for fluorescence and ¹⁹F NMR assay. See below Table 3 for the ON sequences.

^cThe yield for the product was not determined as it was isolated as a duplex (**P2*•T6**).

Table 3: Sequences of modified ONs.

Modified ONs	Sequence
P2*	5' FAM-GGA GCT CAG CCT TCA CTG C- FBFdU 3'
P3*	5' GGA GCT CAG CCT TCA CTG C- FBFdU 3'
P2**	5' FAM-GGA GCT CAG CCT TCA CTG C- FBFdU -GC TAA ACT ATA TTG CTA GCG CAC 3'

3.2.4 Binary complex (duplex-enzyme) formation by fluorescence

To accept multiple modifications to the DNA strand, the enzyme must accommodate the 3'-end modified primer-template duplex in the enzyme pocket before processing the cognate nucleotides to give a full-length primer extension product. Hence, we wanted to investigate the

differences in the binary complex formation by two different family of DNA polymerases using the responsiveness of the nucleotide probe by fluorescence spectroscopy (Figure 5).

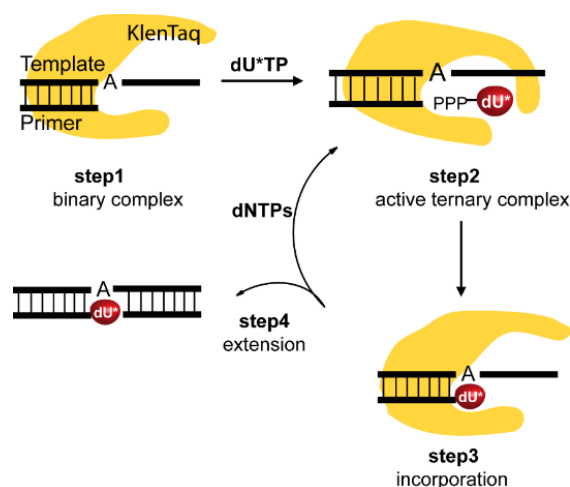


Figure 5. Catalytic steps for enzymatic incorporation.

We first monitored the effect of modification on the binary complex formation by performing the electrophoretic mobility shift assay (EMSA) under native conditions. Unmodified **P2•T7** and 5'-FAM-3'-FBFdU-modified **P2*•T7** duplexes were incubated independently with increasing concentrations of KlenTaq or KOD DNA polymerases. Due to the formation of a stable binary complex between enzyme and duplex it shows slower mobility than the duplex alone in the native PAGE (Figure 6). EMSA data revealed that FBFdU-modified and native-duplexes show almost similar binding affinity for both the KlenTaq ($K_d = 0.28 \pm 0.04, 0.24 \pm 0.03 \mu\text{M}$, respectively) and KOD DNA polymerases ($K_d = 0.23 \pm 0.02, 0.27 \pm 0.02 \mu\text{M}$, respectively, Figure 7). These results demonstrate that modification does not disturb the binary complex formation. Further, the binary complex formation was studied under equilibrium conditions by observing the changes in the fluorescence profile of duplex **P3*•T7** as a function of enzyme concentration. 3'-FBFdU-modified primer (**P3***) was utilized in this study. Modified duplex **P3*•T7** was incubated with varying concentrations of KlenTaq/KOD DNA polymerase at 55 °C for 20 min, and fluorescence was recorded by exciting the sample at 330 nm (Figure 8A). Upon increasing the concentration of the KlenTaq and KOD DNA polymerase, a 6-fold and 3-fold increase, respectively, in the emission intensity with negligible changes in the emission maximum was observed (Figure 8B, 8C). A plot of normalized fluorescence intensity *versus* enzyme concentration gave an apparent K_d of $0.36 \pm 0.01 \mu\text{M}$ for KlenTaq, whereas for KOD polymerase, it was $0.06 \pm 0.02 \mu\text{M}$ (Figure 8D).

FBFdU is highly sensitive towards its micro-polarity and shows enhancement in the fluorescence intensity with significant red-shift with increasing polarity in the medium.¹⁴ For example, in methanol, it shows λ_{em} at 418 nm, whereas in water, it shifts to 437 nm with an

increase in the fluorescence intensity. On the other hand, the presence of a rotatable aryl-aryl bond between fluorobenzofuran and uracil ring in the nucleoside analog makes it a molecular rotor element.^{49,50} FBFdU exhibits an increase in fluorescence intensity and lifetime due to the rigidification of the fluorophore in a more viscous medium. In the duplex state (**P3*•T7**), the FBF ring projected in the major groove is like to experience partial stacking interaction with neighboring bases, and hence, shows low fluorescence. However, once it forms a stable binary complex with the enzyme, the fluorobenzofuran ring likely gets rigidified in the enzyme pocket and exhibits enhanced fluorescence intensity. An emission maximum of 425 nm indicates that the modification is located in an environment close to methanol polarity.

We performed a competitive displacement assay to prove that the increase in the fluorescence intensity is due to the specific binding of the modified duplex into the enzyme pocket. A binary complex, consisting of duplex (**P3*•T7**) and enzyme (duplex concentration 50 nM and KlenTaq/KOD concentration 1 μ M) was prepared, which exhibits fluorescence intensity near saturation. The complex was titrated with increasing concentrations of an unlabeled DNA duplex (**P3•T7**) (Figure 8A). Introducing the unlabeled duplex led to a gradual reduction in fluorescence intensity. This phenomenon can be attributed to the displacement of the less emissive modified duplex from the active site (Figure 9A, 9B). Notably, even with a 20-fold excess of the unmodified duplex, complete displacement of the modified duplex from the enzyme did not occur. An electrophoretic mobility shift assay (EMSA) further validated this competitive displacement process (Figure 10). Similar to the challenging displacement of a binary complex formed by the modified duplex, the complex involving the enzyme and the unmodified duplex exhibited robust stability, requiring an excess of the modified duplex to disrupt the complex (Figures 9 and 10).

Fluorescence studies provide further insights into the processivity of the KlenTaq and KOD DNA polymerases. Moreover, both the enzymes efficiently accommodated the modification in the enzyme cavity and formed a stable binary complex to generate a site-specifically labeled product (Figure 4C). However, the apparent K_d values determined from the fluorescence experiment revealed that KOD binds to the 3'-FBFdU-modified duplex with higher binding affinity than KlenTaq (almost 6 times). The observed differences in the binding efficiency for KOD and KlenTaq DNA polymerase can be speculated by comparing the crystal structures of the binary complex, respectively (Figure 8E, 8F). Two significant factors might be responsible for these differences: (i) in the active site of KOD DNA polymerase, the conformation of the duplex changes to the B-form of DNA, which provides a wide-opening in the major groove to accommodate the larger modification and (ii) additionally, thumb domain

interactions differ between KOD and KlenTaq. In KlenTaq, the domain extends into the major groove, potentially causing clashes with modifications at C5 positions (Figure 8E, 8F).

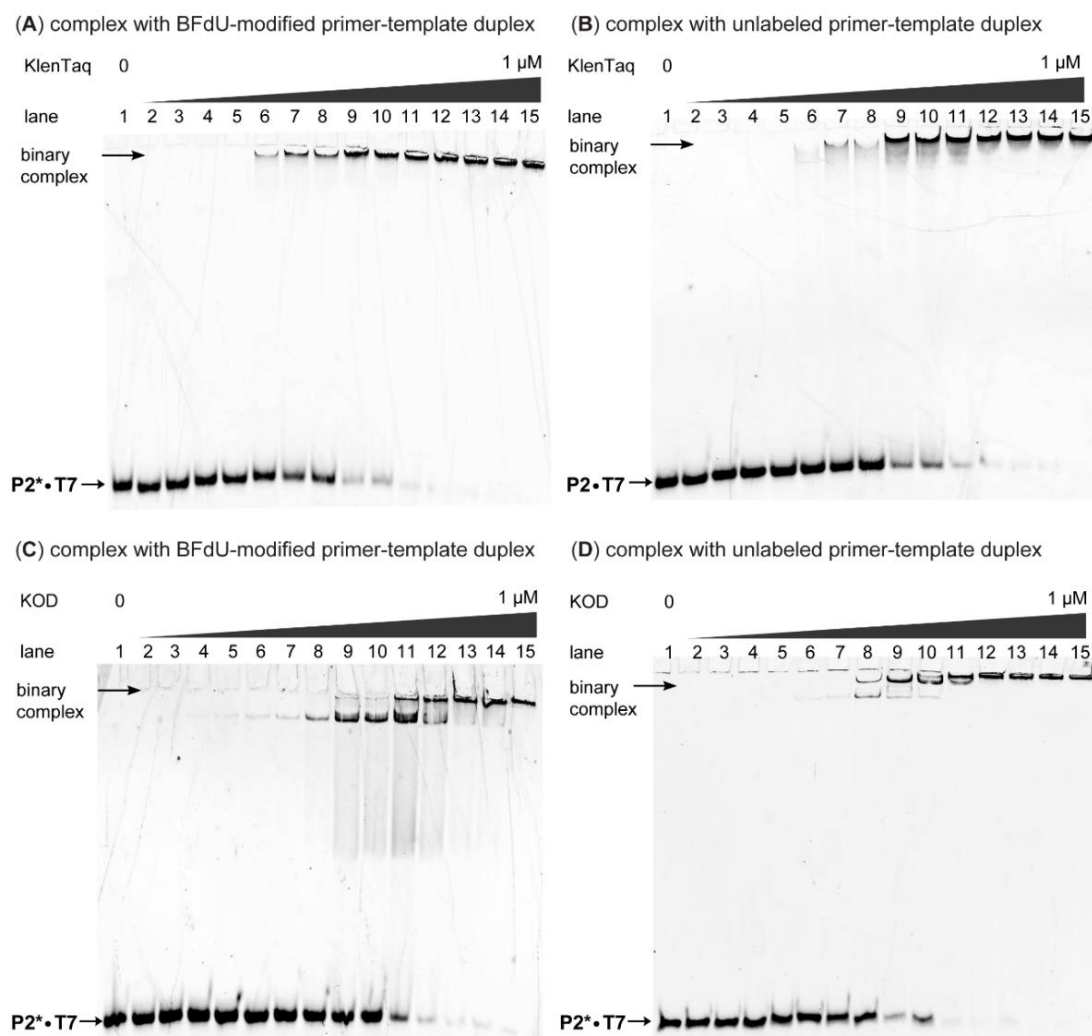


Figure 6. Monitoring binary complex formation using EMSA. Complex formation with (A) BFDU-modified primer-template duplex (P2*•T7) and (B) unmodified primer-template duplex (P2•T7) with increasing concentrations of the KlenTaq. (C) and (D) Complex formation with KOD DNA polymerase with BFDU-modified (P2*•T7) and natural primer-template duplex (P2•T7), respectively.

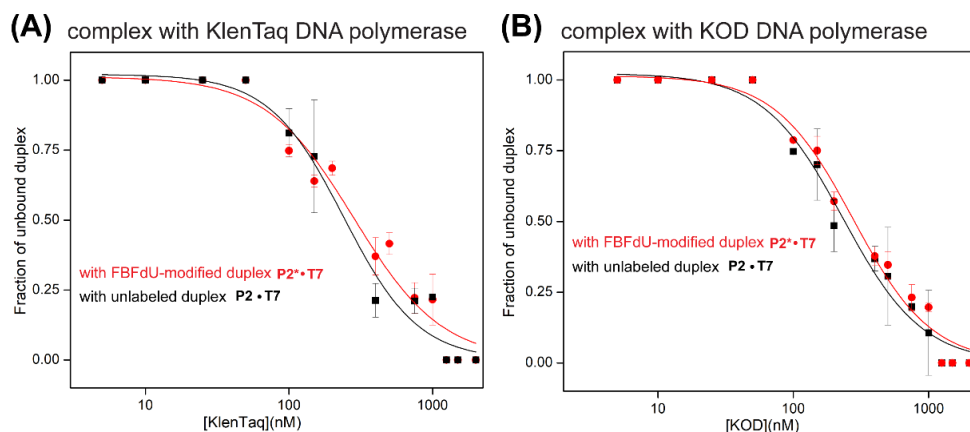


Figure 7. Curve fits for the fraction of unbound duplex with increasing concentration of (A) KlenTaq and (B) KOD DNA polymerase.

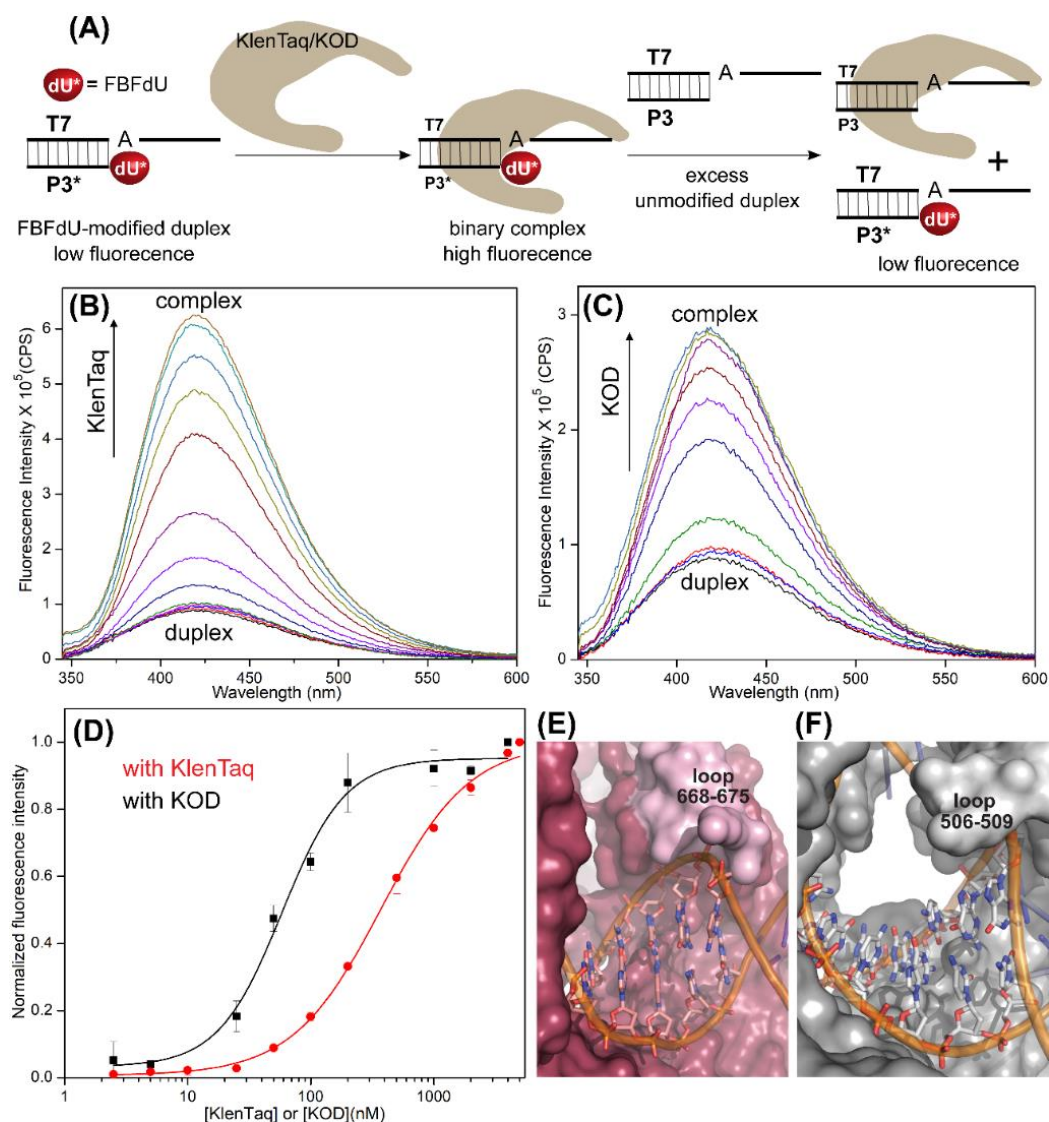


Figure 8. FBFdU reports the binary complex formation by fluorescence. (A) The schematic diagram for monitoring the binary complex formation and competitive displacement assay by fluorescence. (B) and (C) Fluorescence spectra for the duplex P3*•T7 (50 nM) with increasing concentrations of KlenTaq and KOD DNA polymerase, respectively (0–5 μ M). Samples were excited at 330 nm with excitation and emission slit widths of 7 nm and 8 nm, respectively. (D) Curve fits for the binding of the KlenTaq and KOD. Normalized fluorescence intensity λ_{em} (430 nm) versus the concentration of the enzyme was plotted (E) and (F) Active site of KOD DNA polymerase (raspberry, PDB: 5OMF)⁵¹ and KlenTaq DNA polymerase (grey, PDB: 3RTV).⁵²

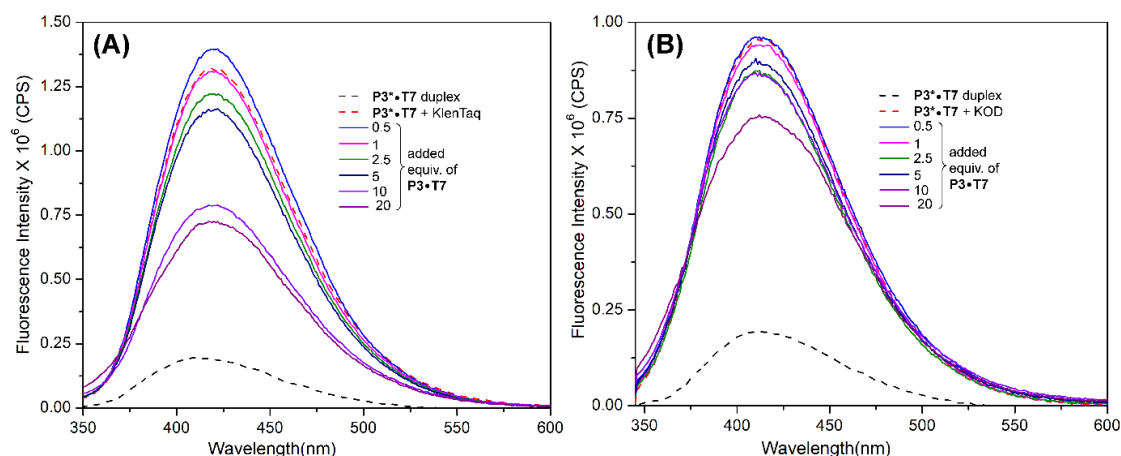


Figure 9. Fluorescence was utilized in a competitive displacement assay. The FBFdU-modified primer-template duplex $P3^* \bullet T7$ (50 nM) exhibited minimal fluorescence (depicted by the black dashed line). (A) A binary complex of $P3^* \bullet T7$ (50 nM) and KlenTaq (1 μM) and (B) KOD (1 μM) shows very high fluorescence (red dashed line). The introduction of increasing amounts of unlabeled duplex ($P3 \bullet T7$) led to the displacement of the complex, resulting in a gradual reduction in fluorescence (illustrated by the solid lines).

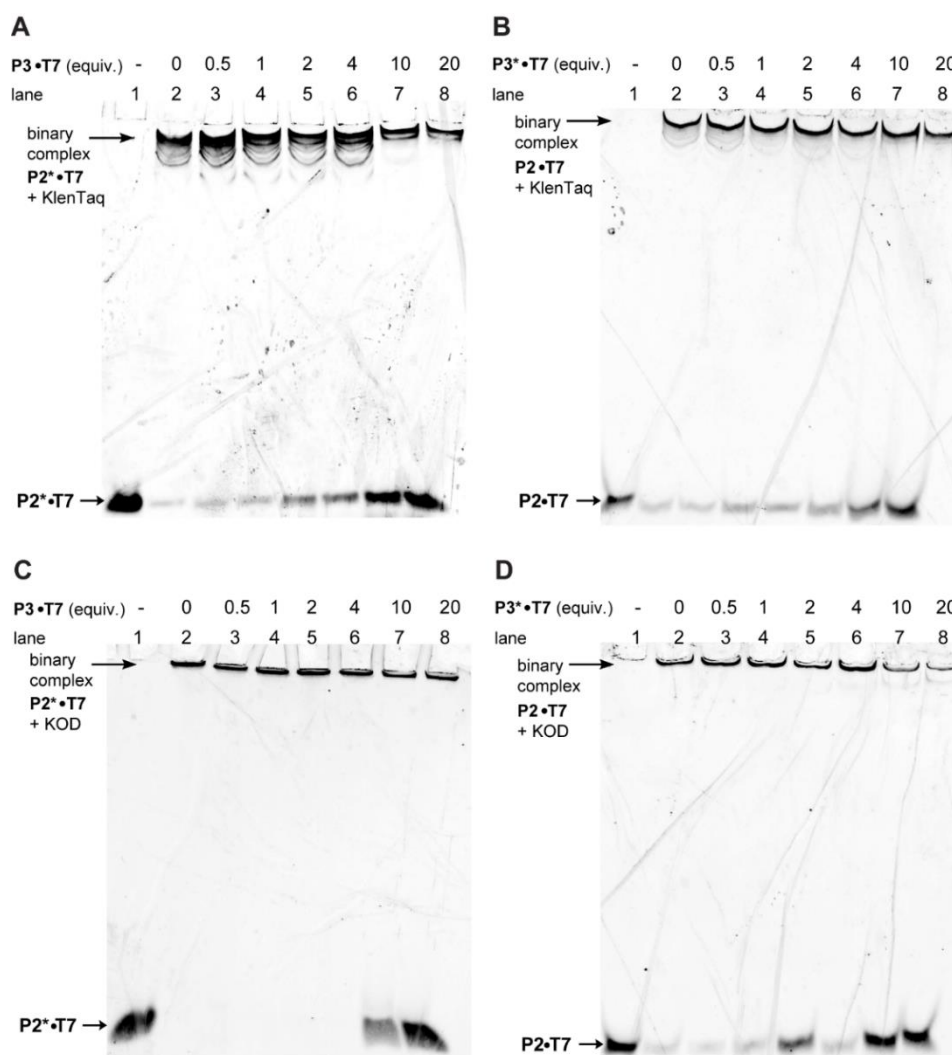


Figure 10. Competitive displacement assay using EMSA. Titration of the unmodified duplex **P3•T7** to a binary complex consisting of **P2*•T7** (5'-FAM, 3'-BFdU-modified primer-template duplex) and (A) KlenTaq or (C) KOD led to the displacement of duplex **P2*•T7**, evident by the appearance of a lower band. **P3** shares the same sequence as **P2** but lacks the 5'-FAM label. The addition of the 3'-BFdU-modified primer-template duplex **P3*•T7** to a binary complex of **P2•T7** and (B) KlenTaq or (D) KOD, resulted in the displacement of duplex **P2•T7**, indicated by the emergence of a lower band. In both cases, the binary complexes formed with native and BFdU-modified duplexes exhibited considerable stability, and even an excess of the respective duplexes could not entirely disrupt the complexes, as also observed in the fluorescence assay.

3.2.5 Nucleotide substrate fluorescently reports its incorporation into DNA ONs

Free nucleotide BFdUTP **3** exhibits higher fluorescence intensity than the 3'-BFdU-labeled primer-template duplex, which encouraged us to design a method to detect the KlenTaq and KOD DNA polymerase activity (Figure 11A). Adding BFdUTP **3** to the primer•template duplex (**P3•T7**) or polymerases alone does not show changes in the fluorescence intensity. Rewardingly, when the BFdUTP **3** (dashed black curve, $\lambda_{em} = \sim 445$ nm) was added to samples of binary complex (enzyme and primer•template duplex **P3•T7**), each containing increasing concentrations of the KlenTaq or KOD DNA polymerase (0–50 nM), a considerable decrease in the fluorescence intensity accompanied by a blue-shift ($\lambda_{em} = \sim 430$ nm) in the emission maximum was observed. The observed outcome is because of the incorporation of a single BFdU analog at the end of the primer resulting in the generation of a less emissive BFdU-labeled duplex **P3*•T7** (Figure 11C and 11D). In a control experiment, a binary complex made of KlenTaq/KOD and a 3'-ddC-modified primer-template duplex **P5•T7** were added to BFdUTP. Due to the presence of a blunt 3'-end, there was no incorporation and thus, no decrease in fluorescence intensity was observed (Figure 11C and 11D, green curve). A plot of decrease in normalized fluorescence intensity against enzyme concentration shows that after a particular enzyme concentration, there is no significant decrease in fluorescence intensity until it reaches saturation (Figure 11E). Point to be noted that KOD DNA polymerase at a very low enzyme concentration shows a sudden drop in fluorescence intensity and quickly reaches saturation compared to KlenTaq. This data suggests that KOD DNA polymerase accepts the modifications more efficiently than KlenTaq. To prove that quenching in fluorescence is due to single nucleotide incorporation, an experiment was performed with a 5'-FAM labeled primer-template duplex (**P2•T7**) under similar conditions and analyzed by denaturing PAGE. It has been observed that KlenTaq and KOD give single-modified products under these conditions (Figure 11B). Taken together, BFdU efficiently reports its incorporation by fluorescence changes, and using this assay, we can distinguish the incorporation efficiency of the two DNA polymerases.

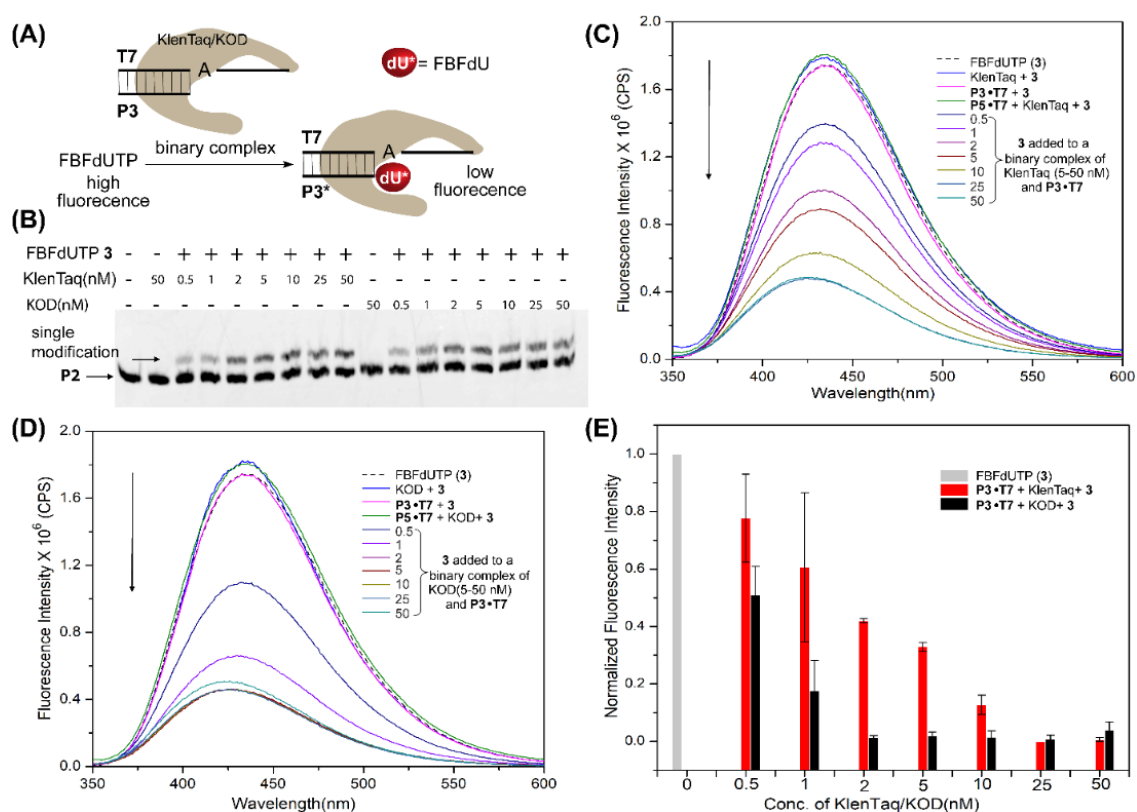


Figure 11. FBFdU enables monitoring the nucleotide analog incorporation by KlenTaq and KOD using fluorescence spectroscopy. **(A)** The schematic diagram for assay setup. **(B)** Gel images showing the single nucleotide incorporation of FBFdUTP (100 nM) in the presence of P2•T7 (200 nM) with increasing concentration of KlenTaq and KOD concentration (0–50 nM). **(C)** and **(D)** A graph illustrates the decrease in fluorescence intensity of FBFdUTP (100 nM) with increasing KlenTaq and KOD concentration (0–50 nM) when the P3•T7 primer•template duplex (200 nM) is present. The incorporation of FBFdUMP at the 3'-end leads to decreased fluorescence intensity, indicating lower fluorescence in the labeled duplex. Control experiments demonstrate minimal changes in fluorescence intensity when KlenTaq/KOD (50 nM) or P3•T7 (200 nM) is added to FBFdUTP (100 nM). The addition of a binary complex composed of (50 nM) of KlenTaq/KOD and (200 nM) P5•T7 duplex (P5 contained a ddC at the 3'-end) to FBFdUTP (100 nM) does not exhibit a significant reduction in fluorescence intensity. All samples are excited at 330 nm with excitation and emission slit widths set at 7 nm and 8 nm, respectively. **(E)** A graphical representation of the decrease in normalized fluorescence intensity versus KlenTaq (red) and KOD (black) concentration.

3.2.6 Binary complex (duplex-enzyme) formation and nucleotide incorporation by ^{19}F NMR

Next, we aimed to exploit the potential of the ^{19}F label in monitoring the binding process involved in the DNA polymerization process by ^{19}F NMR (Figure 12A). For this purpose, a 3'-end FBFdU-labeled primer (P3*) was synthesized via solid-phase ON synthesis and purified by PAGE. The integrity of the product was confirmed by ESI mass analysis. The modified primer P3* in the KlenTaq binding buffer exhibits a peak at -122.0 ppm. When FBFdU-modified primer P3* was annealed with the complementary ON T7, the signal shifts slightly

to -122.23 ppm. Rewardingly, incubation of 3'-FBFdU-modified duplex (**P3*•T7**) with 0.5 equivalent of KlenTaq results in the appearance of a distinct peak at -121.87 ppm corresponding to the binary complex with a concomitant decrease in the peak for the duplex. Further, with increased KlenTaq concentration (1eq.), the peak intensity for the free duplex (**P3*•T7**) almost diminished, and the sharp peak at -121.87 ppm intensified (Figure 11B, purple line). These observations indicate that the distinct peak at -121.87 is due to the binding of the 3'-end modified primer-template duplex in the enzyme pockets (Figure 12B).

In FBFdU, the fluorophore and ^{19}F NMR labels are incorporated into the same electronic environment and the electron density on the fluorine atom will be influenced by various factors like the fluorescence properties as mentioned above. Based on the sensitivity of the analog, the difference in the environment around the nucleoside probe in the free duplex form and the duplex bound to the protein is likely the reason for the observed difference in the chemical shift exhibited by the probe. Further, to prove that the appearance of the new peak is due to the formation of the binary complex, an electrophoretic mobility shift assay (EMSA) was conducted under non-denaturing conditions (Figure 12C). Both unmodified **P2•T7** and 5'-FAM,3'-FBFdU-modified **P2*•T7** duplexes were treated with increasing concentrations of KlenTaq polymerases as used in the ^{19}F NMR experiment. EMSA results also revealed the formation of a binary complex (Figure 12C, lanes 2 and 3). Further, to assess the stability of the binary complex, we subjected it to heating at 95°C for 20 minutes and conducted an NMR experiment. Interestingly, no discernible difference in peak intensity at -121.87 ppm was observed, indicating the stability of the binary complex even at elevated temperatures. These findings further evidenced the thermostable nature of KlenTaq, demonstrated by forming a stable complex with the modified duplex at higher temperatures. Similar experiments were performed with KOD DNA polymerase to monitor its binding with the modified duplex. We observed a decrease in the peak intensity in the duplex with increasing enzyme concentration. Unfortunately, the peaks were not well resolved. Due to the high molecular weight/size of the KOD DNA polymerase, the bound complex might exhibit a broader spectrum (Figure 13).

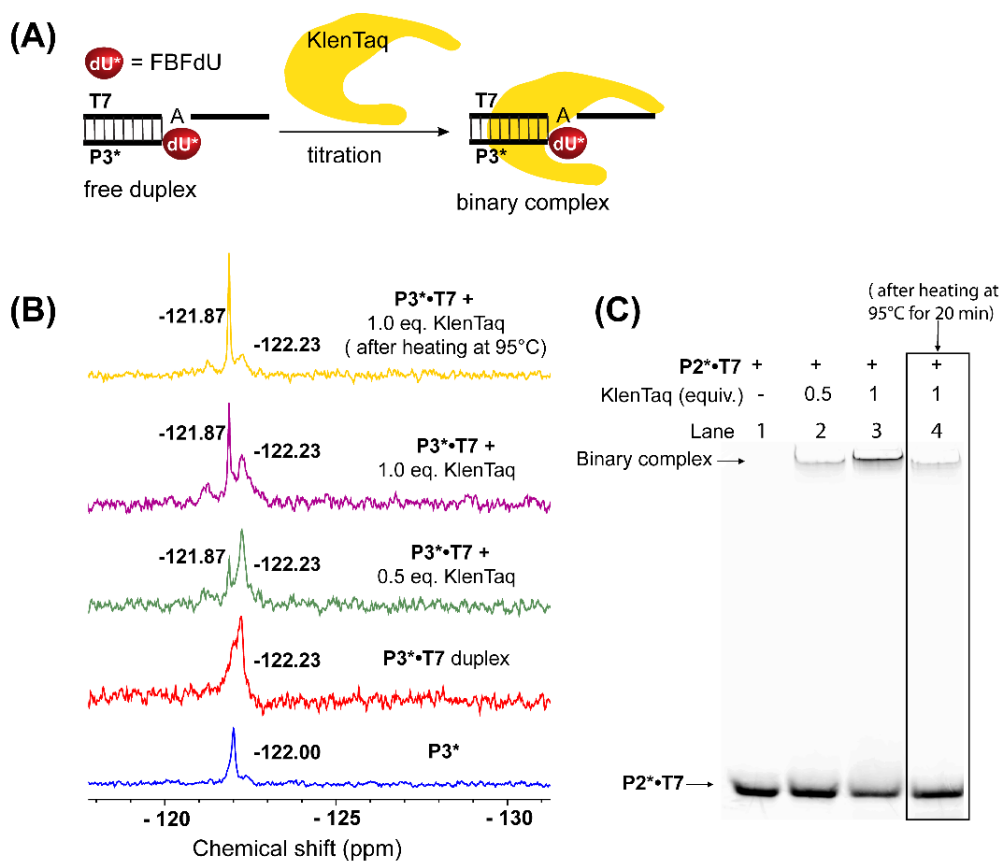


Figure 12. Monitoring the binary complex formation by ^{19}F NMR experiment. **(A)** Schematic representation for the NMR experiment. **(B)** ^{19}F NMR signals for primer **P3***, primer-template duplex **P3*•T7** in 1X binding buffer (10 mM Tris-HCl, 50 mM KCl, 5 mM MgCl_2 , pH 7.9) containing 10% glycerol. Changes in ^{19}F NMR signal of the duplex (10 μM) in KlenTaq binding buffer containing 10 % glycerol as a function of increasing KlenTaq concentration. **(C)** EMSA for monitoring the binary complex formation of **P2*•T7** (10 μM) with increasing concentration of KlenTaq (5–10 μM).

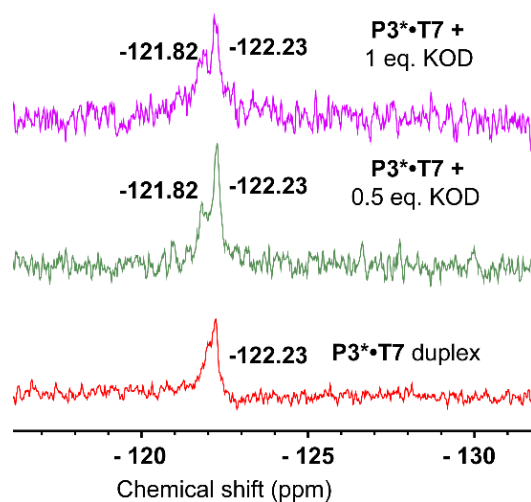


Figure 13. Monitoring the binary complex formation between primer-template duplex **P3*•T7** and KOD in 1X binding buffer (10 mM Tris-HCl, 50 mM KCl, 5 mM MgCl_2 , pH 7.9) containing 15 % glycerol. by ^{19}F NMR experiment.

Next, we became curious to monitor the incorporation of FBFdUTP **3** by using ^{19}F NMR spectroscopy (Figure 14A). Although FBFdUTP faithfully reports its incorporation by KlenTaq *via* changes in its fluorescence properties, it is challenging to predict the relative population of different components present at equilibrium conditions. During the incorporation process, a population of different components can be present, such as (i) FBFdUTP, (ii) free unmodified and modified primer-template duplexes, (iii) binary complex of modified primer-template duplex with the enzyme. Free nucleotide FBFdUTP **3** shows a sharp peak at -122.36 ppm (Figure 14B, cyan line). Once FBFdUTP is added to the binary complex of an unmodified duplex (**P3•T7**) (**T7** guides incorporation of single modification at the 3'-end) and KlenTaq, the peak for the FBFdUTP diminishes, and the peak corresponded to the bound state (-121.87 ppm) of the modified duplex increases (Figure 14B, brown line). This data suggests that after the incorporation, the modified duplex exclusively stays in the bound state for further elongation by primer extension reaction. Taken together, FBFdU provides a greater understanding on the enzymatic incorporation process by providing distinct ^{19}F NMR peaks. This study sheds light on the processivity of KlenTaq. First, it forms a stable binary complex before accepting the modification, and second, after incorporating the modifications, it accommodates the modified duplex in the enzyme cavity to facilitate the subsequent addition of natural nucleotides to produce a full-length product.

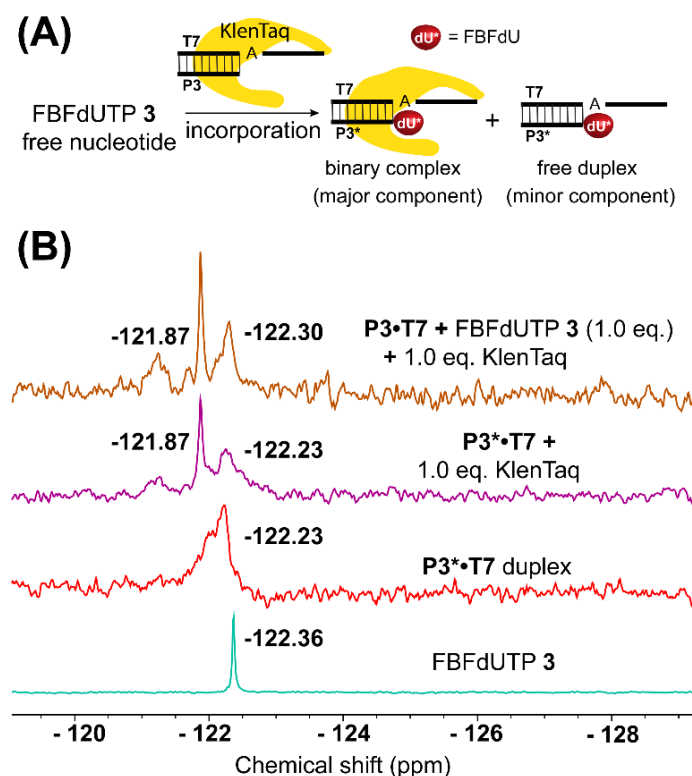


Figure 14. FBFdU reports the enzymatic incorporation by ^{19}F NMR experiment. **(A)** The schematic diagram represents the experimental setup. **(B)** ^{19}F NMR signals for FBFdUTP **3** (10 μM), primer•template duplex **P3•T7** and the binary complex composed of **P3•T7** and KlenTaq in a 1:1 molar ratio in 1X klenTaq binding buffer (10 mM Tris–HCl, 50 mM KCl, 5 mM MgCl_2 , pH 7.9) containing 10% glycerol. Changes in the ^{19}F NMR signal of the FBFdUTP (10 μM) in the presence of KlenTaq DNA polymerase (10 μM) and the unmodified primer-template duplex **P3•T7** (**T7** guides the single nucleotide incorporation).

3.2.7 X-ray crystal structures of ternary complexes provide insights on the incorporation mechanism

Our objective is twofold: firstly, to probe the active-site conformational changes essential for accommodating the additional fluorine atom within the benzofuran ring. Secondly, inspired by the size-dependent processing differences in PEX reactions that were previously noted (Chapter 2), we sought to gain insights into the plasticity of the active site when incorporating modifications of varying heterocycle size. We envisioned to obtain structural basis of the incorporation mechanism by capturing snapshots of the catalytic steps poised for the incorporation of both the smaller heterocyclic modification, SedUTP, and the more expansive moiety, FBFdUTP, and subsequently comparing these ternary crystal structures with BFdUTP.

High throughput crystallization screens with FBFdUTP and SedUTP in the presence of KlenTaq and KOD DNA polymerases were set up as illustrated in Figure 15A. Template **T8** would direct the incorporation of ddCMP at the 3'-end of the primer and terminate the polymerization due to the absence of the 3'-OH group. This setup would allow the modified nucleotides to be captured at the nucleotide-binding site, ready to be added to the growing chain by phosphodiester bond formation. In the presence of KOD, we obtained microcrystals, which did not provide quality diffraction data. Rewardingly, high-resolution structures of the KlenTaq ternary complexes composed of FBFdUTP and SedUTP were obtained (Figure 15B, 15E). Superimposition of structures reveal that the ternary complexes with FBFdUTP and SedUTP adopt a structure that closely resembles the native complex and also the structure with BFdUTP (PDB: 3RTV, see Figure 15B, 15C and 15D, 15E, respectively). The nucleotide analogs form Watson-Crick pairing with the dA residue of the template (see Figure 15D and 15G). Fluorobenzofuran and selenophene moieties with respect to the uracil rings exhibit slight non-coplanarity and partially stack with ddCMP inserted at the 3'-end of the primer (Figure 15D, 15G and figure 16A, 16B).

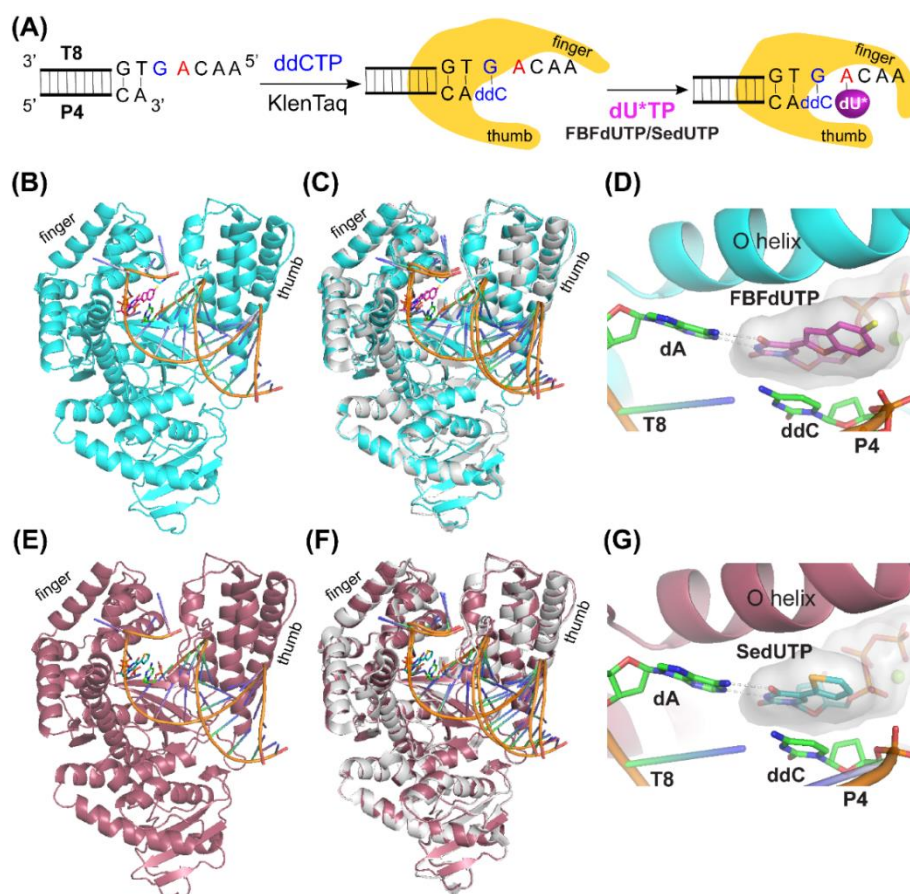


Figure 15. (A) Schematic diagram for the crystallization setup to trap the fluorescent nucleotide probes (FBFdUTP, SedUTP) in the ternary complex. (B) and (E) is the overall structure of the KlenTaq FBFdUTP ternary complex (cyan) and KlenTaq-SedUTP ternary complex (raspberry), respectively. (C) and (F) is the superimposition of KlenTaq-FBFdUTP (cyan) and KlenTaq-SedUTP (raspberry) with a native KlenTaq ternary complex (gray; PDB: 3RTV), respectively. (D, G) Close-up view of the active site showing the bound nucleotide analogs. The Watson-Crick base pairing between the nucleotide analogs and dA of the template strand is indicated with dashed lines. In all structures, FBFdUTP is highlighted in magenta and SedUTP in deep green.

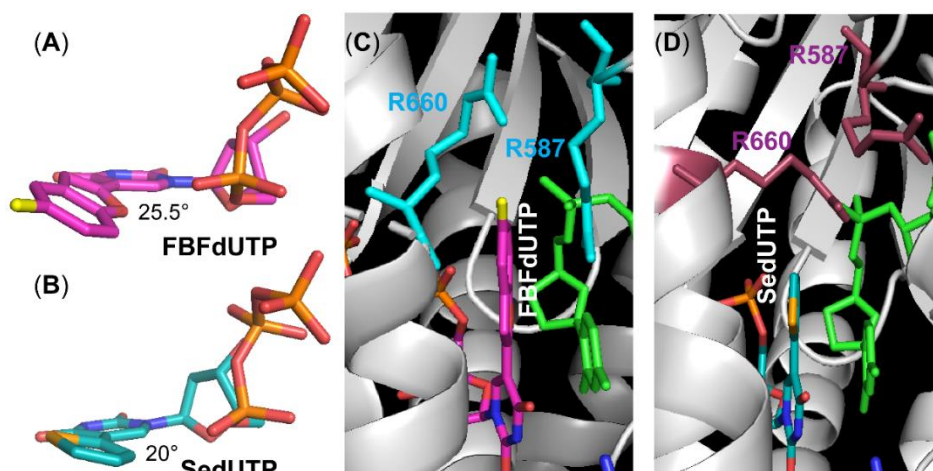


Figure 16. (A) Fluoronenzofuran and (B)selenophene with uracil rings are non-coplanar. The dihedral angle of the non-planar FBFdUTP **3** and SedUTP is depicted. (C) Another view of the cleft in which the FBFdU ring (magenta) is encased is shown. Arg660 and Arg587 residues are colored cyan. Protein and DNA are shown in grey and green colors, respectively. (D) SedU ring (deep green) is encased in enzymatic pocket, Arg660 and Arg587 residues (colored raspberry) have not moved. Protein and DNA are shown in grey and green colors, respectively.

Structure with FBFdUTP

In the active form of the ternary complex, the finger region adopts a closed conformation, and the O helix aligns atop the FBFdUTP bound in the active site much like BFdUTP (Figure 17A). In order to accommodate the FBFdU modification, two crucial amino acid residues namely Arg660 and Arg587, assume an alternate conformation as compared to the native structure (Figures 17A and 17B). Typically, Arg660 makes a hydrogen bond with the primer 3'-terminus phosphate backbone, stabilizing the closed active state of the enzyme.^{53,54} However, Arg660 flips out and makes room to accommodate FBF modification. During this process, Arg587 comes closer to the FBF ring and interacts with the phosphate backbone of the penultimate base of the primer strand. This alternate conformation produces a crevice to encase the heterocycle modification in the active site (Figure 16C). Despite these minor differences, the contacts and distance required to form a phosphodiester bond are well maintained in structures containing the modified nucleotide (3.9 Å) and a natural substrate (3.8 Å for 3RTV).⁵² This similarity indicates that the modification does not impede the primer extension process (Figure 18E, G). Further, the incorporation of a fluorine atom into the benzofuran ring does not interfere with the pocket.

Structure with SedUTP

Interestingly, the structure of the ternary complex with SedUTP reveals a conformation different from FBFdUTP. While the finger domain is in a closed state, and the O-helix is situated on top of the SedUTP (see Figure 17D), the gatekeeper residues Arg587 and Arg660 adopt a conformation similar to the structure formed in the presence of a native nucleotide (Figure 17D, 17E). Further, the distance between the α -phosphorus of SedUTP and the 3'-C of the primer is the same as in for the nature substrate (3.8 Å for 3RTV) (Figure 18E, H). These results indicate that SedUTP fits well into the KlenTaq pocket without disturbing the amino acid residues.

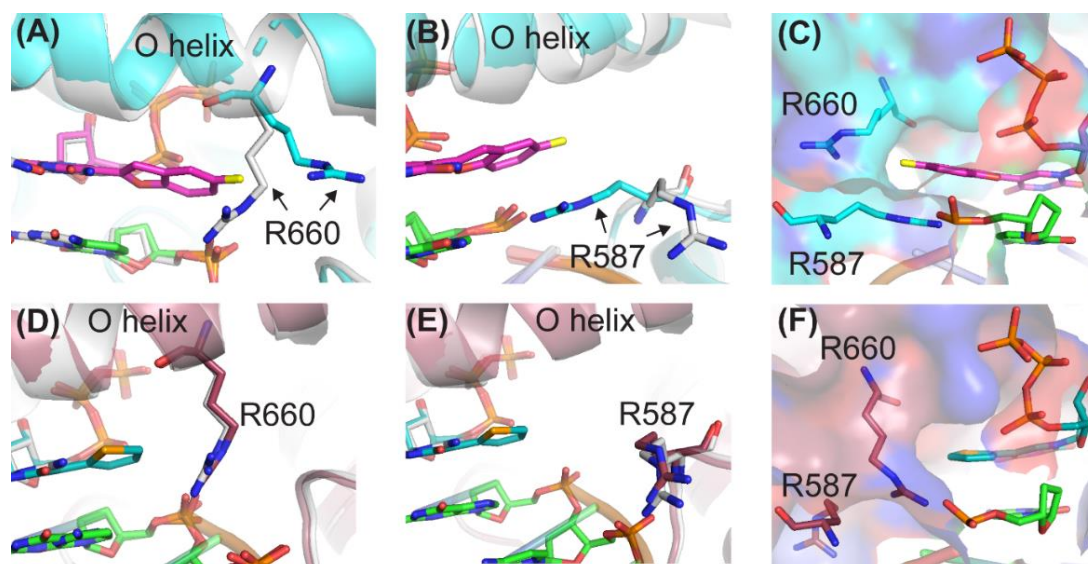


Figure 17. (A, B and D, E): Superimposition of KlenTaq-FBFdUTP (cyan) and KlenTaq-SedUTP (raspberry) and KlenTaq (gray; PDB: 3RTV) highlighting the different orientations of Arg660 and Arg587, respectively. (C, F) A cleft in which the fluorobenzofuran and selenophene ring are located in the active site is shown.

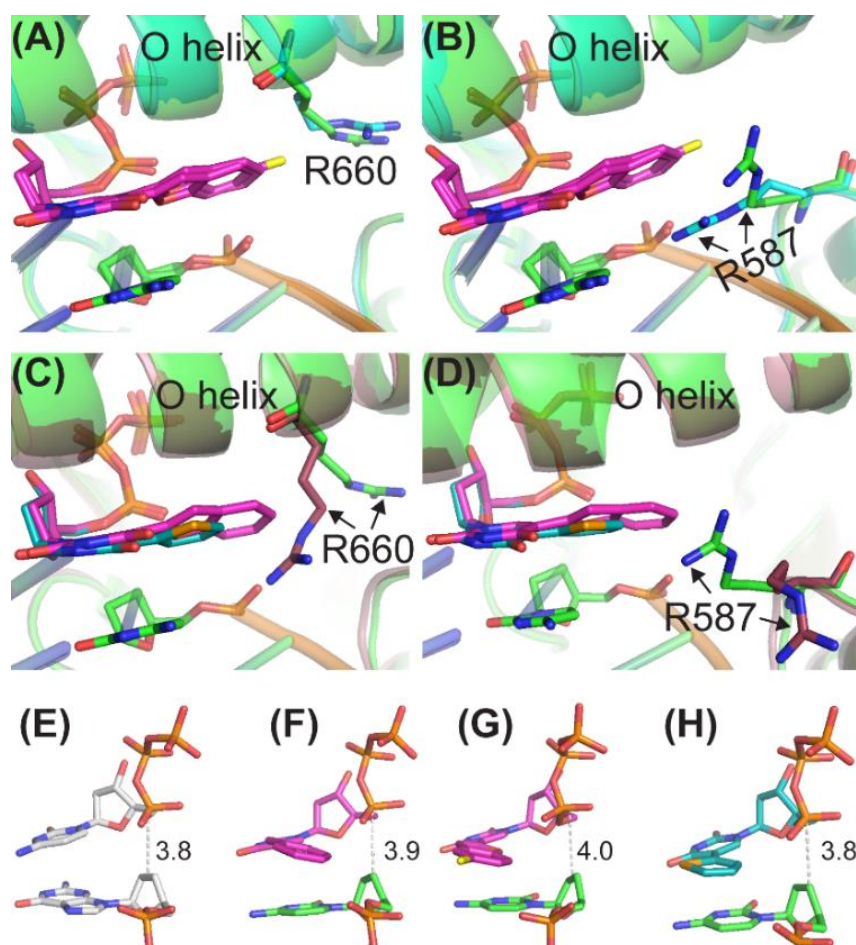


Figure 18. (A, B) superimpositions of active site conformations between KlenTaq-BFdUTP (PDB: 7OWF) (green color) and KlenTaq-FBFdUTP (cyan color). (C, D) superimposition of the active-site of KlenTaq-BFdUTP (green) with KlenTaq-SedUTP (raspberry). (E, F, G, H) Close-up view showing the distance between the α -phosphorus and 3'-end of the primer in the ternary complex structures of KlenTaq binds to dCTP ((PDB: 3RTV), KlenTaq-BFdUTP (PDB: 7OWF), KlenTaq-FBFdUTP and KlenTaq-SedUTP, respectively.

3.2.7.1 Processivity of KlenTaq depends on the size of the heterocycle-modified nucleotides

To assess the size-dependency on the processing of modified nucleotides by KlenTaq DNA polymerase, we compared the crystal structures and performed a competition PEX reaction assay. We observed that in the single nucleotide incorporation assay, all the modified nucleotides gave exclusively a single nucleotide incorporated product (Chapter 2, Figure 4B). Here, we compared the crystal structures of KlenTaq-BFdUTP (PDB: 7OWF), its expanded modification (KlenTaq-FBFdUTP) and the smaller modification (KlenTaq-SedUTP). This comparison sheds light on how the size of the heterocycle modification regulates the efficiency of incorporation by KlenTaq DNA polymerases. It is worth mentioning that the flexibility of the two amino acid residues, Arg660 and Arg587, is responsible for the accommodation of the heterocycle in the enzyme's active site. Due to the integration of the fluorine atom in the benzofuran moiety, Arg587 moved further and placed just below the fluorobenzofuran ring. However, Arg660 holds up its position, which is almost similar to the benzofuran moiety (Figure 18A, B). The incorporation of a smaller heterocyclic moiety (selenophene) does not distort the amino acid structures (Figure 18C, D). To understand how these conformational changes in the amino acid residues affect the overall incorporation efficiency of the modified nucleotides, we performed a competition assay with the natural thymidine triphosphate (dTTP) and the modified nucleotides. The modified nucleotides were mixed at different ratios with the natural dTTP and subjected to single nucleotide incorporation assay in the presence of KlenTaq DNA polymerase and a 5'-FAM-labeled primer-template duplex (**P6•T8**) (**T8** guides the insertion of single thymidine analog at the 3'-end) (Figure 19A). Due to the slower mobility of the modified products compared to the natural one, we obtained two well-separated bands in gel electrophoresis analysis. Further, the band intensities were quantified and plotted to access the incorporation efficiency. In the case of benzofuran and fluorobenzofuran-modified nucleotides, experiments resulted in an approximate 35/1 ratio (dU*TP /dTTP) ratio to afford equal amounts of the single-added products (Figure 19B, C, E and F). This data demonstrates that while both the modified nucleotides were efficiently processed by the DNA polymerase in a single nucleotide incorporation reaction, KlenTaq exhibits greater preference for the natural

nucleotide. Interestingly, in the case of a smaller heterocyclic moiety, SedUTP, experiments showed that a ratio of 1/15 SedUTP/dTTP was required to produce equally modified products. This result indicates that KlenTaq accepts SedUTP almost fifteen times more efficiently than the natural nucleotide (Figure 19D, G).

We hypothesize that two major factors are attributing to these observations. First, the precatalytic state is potentially stabilized by the stacking between the primer 3'-end nucleotide with the incoming nucleotide. Second, due to the steric hindrance, the distortion in the amino acid residues Arg660 and Arg587, to accommodate the modified nucleotides in the enzyme pocket. While the larger heterocyclic modifications efficiently contribute to the stacking interactions in the precatalytic state, the disruption in the conformations is likely the dominating factor. The plasticity of the amino acids residues Arg660 and Arg587 though allows the incorporation of heterobycle-modified nucleotides, it diminishes the overall incorporation efficiency (Figure 17 and Figure 18A–D). Whereas, a smaller heterocyclic moiety like SedUTP perfectly fits in the enzyme pocket without disturbing the conformation of the active site. In addition, it stacks with the 3'-end nucleotide (here ddCMP) to stabilize the complex.

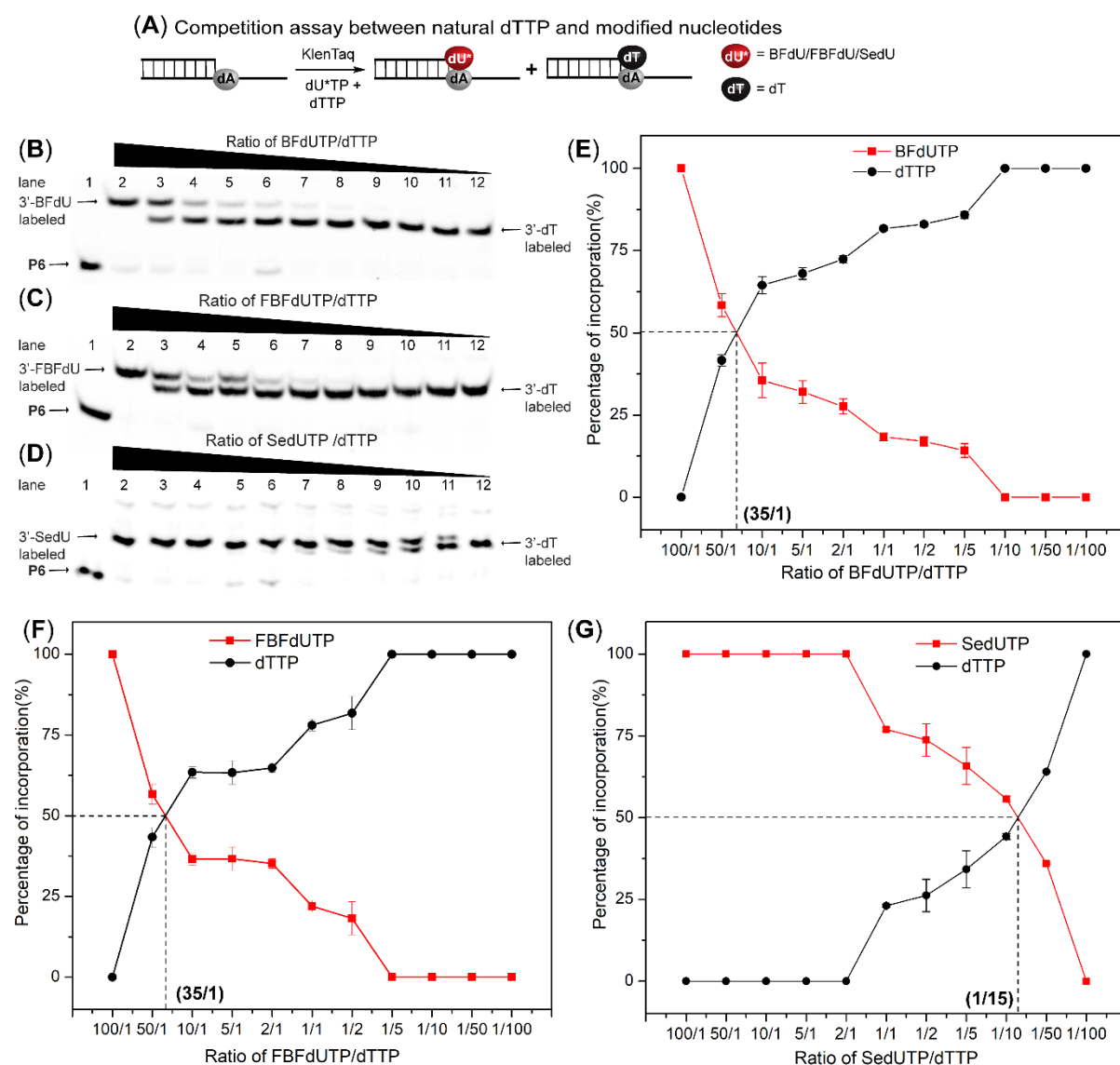


Figure 19. (A) Illustration of the competition reaction assay. (B, C, D) PAGE analysis of the competition experiment using KlenTaq DNA pol for BFdUTP, FBFdUTP, SedUTP, respectively. (E, F, G) Graphical readout of the PAGE analysis. The point of 50 % incorporation is indicated with a dashed line. All the data was collected in triplicate.

3.2.8 Size-dependency in incorporating multiple modifications in PCR reactions by KlenTaq and KOD DNA polymerases

The enzymatic synthesis of highly functionalized ONs is important for a broad range of applications,^{59–62} and this process depends on the incorporation of the multiple modified nucleotides by PEX and PCR. However, restriction in the incorporation of a large number of consecutive modifications into the DNA strand limits the scope of applications. In general, archaeal DNA polymerases from family B are well suited for this purpose compared to A family DNA polymerase.^{59–64} We aimed to understand the processing of the series of heterocyclic modified nucleotide analogs (BFdUTP, BTdUTP, FBFdUTP, and SedUTP) by PCR reactions in the presence of KlenTaq and KOD DNA polymerases. We chose a 98-mer

template ON strand, which would guide multiple modifications in the DNA strand (Table 4). To compare their incorporation efficiency, we have conducted the PCR reactions in the presence of dATP, dGTP, dCTP, and dTTP/dU*TP. To be processed in a PCR reaction, the modified nucleotides have to be efficiently incorporated by the enzyme and furthermore, the modifications must be accommodated in the template binding site in order to be recognized by the enzymes. KlenTaq, belonging to the family A DNA polymerase, is unable to produce full-length products for bicyclic modifications (Figure 20A, lane 4,5,6) and majorly forms truncated products. It is also unable to process smaller nucleotide SedUTP (Figure 20B, lane 6). On the contrary, KOD DNA polymerase belonging to family B DNA polymerases also had difficulty processing the larger modifications (Figure 20B, lane 4,5,6) but efficiently processed SedUTP (Figure 20B, lane 7). As a control experiment, we performed a PCR reaction in the absence of the natural dTTP and modified nucleotides, where both the enzymes were unable to produce the PCR products. These results demonstrate that the obtained PCR product is not because of the misincorporation (Figure 20A and B, lane 3).

Based on our examination of crystal structures, it is evident that the incorporation of a single fluorescent nucleotide analog, irrespective of the heterocycle modification, into the primer does not impede its translocation along the channel, leading to the production of complete PEX products (Chapter 2, section 2.2.2, Figure 4). However, the confined channel volume could potentially constrain KlenTaq's proficiency in incorporating multiple bulky modifications in close proximity. The smaller modification, SedUTP, is efficiently accepted by KlenTaq to produce full-length products for all the templates in the primer extension reaction and also demonstrates better efficiency than the natural nucleotide in the single nucleotide incorporation assay. However, in a PCR reaction with a longer template much less full-length product was formed. Additionally, in the active site of the KlenTaq DNA polymerase, A form of the duplex strands provides a smaller major groove volume to accommodate multiple modifications. Meanwhile, KOD DNA polymerase also faces difficulty in processing bulky modifications but efficiently accommodates the multiple smaller modifications in the enzyme and processes as effectively as natural dTTP. The ternary closed complex of KOD DNA polymerase features a larger channel volume in comparison to KlenTaq DNA polymerase (Chapter 2, section 2.2.7, Figure 7E, F) and also exhibits a more spacious crevice between its finger and thumb/palm domains, which provides greater room for accommodating multiple modifications. Even in our fluorescence data, we observed the differences in the channel volume reflected in the overall binary complex formation (Figure 8) as well as the incorporation efficiency (Figure 11). Within the active site of KOD DNA polymerase, the

presence of the B-form of DNA (section 3.2) provides an expanded area, particularly advantageous for accommodating major groove modifications, including those at C5-modified pyrimidines analogs. This characteristic appears to facilitate the seamless incorporation of consecutive modified nucleotides. Consequently, KOD DNA polymerase demonstrates comparable efficiency in processing both natural and unnatural substrates.

Table 4: sequences for PCR reactions.

ONs	Sequence
P_F	5' GACATCATGAGAGACATCGC 3'
P_R	5' CAAGGACAAAATACCTGTATTCCCTT 3'
T_{pcr}	5'GACATCATGAGAGACATCGCCTCTGGGCTAATAGGACTACTTCTAATCTGTACAGAGC AGATCCCTGGACAGGAAGGAATACAGGTATTTTGTCTTG 3'

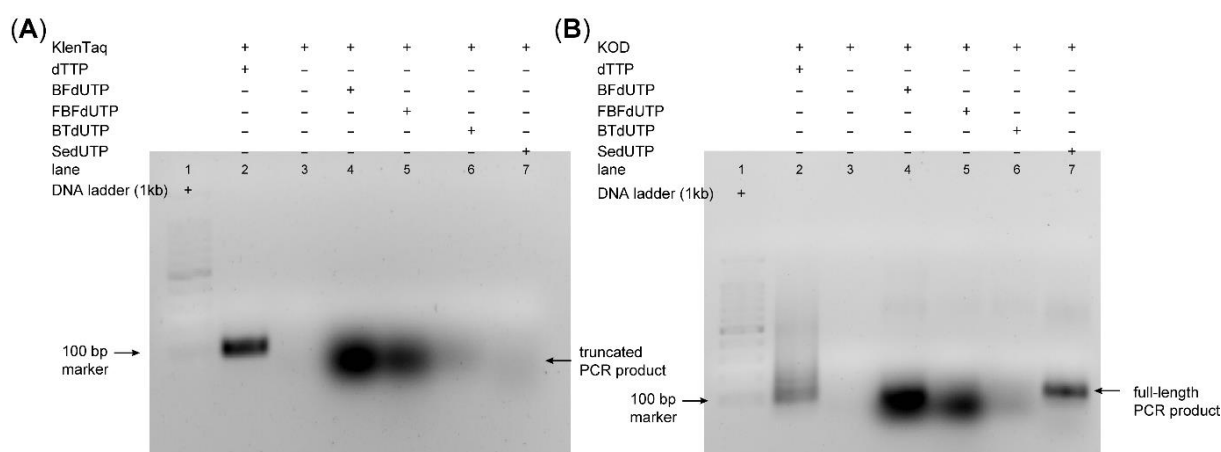


Figure 20. 2% agarose gel electrophoresis (AGE) image of the PCR products (EtBr stained). PCR with (A) KlenTaq and (b) KOD DNA polymerase. Lane 1 in Figure (A, B) is a ladder of 1Kb.

3.3 Conclusion

Employing biochemical and structural analyses utilizing a responsive fluorescent nucleotide analog, we gathered valuable insights into the functional capabilities of A and B family DNA polymerases. Our studies involving different templates guiding the insertion of single and multiple nucleotide analogs at various sites, reveal distinct incorporation behaviour. The introduction of an electronegative fluorine atom into the benzofuran ring does not compromise the overall incorporation efficiency, demonstrating comparable effectiveness to BFdUTP. B family polymerases efficiently process the nucleotide analog, generating ONs with labels at different positions. In contrast, KlenTaq exhibits stringency in incorporating the nucleotide, particularly at consecutive positions in primer extension setups. In single nucleotide

incorporation scenarios, KlenTaq accurately adds only one modified nucleotide, while KOD misincorporates a second nucleotide, resulting in a double-modified product. Consequently, KlenTaq proves advantageous in the initial step for site-specific labeling, followed by primer extension with either KOD or KlenTaq after purification. After a purification step, the modified primer can be fully extended with either KOD or KlenTaq allowing site-specific incorporation of the nucleotide analog.

The FBFdU label emerges as an effective fluorescent reporter, offering real-time insights into the binary complex formation and enzyme activity under equilibrium conditions. Fluorescence studies provide crucial information on the binding affinity of KlenTaq and KOD to the 3'-end-labeled primer, highlighting differences in substrate promiscuity. The ^{19}F -label of the nucleoside yields distinct signatures facilitating direct comparison of the relative population of states adopted by the nucleotide and modified ON under equilibrium conditions. These findings extend the utility of our dual-app nucleotide analog as not only a fluorescent probe for monitoring polymerization events but also as a tool for NMR spectroscopy to observe catalytic processes. Further, examination of three ternary complex crystal structures of BFdUTP, FBFdUTP, and SedUTP reveals critical points about the incorporation mechanism. The flexibility of Arg660 and Arg587 around the active site emerges as a key factor controlling the incorporation of the modified heterocyclic moiety. Additionally, the stacking interaction in the pre-catalytic state enhances the substrate's efficacy. Furthermore, the narrow channel within the enzyme is well-suited for accommodating a single five-membered ring, mimicking natural amino acid conformations. However, the insertion of multiple bicyclic moieties disrupts the processivity of the enzyme. Taken together, this multi-disciplinary analysis provided valuable insights into the incorporation mechanism of nucleobase-modified nucleotide probes, which is expected to pave ways to design polymerase-compatible nucleotide probes and inhibitors.

3.4 Experimental sections

3.4.1 Materials

5-Iodo-2'-deoxy uridine was purchased from TCI chemicals. *Bis* (triphenylphosphine)-palladium (II) chloride, butyllithium and tributyltin chloride were purchased from Sigma Aldrich. 5'-FAM labeled DNA ON, primers and templates were purchased from Integrated DNA Technologies, Inc. DNA templates were purified by polyacrylamide gel electrophoresis (PAGE) under denaturing conditions, and desalted using Sep-Pak Classic C18 cartridges (Waters Corporation). dNTPs were purchased from Thermo Fischer Scientific.

3.4.2. Instrumentation

NMR spectra were recorded on a 400 MHz Jeol ECS-400 and Bruker Ultrashield 400 MHz WB Plus spectrometers. ^{19}F NMR experiments were performed Bruker AVANCE III HD ASCEND600MHz spectrometer equipped with a Cryo-Probe (CP2.1QCI600S3 H/F-C/N-D-05Z XT). Mass measurements were performed using a Water Synapt G2 High Definition mass spectrometer. HPLC analysis was done using Agilent Technologies 1260 Infinity HPLC. Reverse-phase flash chromatographic (C18 RediSepRf column) purifications of nucleotides were carried out using Teledyne ISCO Combi Flash Rf. ESI-MS analysis of DNA ONs was performed using a Water Synapt G-Si Mass spectrometry instrument in negative mode. DNA and nucleotides were quantified by measuring absorbance in Shimadzu UV-2600 or NanoDropTM 2000c spectrophotometer. Fluorescent DNA ONs were resolved by PAGE on an OWL S4S sequencing gel electrophoresis instrument and were imaged using Typhoon FLA 9500 Variable mode imager. Steady-state fluorescence spectra were recorded using a micro-fluorescence cuvette (Hellma, path length 1.0 cm) on a Horiba Fluoromax-4 spectrophotometer.

3.4.3 Synthesis of modified nucleotide analog 3

5-fluorobenzofuran-modified-2'-deoxyuridine **2** was synthesized using a previously reported procedure. The compound was characterized by NMR and HRMS analysis and matched the previously reported data.⁴⁴

FBFdUTP 3. The nucleotide analog was synthesized as mentioned in Chapter 2 (section 2.4.3) using FBFdU **2** (31 mg, 17 % yield). ^1H -NMR (400 MHz, D_2O): δ (ppm) 8.25 (s, 1H), 7.62–7.59 (m, 1H), 7.31–7.28 (m, 1H), 7.23 (s, 1H), 7.07 (t, $J = 9.2$ Hz, 1H), 6.34 (t, $J = 7.2$ Hz, 1H), 4.68 (s, 1H), 4.25 (br, 3H), 2.50–2.37 (br, 2H); ^{13}C NMR (100 MHz, D_2O): δ (ppm) 181.4, 159.8, 158.2, 150.8, 149.8, 137.3, 129.7, 112.1, 111.9, 111.8, 106.3, 104.6, 85.6, 70.5, 58.9, 38.5; ^{31}P NMR (162 MHz, D_2O): δ (ppm) -6.46 (d, $J = 20.7$ Hz, P_α), -11.48 (d, $J = 19.9$ Hz, P_γ), -22.51 (t, $J = 20.4$ Hz, P_β); ^{19}F NMR (376.6 MHz, D_2O): δ (ppm) -121.62; HRMS: m/z calculated for $\text{C}_{17}\text{H}_{17}\text{FN}_2\text{O}_{15}\text{P}_3$ $[\text{M}-\text{H}]^- = 600.9831$, found $[\text{M}-\text{H}]^- = 600.9831$.

3.4.4 Solid-phase synthesis of FBFdU-modified oligonucleotides

FBFdU-Modified phosphoramidite was synthesized using a previously reported procedure.⁴⁴ 3'-end FBFdU-modified DNA ONs (**P3**, **P3***, **P2**, and **P2***) were synthesized using a K&A DNA/RNA synthesizer in 1 μmole scale (1000 Å Universal support III). At the end of the synthesis, cleavage of the ONs from the solid support was achieved using 2M ammonia in methanol at room temperature for 60 minutes, and afterward, these solutions were evaporated

to dryness using SpeedVAC. Next, the solid support was treated with 30% aqueous ammonium hydroxide solution and heated at 55°C for 23 h. Samples were further cooled to RT and centrifuged. The supernatant was evaporated, and the ON was purified by 20% denaturing preparative PAGE. Under UV light, the corresponding gel bands for the ONs were identified and cut from the Gel. Further gel bands were crushed into small pieces by a glass rod in a poly-prep column (Bio-Rad). ONs extracted in aqueous ammonium acetate (0.5 M, 3 mL) by shaking it for 12 h. Extra salt from the oligos was removed by using Sep-PakC-18 cartridges. The purity and integrity of the oligos were checked by RP-HPLC and ESI-mass analysis, respectively.

3.4.5 Reaction using KlenTaq and KOD DNA polymerases

To form primer-template duplexes, a 5'-FAM-labeled primer **P1** (50 μ M) was annealed with templates **T1–T5** (55 μ M) in a reaction buffer (50 mM Tris–HCl, 16 mM $(\text{NH}_4)_2\text{SO}_4$, 2.5 mM MgCl_2 , 0.1% (v/v) Tween 20, pH 8.0) by heating at 95 °C for 3 minutes. Subsequently, the samples were gradually cooled to room temperature and further chilled in an ice bath for 30 minutes. All primer extension reactions followed procedures similar to those outlined in Chapter 2 (section 2.4.4). The concentrations of each component are provided in the accompanying table.

Components	Amount (μ L)	Stock conc.		Resultant Conc.	
Reaction buffer	4	5.0	x	1.0	x
dNTP-dUTP mix	3	1	mM	150	μ M
FBFdUTP 3/dTTP	3	1	mM	150	μ M
P1•(T1–T5)	0.6	50	μ M	1.50	μ M
KlenTaq/KOD	1.2	50	nM	3	nM
Water	8.2				
Total volume	20				

A 5 μ L aliquot (0.94 pmol) of the sample was applied to an analytical 18% denaturing polyacrylamide gel and subjected to electrophoresis at 30 W for a duration of 2.5 hours. Gel visualization was carried out utilizing a Typhoon gel scanner with excitation at the FAM wavelength. Control reactions were performed to investigate misincorporation involving the primer-template duplex, enzyme, dATP, dGTP, and dCTP, excluding dTTP or modified dUTP.

3.4.6 Single nucleotide incorporation assay

A 5'-FAM-labeled primer, **P2** (50 μ M), and template **T6** (55 μ M) were hybridized using a similar procedure as mentioned previously. The single nucleotide incorporation assay was carried out using the same procedure as mentioned in earlier chapter 2 (section 2.4.5). The concentration of each component is mentioned in the table below.

Components	Amount (μL)	Stock conc.		Resultant Conc.	
Reaction buffer	4	5.0	x	1.0	x
FBFdUTP 3/dTTP	3	1	mM	150	μM
P2•T6	0.6	50	uM	1.50	μM
KlenTaq/KOD	1.2	50	nM	3	nM
Water	11.2				
Total volume	20				

Aliquots of the reaction (10 μL) at 15 and 45 minutes were mixed with 35 μL of denaturing loading buffer (8.3 M urea in 10 mM Tris–HCl, 100 mM EDTA, 0.05% bromophenol blue, pH 8) and heat denatured at 95 °C for 5 minutes followed by cooling on an ice bath. Subsequently, 5 μL (1.6 pmol) of the sample was loaded onto an analytical 20% denaturing polyacrylamide gel, electrophoresed at 30 W for 2.5 hours, and visualized using a Typhoon gel scanner at a FAM wavelength.

3.4.7 Large-scale synthesis of single FBFdU-modified ONs using KLenTaq

Large scale Single nucleotide incorporation reaction was conducted in the presence of 150 μM FBFdUTP 3, 1.50 μM pre-annealed primer•template duplex (P2•T6), and 3 nM of klenTaq in nM KlenTaq in reaction buffer (pH 8.0, 16 mM (NH₄)₂SO₄, 2.5 mM MgCl₂, 0.1% (v/v) Tween 20) in the total volume of 840 μL and purified using the same protocol mentioned in Chapter 2 (section 2.4.5). Approximately 60% of the pure product **P2*** was isolated and 500 pmol of DNA was dissolved in LCMS buffer (10 mM triethylamine and 100 mM hexafluoro-2-propanol in autoclaved water) containing 50% ACN and subjected to ESI mass analysis in negative mode.

3.4.8 Two-step site-specific labeling

All the reaction steps and the collection of the DNA pellets were performed under similar conditions as we mentioned in previous Chapter 2 (section 2.4.6). Four reactions were conducted, and the collective DNA pellets were dissolved in 30 μL of denaturing loading buffer (8.3 M urea in 10 mM Tris–HCl, 100 mM EDTA, 0.05% bromophenol blue, pH 8). Subsequently, the mixture was loaded onto a preparative 15% polyacrylamide gel and subjected to electrophoresis under denaturing conditions. The gel was UV-shadowed and the band corresponding to the full-length product was excised. The resulting ON (oligonucleotide) product was then extracted from the gel using 0.5 M ammonium acetate and desalting was carried out using a Sep-Pak C18 cartridge.

3.4.9 Fluorescence binding assay to determine binary complex formation.

FBFdU-modified- primer•template duplexes (**P3*•T7**, 10 μ M) were prepared in a binding buffer (10 mM Tris-HCl, 50 mM KCl, 5 mM MgCl₂, pH 7.9) by taking 1:1.05 ratio of primer **P3*** and template **T7**, respectively, using previously mentioned procedure. 200 μ L volume of samples containing Primer•template duplexes (**P3*•T7**, 50 nM) with increasing concentrations of KlenTaq/KOD (0–5 μ M) in the binding buffer were incubated at 55 °C for 20 minutes. Fluorescence was recorded by exciting the samples at 330 nm with excitation and emission slit widths of 7 nm and 8 nm, respectively at 25°C temperature. Overall glycerol concentrations were maintained at 0 –1.5% for all the samples. The fluorescence experiments were conducted in triplicate. The increase in fluorescence intensity at the emission maximum (430 nm) was fitted to the Hill equation (Origin 8.5) by plotting normalized fluorescence intensity (F_N) against KlenTaq/KOD concentration. A Hill coefficient (n) close to 1 was observed for the binding of the FBFdU-labeled primer•template duplex (**P3*•T7**) to KlenTq/KOD DNA polymerase, with R^2 approaching unity.

$$F_N = \frac{F_i - F_0}{F_S - F_0}$$

F_i = fluorescence intensity measured corresponding to each protein concentration. F_0 and F_S represent the fluorescence intensity measured in the absence of protein and at the saturation point, respectively. n =Hill coefficient (the degree of cooperativity in binding). K_d is the apparent dissociation constant.

$$F_N = F_0 + (F_S - F_0) \left(\frac{[\text{Polymerase}]^n}{[K_d]^n + [\text{Polymerase}]^n} \right)$$

3.4.10 Monitoring the enzymatic incorporation of FBFdUTP 3 by fluorescence assay by KlenTaq and KOD DNA polymerase

For fluorescence assay

Firstly, unmodified primer•template duplexes (**P3•T7**, 200 nM) were incubated with increasing concentrations of KlenTaq/KOD (0–50 nM) in the binding buffer (total volume of 200 μ L) at 55 °C for 20 minutes. FBFdUTP 3 (final concentration 100 nM) was added to each sample and incubated at 55 °C for 20 minutes. The glycerol concentrations were below 0.05% for all the samples. Subsequently, samples were cooled to 25 °C, and fluorescence was measured at that temperature. As a control experiment, the **P5•T7** duplex (200 nM) was incubated with KlenTaq/KOD (50 nM) at 55 °C for 20 minutes, and FBFdUTP (final concentration of 100 nM) was added and incubated at 55 °C for 20 minutes. Another control experiment used the same procedure with a sample containing only KlenTaq/KOD (50 nM) or the primer-template

duplex. All samples were excited at 330 nm with excitation and emission slit widths of 7 and 8 nm, respectively. Measurements were carried out in triplicate.

For PAGE analysis

Using a similar procedure to fluorescence assay, all the reactions were performed in the presence of 5'-FAM labeled primer-template duplex (**P2•T7**, 200 nM), FBFdUTP **3** (100 nM), increasing concentration of KlenTaq/KOD (ranging from 0 to 50 nM) in binding buffer (total volume of 20 μ L) at 55 °C for 20 minutes. After completion of the reaction, 10 μ L of the denaturing loading buffer was added and heated at 95 °C for 5 minutes and cooled to RT. 5 μ L of the reaction mixtures were loaded into a 20% denaturing PAGE and ran for 3 h at 35 W. Gel image was taken using Typhoon Gel scanner at FAM wavelength.

3.4.11 ^{19}F NMR spectroscopy assay to determine the binary complex formation and incorporation by KlenTaq

For binding assay

FBFdU-modified-primer (**P3***) (100 μ M) was annealed with the complementary ON (**T7**) (100.5 μ M) in 1X KlenTaq binding buffer to form the DNA duplex (**P3*•T7**) by heating at 95°C for 3 minutes and then slowly cooled down to RT. The primer only (**P3***) (10 μ M), primer-template duplex (**P3*•T7**) (10 μ M) and primer-template duplex with KlenTaq (5,10 μ M) were incubated at 55°C for 20 minutes in 1X KlenTaq binding buffer (containing 10% glycerol and 20% D_2O), cooled to room temperature to perform the NMR experiments. Samples were transferred into a Shigemi tube (5 mm advance NMR microtube) and ^{19}F NMR data were recorded on a Bruker AVANCE III HD ASCEND600MHz spectrometer equipped with a Cryo-Probe (CP2.1QCI600S3 H/F-C/N-D-05Z XT). To check the stability of the binary complex, the sample containing duplex (**P3*•T7**) (10 μ M) and KlenTaq (10 μ M) in binding buffer was first incubated at 55°C for 20 minutes and followed by heated at 95°C for 20 minutes. The sample was cooled to room temperature and performed the experiment.

For incorporation assay

Unmodified primer (**P3**) (100 μ M) was heated with the complementary ON (**T7**) (100.5 μ M) in 1X KlenTaq binding buffer at 95°C for 3 minutes, then slowly cooled down to RT to form the primer-template duplex (**P3•T7**). Samples containing FBFdUTP **3** (10 μ M), or FBFdUTP (10 μ M) with primer-template duplex (**P3•T7**) (10 μ M) and KlenTaq (10 μ M) were heated at 55°C for 20 minutes in binding buffer (contains 10% glycerol and 20% D_2O) and cooled to RT. Further samples were subjected to an NMR experiment.

3.4.12 EMSA to determine binary complex formation

For Fluorescence

A 20 μL solution containing primer-template duplex samples (**P2*•T7/P2•T7**, 50 nM) with varying concentrations of KlenTaq/KOD (ranging from 0 to 2 μM) in 1X binding buffer (glycerol concentration was 0–1%), was incubated at 55 °C for 20 minutes. Following this, 10 μL of native loading buffer (containing 50% sucrose, 10 mM Tris-HCl, 100 mM EDTA, 0.05% bromophenol blue, and adjusted to pH 8) was added to each sample, and 10 μL of each mixture was loaded onto a native 12% polyacrylamide gel. Electrophoresis was performed in EDTA-free 1x TBE buffer supplemented with 5 mM MgCl_2 at 4 °C. The gel was visualized using a Typhoon gel scanner with excitation at the FAM wavelength. This binding assay was conducted in duplicate. The DNA duplex and the enzyme-binary complex exhibited distinct mobilities on the gel. The intensity of the DNA bands at various KlenTaq/KOD concentrations was quantified using Fiji Image analysis software. A plot correlating the fraction of free DNA to the KlenTaq/KOD concentration was fitted to the Hill equation using Origin 8.5 to determine the dissociation constant (K_d). The Hill coefficient (n) fell within the range of 1.3 to 1.4, and the coefficient of determination (R^2) was close to 1, indicating a strong correlation.

For the NMR experiment

A 100 μM solution of either 5'-FAM labeled primer **P2** or 5'-FAM, 3'-BFdU-labeled primer **P2*** was subjected to annealing with a 100.5 μM template **T7** in binding buffer was annealed as mentioned in the previous experiment to form the primer-template duplex(**P2*•T7/P2•T7**). Subsequently, samples containing primer-template duplex samples (**P2*•T7/P2•T7**, 10 μM) with varying concentrations of KlenTaq/KOD (5, 10 μM) were prepared in binding buffer (total volume 20 μL) and incubated at 55 °C for 20 minutes. After incubation, 10 μL of native loading buffer (containing 50% sucrose, 10 mM Tris-HCl, 100 mM EDTA, 0.05% bromophenol blue, and adjusted to pH 8) was added to each sample, and 5 μL the mixtures (33.3 pmol) was loaded into the 12% native PAGE and electrophoresed using the previous procedure. Gel images were taken at FAM wavelength using a Typhoon Gel scanner.

3.4.13 Competitive displacement assay

By fluorescence

P3*•T7 duplex was prepared as mentioned above. A volume of 200 μL samples containing a binary complex of **P3*•T7** (50 nM) and KlenTaq/KOD (1 μM) with varying concentration unlabeled primer-template duplex (**P3•T7**, 0–1 μM) in the binding buffer was kept at 55 °C for 15 minutes and subjected to a steady-state fluorescence experiment. Fluorescence was

measured by exciting at 330 nm with slit widths of 7 nm and 8 nm for excitation and emission respectively.

By EMSA

P2*•T7 and **P2•T7** primer-template duplexes were made using the above-mentioned procedure. A sequence of binary complex samples comprising **P2*•T7** duplex (50 nM) and either KlenTaq (1 μ M) or KOD (1 μ M) was subjected to titration with increasing concentrations of unlabeled primer-template duplex **P3•T7** (0 –1 μ M) in a binding buffer (10 mM Tris-HCl, 50 mM KCl, 5 mM MgCl₂, pH 7.9) of a final volume of 20 μ L. Following a 20-minute incubation at 55 °C, 10 μ L of native loading buffer was added. The samples were then separated by loading onto a 12% native polyacrylamide gel (PAGE) and subjected to electrophoresis for 2.5 hours in 1X EDTA-free TBE buffer supplemented with 5 mM MgCl₂ at 4 °C. Gel was imaged by a Typhoon Gel scanner with excitation at the FAM wavelength. A parallel experiment was conducted using the primer-template duplex **P2•T7**. Binary complexes were formed with **P2•T7** duplex (50 nM) and either KlenTaq (1 μ M) or KOD (1 μ M). The binding buffer was supplemented with escalating concentrations of the FBFdU-labeled primer-template duplex **P3*•T7** (ranging from 0 to 1 μ M). EMSA was performed following the same procedure as described above.

3.4.14 Competition reaction between modified nucleotides and dTTP

P6•T8 duplex was prepared in 1X reaction buffer as the previously mentioned procedure. In each 20 μ L reaction, 1.5 μ M of **P6•T8** duplex 150 μ M of dNTP mixture containing dU*TP (BFdUTP, FBFdUTP, SedUTP)/dTTP in 1X reaction buffer and 3 nM of KlenTaq were added. The dTTP to dU*TP ratio varied from 0:1 to 1:0 (0/1, 1/50, 1/10, 1/5, 1/2, 1/1, 2/1, 5/1, 10/1, 50/1, and 1/0). Reactions were performed at 55 °C for 40 minutes and followed by 130 μ L of denaturing loading buffer (8.3 M urea in 10 mM Tris-HCl, 100 mM EDTA, 0.05% bromophenol blue, pH 8) was added and heat denatured at 95 °C for 5 minutes. 5 μ L (1 pmol) of the samples were loaded onto a 20% denaturing PAGE and electrophoresed in 1X TBE buffer for 4 hours at 30W. Gel images at FAM emission wavelength were captured using a Typhoon gel scanner. Relative band intensity was determined using Fiji image analysis software. All experiments were performed in triplicate to ascertain standard deviation and average incorporation efficiency values.

3.4.15 Crystallization and structure determination.

The experimental setup for obtaining the ternary complex of KlenTaq with primer-template duplex and modified dNTPs was performed as mentioned in Chapter 2 (2.4.10). Using a molar ratio of 1:1.2:3:10, a mixture was prepared with KlenTaq, primer-template duplex, ddCTP, and

modified nucleotide (FBFdUTP/SedUTP) in the presence of 20 mM MgCl₂, resulting in final concentrations of 6.3 mg/mL, 0.12 mM, 0.31 mM, and 1.03 mM, respectively. The crystals were grown using the hanging drop vapor diffusion method by mixing with the reservoir solution in different ratios.

FBFdUTP.

Diffraction-quality crystals with FBFdUTP were obtained in a drop composed of a 1:2 ratio of protein to the reservoir. Reservoir composition was 15% (w/v) PEG 3K, 20% (v/v) 1,2,4-butanetriol, 1% (w/v) NDSB 256, 0.3 M Li₂SO₄, 0.3M Na₂SO₄, and 0.3 M K₂SO₄.

SedUTP.

For crystals with SedUTP, the reservoir composition was 1M di-sodium tartrate, 50 mM Tris, pH 7.5, and 30 mM MgCl₂.

Data collection followed the protocol outlined in Chapter 2. Figures were generated using PyMOL, and for the alignment of structures in figures and calculation of RMSD values (after excluding outliers), the "align" command in PyMOL was utilized.

3.4.16 PCR reaction using KlenTaq and KOD DNA polymerase

The PCR mixture (50 µL) contained KOD/KlenTaq DNA polymerase (50 nM, 100 nM, 200 nM), natural dNTPs (dATP, dGTP, and dCTP 200 µM), dU*TP (200 µM), primers **P_F**(Forward primer) and **P_R** (Reverse primer) (2 µM), a 98 -mer template **T_{PCR}**(template for PCR) in 1X Phusion buffer. After preheating at 95 °C for 3 min, 30 PCR cycles were run under the following conditions: denaturation for 1 min at 95 °C, annealing at 55 °C for 1 min, extension at 72 °C for 1.5 min, followed by final extension step 5 min at 75 °C. 2 µL of the PCR mixture was added with 3 µL of 6 X loading buffer and analyzed on 2 % agarose gel (EtBr stained) in 0.5 X TAE buffer. Images were taken using E-gel imager.

3.5 References:

1. Kranaster, R.; Marx, A. Engineered DNA polymerases in biotechnology. *ChemBioChem* **2010**, *11*, 2077–2084.
2. Chen, T.; Hongdilokkul, N.; Liu, Z.; Thirunavukarasu, D.; Romesberg, F. E. The expanding world of DNA and RNA. *Curr. Opin. Chem. Biol.* **2016**, *34*, 80–87.
3. Hocek, M. Enzymatic synthesis of base-functionalized nucleic acids for sensing, cross linking, and modulation of protein–DNA binding and transcription. *Acc. Chem. Res.* **2019**, *52*, 1730–1737.
4. Tomkuvienė, M.; Mickutė, M.; Vilkaitis, G.; Klimašauskas, S. Repurposing enzymatic transferase reactions for targeted labeling and analysis of DNA and RNA. *Curr. Opin. Biotechnol.* **2019**, *55*, 114–123.
5. Klöcker, N.; Weissenboeck, F. P.; Rentmeister, A. Covalent labeling of nucleic acids. *Chem. Soc. Rev.* **2020**, *49*, 8749–8773.

6. Espinasse, A.; Lembke, H. K.; Cao, A. A.; Carlson, E. E. Modified nucleoside triphosphates in bacterial research for in vitro and live-cell applications. *RSC Chem. Biol.* **2020**, *1*, 333–351.
7. McKenzie, L. K.; El-Khoury, R.; Thorpe, J. D.; Damha, M. J.; Hollenstein, M. Recent progress in non-native nucleic acid modifications. *Chem. Soc. Rev.* **2021**, *50*, 5126–5164.
8. Zhu, G.; Song, P.; Wu, J.; Luo, M.; Chen, Z.; Chen, T.; Application of Nucleic Acid Frameworks in the Construction of Nanostructures and Cascade Biocatalysts: Recent Progress and Perspective. *Front. Bioeng. Biotechnol.* **2022**, *9*, 792489.
9. Vyas, R.; Zahurancik, W. J.; Suo, Z. Structural basis for the binding and incorporation of nucleotide analogs with L-stereochemistry by human DNA polymerase λ . *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, E3033–E3042.
10. Bergen, K.; Steck, A. L.; Strütt, S.; Baccaro, A.; Welte, W.; Diederichs, K.; Marx, A. Structures of KlenTaq DNA polymerase caught while incorporating C5-modified pyrimidine and C7-modified 7-deazapurine nucleoside triphosphates. *J. Am. Chem. Soc.* **2012**, *134*, 11840–11843.
11. Obeid, S.; Bußkamp, H.; Welte, W.; Diederichs, K.; Marx, A. Interactions of non-polar and “Click-able” nucleotides in the confines of a DNA polymerase active site. *Chem. Commun.* **2012**, *48*, 8320–8322.
12. Obeid, S.; Baccaro, A.; Welte, W.; Diederichs, K.; Marx, A. Structural basis for the synthesis of nucleobase modified DNA by *Thermus aquaticus* DNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 21327–21331.
13. Kropp, H. M.; Dürr, S. L.; Peter, C.; Diederichs, K.; Marx, A. Snapshots of a modified nucleotide moving through the confines of a DNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115*, 9992–9997.
14. Manna, S.; Sarkar, D.; Srivatsan, S. G. A Dual-App Nucleoside Probe Provides Structural Insights into the Human Telomeric Overhang in Live Cells. *J. Am. Chem. Soc.* **2018**, *140*, 12622–12633.
15. Chen, H.; Viel, S.; Ziarelli, F.; Peng, L. ^{19}F NMR: a valuable tool for studying biological events. *Chem. Soc. Rev.* **2013**, *42*, 7971–7982.
16. Bao, H.-L.; Ishizuka, T.; Sakamoto, T.; Fujimoto, K.; Uechi, T.; Kenmochi, N.; Xu, Y. Characterization of human telomere RNA G-quadruplex structures in vitro and in living cells using ^{19}F NMR spectroscopy. *Nucleic Acids Res.* **2017**, *45*, 5501–5511.
17. Ishizuka, T.; Zhao, P.-Y.; Bao, H.-L.; Xu, Y. Oligonucleotides DNA containing 8-trifluoromethyl-2'-deoxyguanosine for observing Z-DNA structure. *Analyst* **2017**, *142*, 4083–4088.
18. Pham, L. B. T.; Costantino, A.; Barbieri, L.; Calderone, V.; Luchinat, E.; Banci, L.; *J. Am. Chem. Soc.* **2023**, *145*, 1389–1399.
19. Himmelstoß, M.; Erharter, K.; Renard, E.; Ennifar, E.; Kreutz, C.; Micura, R. 2'-O-Trifluoromethylated RNA—a powerful modification for RNA chemistry and NMR spectroscopy. *Chem. Sci.* **2020**, *11*, 11322–11330.
20. Li, Q.; Chen, J.; Trajkovski, M.; Zhou, Y.; Fan, C.; Lu, K.; Tang, P.; Su, X.; Plavec, J.; Xi, Z.; Zhou, C. 4'-Fluorinated RNA: synthesis, structure, and applications as a sensitive ^{19}F NMR probe of RNA structure and function. *J. Am. Chem. Soc.* **2020**, *142*, 4739–4748.
21. Bao, H.-L.; Ishizuka, T.; Sakamoto, T.; Fujimoto, K.; Uechi, T.; Kenmochi, N.; Xu, Y. Characterization of human telomere RNA Gquadruplex structures in vitro and in living cells using ^{19}F NMR spectroscopy. *Nucleic Acids Res.* **2017**, *45*, 5501–5511.

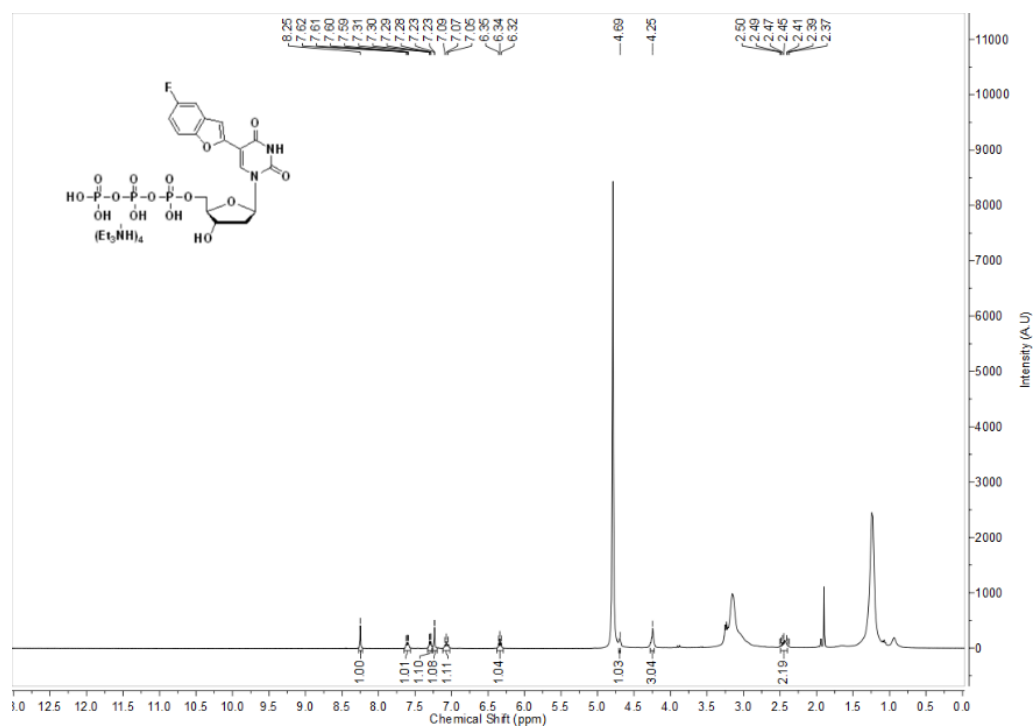
22. Bao, H.-L.; Liu, H.-S.; Xu, Y. Hybrid-type and two-tetrad antiparallel telomere DNA G-quadruplex structures in living human cells. *Nucleic Acids Res.* **2019**, *47*, 4940–4947.
23. Baranowski, M. R.; Warminski, M.; Jemielity, J.; Kowalska, J. 5'-fluoro(di)phosphate-labeled oligonucleotides are versatile molecular probes for studying nucleic acid secondary structure and interactions by ^{19}F NMR. *Nucleic Acids Res.* **2020**, *48*, 8209–8224.
24. Riedl, J.; Pohl, R.; Rulísek, L.; Hocek, M. Synthesis and photophysical properties of biaryl-substituted nucleos(t)ides. Polymerase synthesis of DNA probes bearing solvatochromic and pH-sensitive dual fluorescent and ^{19}F NMR labels. *J. Org. Chem.* **2012**, *77*, 1026–1044.
25. Khatik, S. Y.; Srivatsan, S. G. Environment-Sensitive Nucleoside Probe Unravels the Complex Structural Dynamics of i-Motif DNAs. *Bioconjugate Chem.* **2022**, *33*, 1515–1526.
26. Wee, W. A.; Yum, J. H.; Hirashima, S.; Sugiyama, H.; Park, S. Synthesis and application of a ^{19}F -labeled fluorescent nucleoside as a dual-mode probe for i-motif DNAs. *RSC Chem. Biol.* **2021**, *2*, 876–882.
27. Saady, A.; Wojtyniak, M.; Varon, E.; Böttner, V.; Kinor, N.; Shav-Tal, Y.; Ducho, C.; Fischer, B. Specific, sensitive, and quantitative detection of HER 2 mRNA breast cancer marker by fluorescent light-up hybridization probes. *Bioconjugate Chem.* **2020**, *31*, 1188–1198.
28. Fauster, K.; Kreutz, C.; Micura, R. 2'-SCF₃ Uridine—A Powerful Label for Probing Structure and Function of RNA by ^{19}F NMR Spectroscopy. *Angew. Chem. Int. Ed.* **2012**, *51*, 13080–13084.
29. Olszewska, A.; Pohl, R.; Hocek, M. Trifluoroacetophenone-Linked Nucleotides and DNA for Studying of DNA–Protein Interactions by ^{19}F NMR Spectroscopy. *J. Org. Chem.* **2017**, *82*, 11431–11439.
30. Bao, H.-L.; Ishizuka, T.; Iwanami, A.; Oyoshi, T.; Xu, Y. A Simple and Sensitive ^{19}F NMR Approach for Studying the Interaction of RNA G-Quadruplex with Ligand Molecule and Protein. *ChemistrySelect* **2017**, *2*, 4170–4175.
31. Dow, B. J.; Malik, S. S.; Drohat, A. C. Defining the Role of Nucleotide Flipping in Enzyme Specificity Using ^{19}F NMR. *J. Am. Chem. Soc.* **2019**, *141*, 4952–4962.
32. Baranowski, M. R.; Warminski, M.; Jemielity, J.; Kowalska, J. 5'-fluoro(di)phosphate-labeled oligonucleotides are versatile molecular probes for studying nucleic acid secondary structure and interactions by ^{19}F NMR. *Nucleic Acids Res.* **2020**, *48*, 8209–8224.
33. Holzberger, B.; Rubini, M.; Möller, H. M.; Marx, A. A Highly Active DNA Polymerase with a Fluorous Core. *Angew. Chem. Int. Ed.* **2010**, *49*, 1324–1327.
34. Ghadessy, F. J.; Ramsay, N.; Boudsocq, F.; Loakes, D.; Brown, A.; Iwai, S.; Vaisman, A.; Woodgate, R.; Holliger, P. Generic expansion of the substrate spectrum of a DNA polymerase by directed evolution. *Nat. Biotechnol.* **2004**, *22*, 755–759.
35. Kuwahara, M.; Nagashima, J.; Hasegawa, M.; Tamura, T.; Kitagata, R.; Hanawa, K.; Hososhima, S.; Kasamatsu, T.; Ozaki, H.; Sawai, H. Systematic characterization of 2'-deoxynucleoside-5'-triphosphate analogs as substrates for DNA polymerases by polymerase chain reaction and kinetic studies on enzymatic production of modified DNA. *Nucleic Acids Res.* **2006**, *34*, 5383–5394.

36. Gierlich, J.; Gutmiedl, K.; Gramlich, P. M. E.; Schmidt, A.; Burley, G. A.; Carell, T. Synthesis of highly modified DNA by a combination of PCR with alkyne-bearing triphosphates and click chemistry. *Chem. Eur. J.* **2007**, *13*, 9486–9494.
37. Brudno, Y.; Liu, D. R. Recent progress toward the templated synthesis and directed evolution of sequence-defined synthetic polymers. *Chem. Biol.* **2009**, *16*, 265–276.
38. Cherkasov, D.; Biet, T.; Bäuml, E.; Traut, W.; Lohoff, M. New nucleotide analogues with enhanced signal properties. *Bioconjugate Chem.* **2010**, *21*, 122–129.
39. Hollenstein, M. Synthesis of deoxynucleoside triphosphates that include proline, urea, or sulfonamide groups and their polymerase incorporation into DNA. *Chem. Eur. J.* **2012**, *18*, 13320–13330.
40. Goubet, A.; Chardon, A.; Kumar, P.; Sharma, P. K.; Veedu, R. N. Synthesis of DNA oligonucleotides containing C5-ethynylbenzenesulfonamidemodified nucleotides (EBNA) by polymerases towards the construction of base functionalized nucleic acids. *Bioorganic Med. Chem. Lett.* **2013**, *23*, 761–763.
41. Dutson, C.; Allen, E.; Thompson, M. J.; Hedley, J. H.; Murton, H. E.; Williams, D. M. Synthesis of polyanionic C5-modified 2'-deoxyuridine and 2'-deoxycytidine-5'-triphosphates and their properties as substrates for DNA polymerases. *Molecules* **2021**, *26*, 2250.
42. Nuthanakanti, A.; Ahmed, I.; Khatik, S. Y.; Saikrishnan, K.; Srivatsan S. G. Probing G-quadruplex topologies and recognition concurrently in real time and 3D using a dual-app nucleoside probe. *Nucleic Acids Res.* **2019**, *47*, 6059–6072.
43. Nuthanakanti, A.; Boerneke, M. A.; Hermann, T.; Srivatsan S. G. Structure of the ribosomal decoding site RNA containing a selenium-modified responsive fluorescent ribonucleoside probe. *Angew. Chem. Int. Ed.* **2017**, *56*, 2640–2644.
44. Tanpure, A. A.; Srivatsan, S. G. Conformation-sensitive nucleoside analogues as topology-specific fluorescence turn-on probes for DNA and RNA G-quadruplexes. *Nucleic Acids Res.* **2015**, *43*, e149.
45. Sinkeldam, R. W.; Wheat, A. J.; Boyaci, H.; Tor, Y. Emissive nucleosides as molecular rotors. *ChemPhysChem* **2011**, *12*, 567–570.
46. Noé, M. S.; Sinkeldam, R. W.; Tor, Y. Oligodeoxynucleotides containing multiple thiophene-modified isomorphous fluorescent nucleosides. *J. Org. Chem.* **2013**, *78*, 8123–8128.
47. Wee, W. A.; Yum, J. H.; Hirashima, S.; Sugiyama, H.; Park, S.; Synthesis and application of a 19F-labeled fluorescent nucleoside as a dual-mode probe for i-motif DNAs. *RSC Chem. Biol.* **2021**, *2*, 876–882.
48. Guo, F.; Li, Q.; Zhou, C. Synthesis and biological applications of fluoro-modified nucleic acids. *Org. Biomol. Chem.* **2017**, *15*, 9552–9565.
49. Haidekker, M. A.; Theodorakis, E. A. Molecular rotors—fluorescent biosensors for viscosity and flow. *Org. Biomol. Chem.* **2007**, *5*, 1669–1678.
50. Förster, T.; Hoffmann, G. Die viskositätsabhängigkeit der fluoreszenz quantenausbeuten einiger farbstoffsysteme. *Z. Phys. Chem.* **1971**, *75*, 63–76.
51. Kropp, H. M.; Betz, K.; Wirth, J.; Diederichs, K.; Marx, A. Crystal structures of ternary complexes of archaeal B-family DNA polymerases. *PLoS ONE*. **2017**, *12*, e0188005.
52. Betz, K.; Malyshev, D.A.; Lavergne, T.; Welte, W.; Diederichs, K.; Dwyer, T.J.; Ordoukhanian, P.; Romesberg, F.E.; Marx, A. KlenTaq polymerase replicates unnatural base pairs by inducing a Watson-Crick geometry. *Nat Chem Biol.* **2012**, *8*, 612–614.

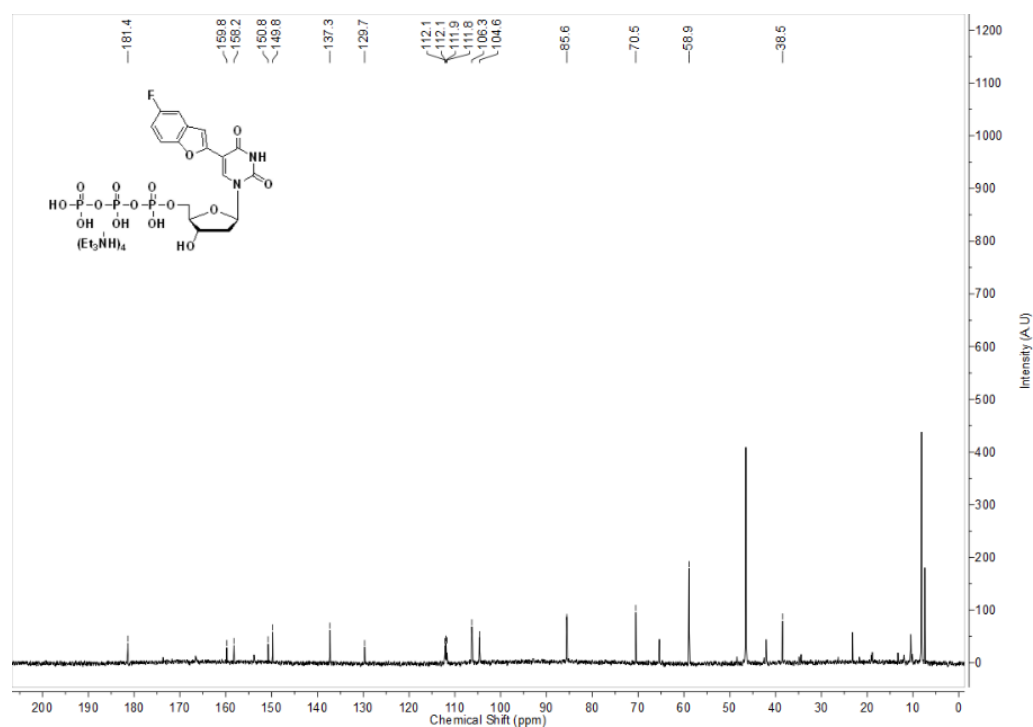
53. Li, Y.; Kong, Y.; Korolev, S.; Waksman, G. Crystal structures of the Klenow fragment of *Thermus aquaticus* DNA polymerase I complexed with deoxyribonucleoside triphosphates. *Protein Sci.* **1998**, *7*, 1116–1123.
54. Li, Y.; Waksman, G. Crystal structures of a ddATP-, ddTTP-, ddCTP, and ddGTP-trapped ternary complex of KlenTaq1: insights into nucleotide incorporation and selectivity. *Protein Sci.* **2001**, *10*, 1225–1233.
55. Kabsch, W. XDS. *Acta Crystallogr. D.* **2010**, *66*, 125–132.
56. Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.; Kapral, G. J.; GrosseKunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilligere, T. C.; Zwart, P. H. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D.* **2010**, *66*, 213–221.
57. Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D.* **2004**, *60*, 2126–2132.
58. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and development of Coot. *Acta Crystallogr. D.* **2010**, *66*, 486–501.
59. Sawai, H.; Nagashima, J.; Kuwahara, M.; Kitagata, R.; Tamura, T.; Matsui, I. Differences in Substrate Specificity of C(5)-Substituted or C(5)-Unsubstituted Pyrimidine Nucleotides by DNA Polymerases from Thermophilic Bacteria, Archaea, and Phages. *Chem. Biodiversity* **2007**, *4*, 1979–1995.
60. Thum, O.; Jager, S.; Famulok, M. Functionalized DNA: A New Replicable Biopolymer. *Angew. Chem. Int. Ed.* **2001**, *40*, 3990–3993.
61. Jager, S.; Famulok, M. Generation and Enzymatic Amplification of High-Density Functionalized DNA Double Strands. *Angew. Chem. Int. Ed.* **2004**, *43*, 3337–3340.
62. Gierlich, J.; Gutmiedl, K.; Gramlich, P. M. E.; Schmidt, A.; Burley, G. A.; Carell, T. Synthesis of Highly Modified DNA by a Combination of PCR with Alkyne-Bearing Triphosphates and Click Chemistry. *Chem.-Eur. J.* **2007**, *13*, 9486–9494.
63. Sawai, H.; Ozaki-Nakamura, A.; Mine, M.; Ozaki, H. Synthesis of New Modified DNAs by Hyperthermophilic DNA Polymerase: Substrate and Template Specificity of Functionalized Thymidine Analogues Bearing an sp³-Hybridized Carbon at the C5 α -Position for Several DNA Polymerases. *Bioconjugate Chem.* **2002**, *13*, 309–316.
64. Mehedi Masud, M.; Ozaki-Nakamura, A.; Kuwahara, M.; Ozaki, H.; Sawai, H. Modified DNA Bearing 5(Methoxycarbonylmethyl)-2' deoxyuridine: Preparation by PCR with Thermophilic DNA Polymerase and Postsynthetic Derivatization. *ChemBioChem* **2003**, *4*, 584–588.

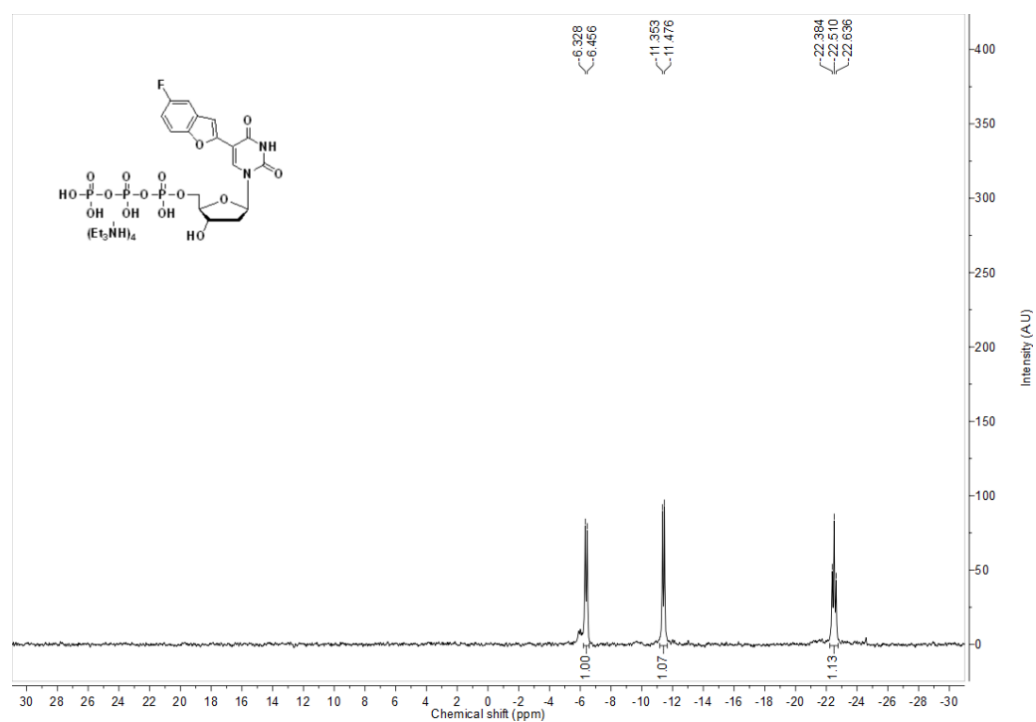
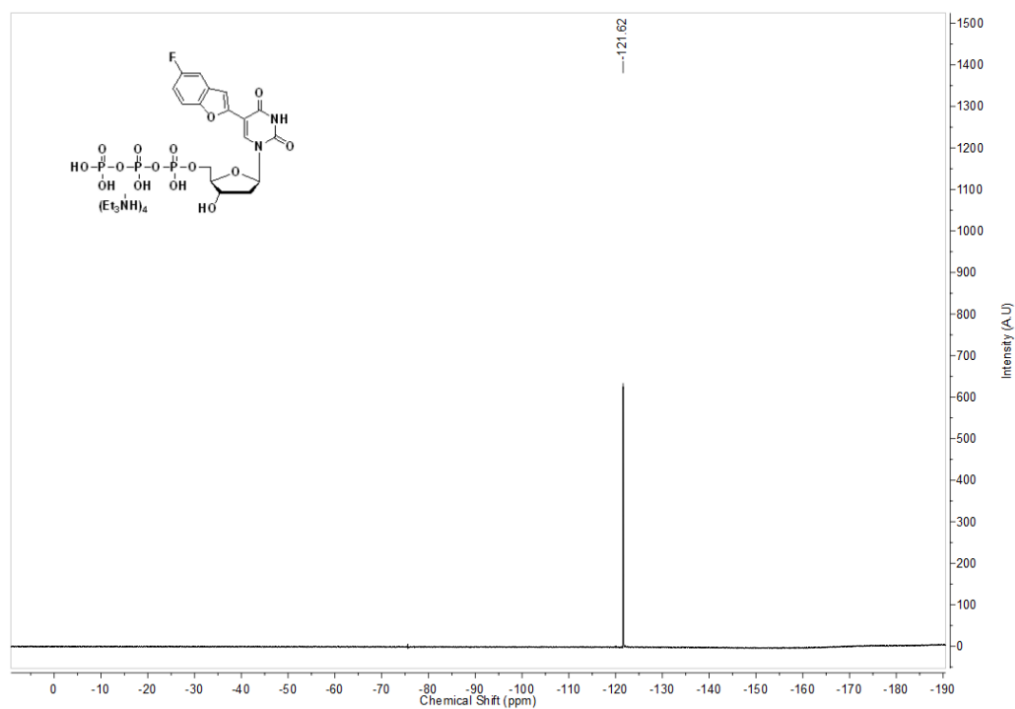
3.6 Appendix-II NMR and Mass spectra of synthesized compounds

¹H-NMR of FBFdUTP **3** (400 MHz, D₂O), a trace amount of triethylammonium acetate buffer is present.

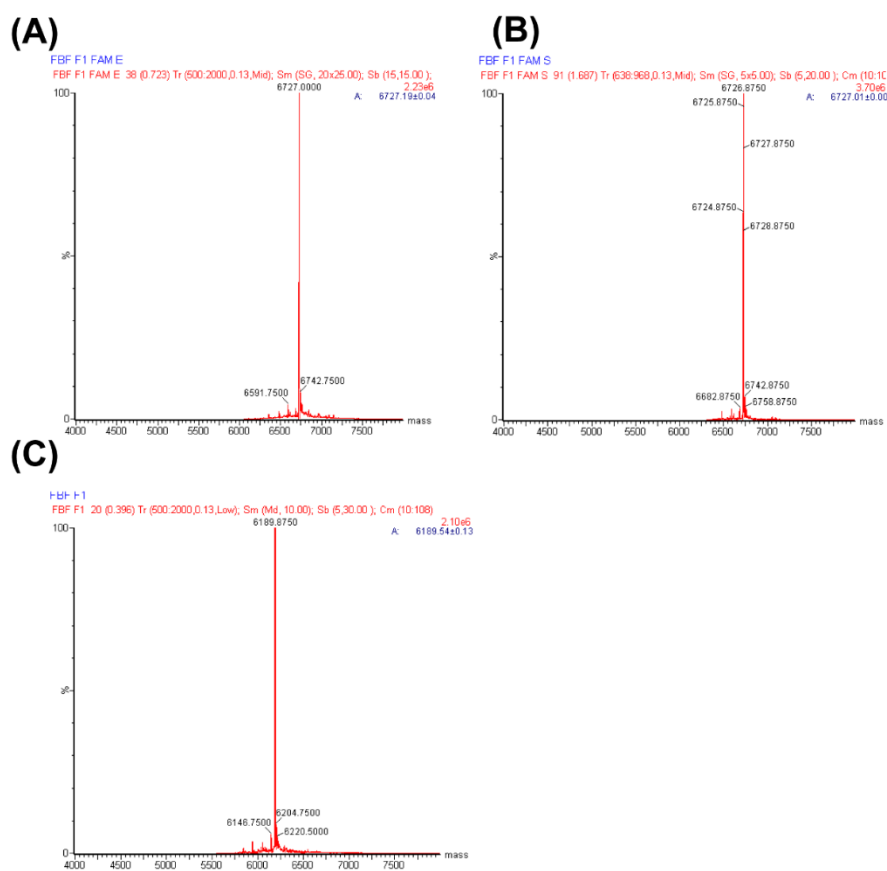
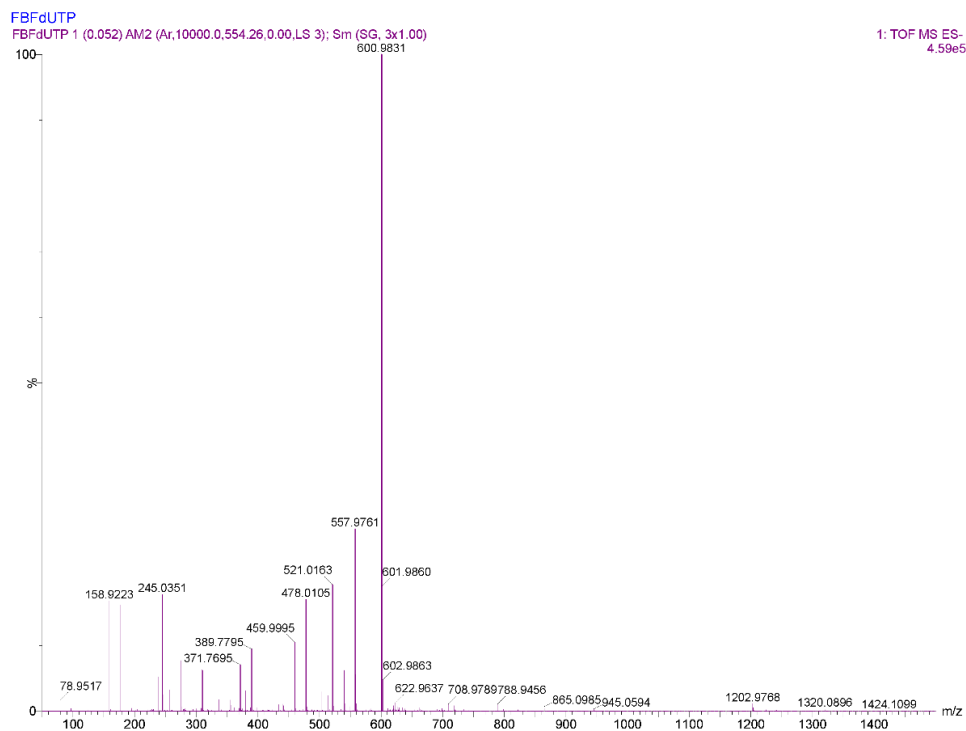


¹³C-NMR of FBFdUTP 3 (100 MHz, D₂O), a trace amount of triethylammonium acetate buffer is present.



^{31}P -NMR of FBFdUTP **3** (162 MHz, D_2O) ^{19}F -NMR of FBFdUTP **3** (376.6 MHz, D_2O)

HRMS spectrum of FBFdUTP **3** in negative mode. Calculated for $C_{17}H_{17}FN_2O_{15}P_3$ $[M-H]^- = 600.9831$, found $[M-H]^- = 600.9831$



ESI-MS spectra for FBFdU-labeled DNA ON products (A, B) **P2*** and (C) **P3***. See Table 3 for details. ESI-MS spectra of (D) site-specifically labeled ON **P2****.

Chapter 4

Microenvironment-responsive fluorescent nucleotide probes serve as good substrates to investigate the activity of terminal deoxynucleotidyl transferase (TdT)

4.1 Introduction

Unlikely template-dependent DNA polymerases, TdT (terminal deoxynucleotidyl transferase) owns a unique ability to insert nucleotides at the 3'-end of a ssDNA in a template-independent manner.¹⁻³ Due to this property and a wide range of substrate acceptability, it has been extensively utilized to synthesize functionalized oligonucleotides (ONs) for several biotechnological applications such as DNA nanotechnology⁴⁻⁶ and DNA-based data storage.⁷⁻¹⁰ TdT has been used in the enzymatic *de novo* synthesis of ONs,^{11,12} where it manifests a rising alternative method for solid phase synthesis of ONs. In recent decades, applications of TdT in molecular biology,¹³⁻¹⁶ biosensing,¹⁷⁻²² and aptamer development have emerged extensively.²³⁻²⁵

Apart from its biotechnological applications, TdT plays a pivotal role in developing vertebrates' adaptive immune system by promoting antibody repertoire diversity by catalyzing the random addition of nucleotides at the N regions of T cell receptors and immunoglobulin molecules during the V(D)J recombination process.²⁶⁻²⁸ Generally, expression of TdT occurs during the development of B cells, ceasing before the expression of Ig light chains.²⁷ Overexpression of TdT in some cancer types causes mutagenesis, where it competes with pol μ DNA polymerase and shows error-prone nonhomologous end joining (NHEJ) pathway during the repair of double-strand breaks (DSB).²⁸ Various clinical studies revealed that elevation of TdT expression is associated with several leukemia diseases (such as acute lymphocytic leukemia (ALL) and acute Myeloid leukemia (AML)). It has gained interest as a potential biomarker and an important therapeutic target for leukemia diseases.²⁹⁻³² Therefore, it is essential to understand the incorporation mechanism and effectively detect TdT activity to delve deeper into the possible pathological roles of TdT, its application in biosensing and for diagnosing TdT-related diseases. However, the traditional detection techniques to monitor the TdT activity rely on the immunoglobulin assay and gel electrophoresis analysis of the incorporation of radiolabeled nucleotide by TdT.³³⁻³⁷ These methods are time-consuming,

depend on the stability of the antibodies, are expensive, use hazardous radioisotopes and cannot provide real-time data for quantification of the TdT activity. Nevertheless, in recent years, nanomaterial-based fluorescent methods have been developed for this purpose. Some of these methods focused on emphasizing enzyme-based nanomaterial-enhanced signal amplification, the process entails creating a sequence-specific DNA, utilizing this DNA as a template for synthesizing nanomaterials with fluorescent or electrochemical properties and measuring the resulting signals.³⁸⁻⁴¹ On the other hand, some techniques involve the extended product formation by the TdT enzyme, which folds into a secondary structure and binds to the nanoparticles or a specific binder to collect a fluorescent signal. For example, G-quadruplexes can bind to thioflavin T and function as a peroxidase-analog⁴² or i-motifs that can form cyclometalated iridium (III) complexes, resulting in the amplified luminescent signal.⁴³ Another category involves nanosensors employing polymerization-directed exonuclease assistance in conjunction with quantum dots (QDs) for fluorescence resonance energy transfer (FRET).^{44,45} Despite their efficacy, these approaches exhibit significant drawbacks, including intricate chemical reactions required for nanomaterial synthesis, elevated expenses associated with labeled probes, and time-consuming procedures. Consequently, they fall short of fulfilling the requirements for easy analysis and clinical diagnosis. Hence, there is an urgent requirement for efficient methodologies that facilitate fast detection and quantification of TdT. Recently, fluorescent nucleobase probe systems have shown good promise for directly analyzing TdT activity.⁴⁶ The Velema group utilized a FRET-based probe, namely a size-expanded deoxyadenosine, which displays selective incorporation by TdT over other polymerases and phosphatase enzymes and enables the detection of TdT activity in cellular conditions.⁴⁶ However, this size-expanded nucleotide exhibits a low incorporation rate and a longer time for detection of the TdT activity. Here, our idea is to establish a simple and efficient platform to understand the mechanism of TdT and implement a real-time fluorescence-based detection technique to monitor the TdT activity by harnessing the inherent responsiveness of the environment-sensitive nucleotide analogs. In this context, we sought to monitor the incorporation of the C5 heterocyclic-modified environment-sensitive nucleotide analogs (Figure 1). These analogs demonstrate good substrate efficacy for enzymatic incorporation by various DNA, RNA polymerases, and PUP transferase.⁴⁷⁻⁵¹ Over the years, these responsive nucleotide analogs have been successfully employed as a biophysical toolbox to investigate the structural dynamics of several therapeutically relevant nucleic acid sequences and their interaction with small molecules and ligands *in vitro* and cellular conditions.⁵²⁻⁵⁴ Moreover, it has been utilized as an intelligent probe system to sense the microenvironment of the enzyme's

active site of template-dependent DNA polymerases, without compromising its processivity, enabling it to report the mechanistic events by monitoring changes in fluorescence signal (Chapter 2 and 3). Hence, incorporating these smart probes by TdT enzyme has several advantages: (i) systematic incorporation studies of these modified nucleotide analogs offer a comprehensive understanding of the mechanism of TdT enzyme and aid in the design of base-modified nucleotide analogs for various applications, (ii) the modified nucleotides themselves detect the binding process involved in the enzymatic incorporation and generate live incorporation data through fluorescence signals, facilitating the establishment of a screening platform to identify inhibitors of TdT activity.

Herein, we report the incorporation of a set of microenvironment-sensitive probes carrying different sizes of heterocyclic moiety at the C5 positions using TdT DNA polymerase (Figure 1). The high conservation of TdT extends across vertebrate phyla, ranging from cartilaginous fish to birds and humans.^{55–61} Also, the human-TdT structures share similarities of almost 80.6% to the murine-TdT.^{57,60} For our purpose, we have expressed and purified murine TdT (PDB ID: 4I27). The murine TdT demonstrates comparable efficiency in accepting these modifications, predominantly yielding double-modification products. Further, we have synthesized a benzofuran-modified dideoxy nucleotide analog and investigated the incorporation of this analog by different primer initiators. A competition experiment with natural dideoxy thymidine nucleotides was conducted to elucidate the key interactions responsible for the TdT reactions. Finally, we report the use of an environment-sensitive fluorescent nucleotide (BFdUTP) to monitor the binding process involved during the enzymatic incorporation process.

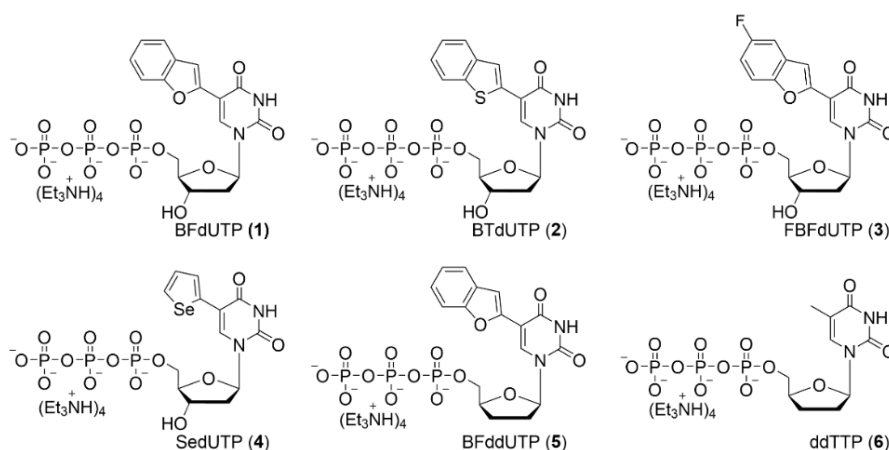
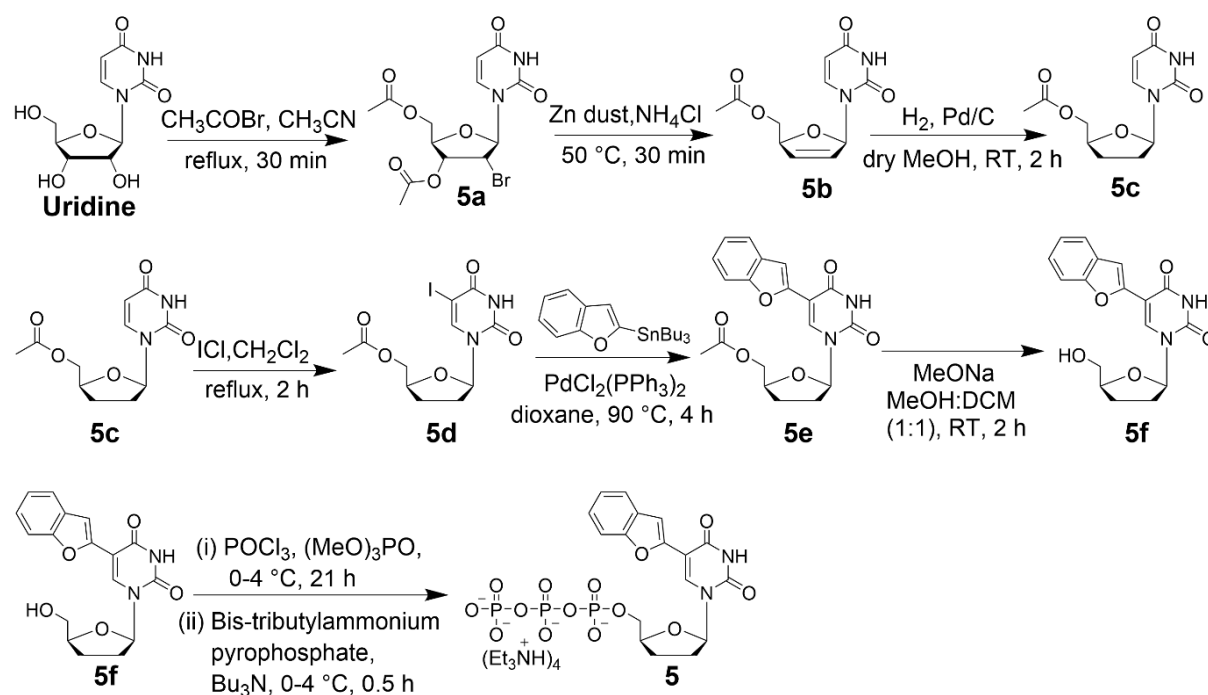


Figure 1. Microenvironment-sensitive nucleotide analogs (1–5).

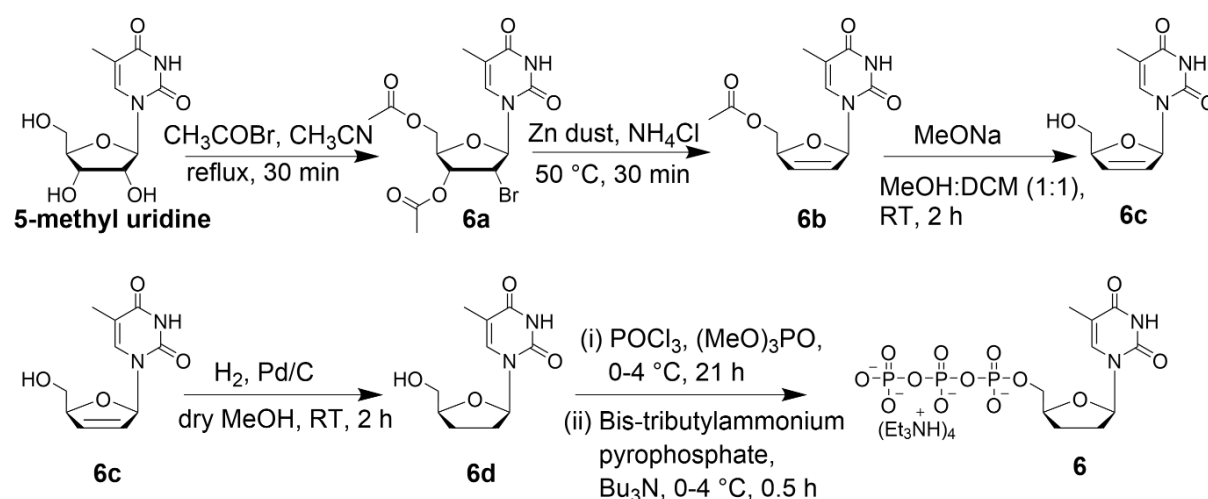
4.2 Result and discussions

4.2.1 Modified nucleotide substrates (1–5)

The ability of TdT to functionalize ssDNA was evaluated using microenvironment-responsive nucleotide analogs (1–5, Figure 1). As described in Chapter 2 and 3, these analogs are minimally perturbing and exhibit useful photophysical properties and ^{19}F NMR signatures that helped in developing biophysical assays to probe nucleic acid structure and recognition. To study the acceptance of these modified nucleotides analog by TdT, we report a strategy for the synthesis of C5-benzofuran modified-2',3'- dideoxy nucleotide (BFddUTP **5**, Scheme 1) and the natural counterpart 2',3'- dideoxy thymidine (ddTTP **6**, Scheme 2). We have used these analogs in various biochemical assays to deduce the key interactions responsible for the template-independent incorporation. Specifically, we have chosen the BFddUTP analog, as it is more emissive than other heterocyclic modified analogs.



Scheme 1. Schematic diagram for the synthesis of 5-benzofuran-2',3'-dideoxyuridine triphosphate (**5**)



Scheme 2. Schematic diagram for the synthesis of 5-methyl-2',3'-dideoxyuridine triphosphate (**6**)

Table 1. ON Sequences used in TdT reaction and Fluorescence assays.

ONs	Sequence
P1	5' FAM-GGA GCT CAG CCT TCA CTG C 3'
P2	5' FAM-GGA GCT CAG CCT TCA CTG CC 3'
P3	5' FAM-GGA GCT CAG CCT TCA CTG CT 3'
P4	5' FAM-GGA GCT CAG CCT TCA CTG CG 3'
P5	5' FAM-GGA GCT CAG CCT TCA CTG CA 3'
P6	5' FAM-GGA GCT CAG CCT TCA CTG CBFdu 3'
P3*	5' GGA GCT CAG CCT TCA CTG CT 3'
P4*	5' GGA GCT CAG CCT TCA CTG CG 3'
P6*	5' GGA GCT CAG CCT TCA CTG CBFdu 3'
P7	5' GGA GCT CAG CCT TCA CTG ddC 3'

ON primer **P1** was used in tailing reactions. ONs **P2**, **P3**, **P4**, **P5**, and **P6** were used in TdT reactions. **P6** is 3'-BFdu-labeled primer **P1**. **P3** and **P4** were used for competition assay. **P3*** and **P4*** are without 5'-FAM-labeled primers **P3** and **P4**, respectively, and have been used for fluorescence assay. **P6*** is the 3'-BFdu-labeled, without 5'-FAM-labeled primer **P6**. **P7** is a 3'-ddC modified primer.

4.2.2 Incorporation of modified deoxynucleotides by TdT

We conducted a tailing reaction to assess the effectiveness of the modified nucleotides (**1–4**) as substrates for TdT. The tailing reactions were carried out in a TdT reaction buffer supplemented with 0.25 mM of CoCl₂, using a 5'-FAM-labeled single-stranded DNA (ssDNA) **P1** in the presence 20 equivalent of either dTTP or the modified nucleotides (**1–4**) at 37 °C (Figure 2A). Aliquots of a reaction were terminated by heating at 75 °C at different time points (15, 30 and 60 min), analyzed by 20% denaturing PAGE, and visualized by scanning at the FAM wavelength. After 60 minutes of the reaction time, the primer was almost consumed. TdT shows a processive incorporation mechanism in the presence of a natural dTTP substrate,

resulting in multiple modifications (Figure 2B, lanes 3–5). In contrast, in the presence of modified nucleotides (**1–4**), it exhibited distributive incorporation and led to the formation of majorly double modifications (Figure 2B, lanes 6–17). Notably, the formation of the double modification products depends on the size of the heterocyclic moiety attached to the uridine ring. Specifically, larger heterocyclic modifications BFdUTP **1**, FBFdUTP **2** and BTdUTP **3** show good substrate properties and largely gave doubly-modified products after 60 min of reaction time (Figure 2B, lanes 6–14). However, SedUTP **4** shows a lower percentage of double modifications and a higher amount of multiple incorporations (Figure 2B, lanes 15–17). The crystal structure of the pre-catalytic state of TdT DNA polymerase with ddTTP revealed that no significant amino acid residue directly interacts with the nucleobase side.^{63,64} However, the insertion of the L398 amino acid sidechain kinked the ssDNA. It disrupts the stacking interaction between the last base and the penultimate base. Further, due to the insertion of this amino acid residue, the incoming nucleotide gets stacked below the 3'-end base of the primer, and the C5 position of the pyrimidine base moves closer to the penultimate base (approx. 3.5Å). Also, it has been observed that the L398 amino acid residue plays an important role in the incorporation process, and mutations of the L398 with the alanine moiety significantly increase the polymerization rate of the enzyme.⁶⁵ This observation raises the suspicion that the interaction between the 3'-end base of the primer and the incoming nucleotide plays a crucial role in the pre-catalytic step to make discrepancies observed in the tailing reaction. Further, once the modifications get incorporated at the 3'-end, the interaction between modifications at C5 positions and the second last base might control further elongation process (Figure 2C). The formation of the double modification product depends on the size of the heterocyclic moiety attached to the uridine ring. Specifically, BFdUTP **1**, FBFdUTP **2** and BTdUTP **3** show good substrate properties and exclusively gave dual modifications. However, upon incorporation into the primer strand due to larger heterocyclic moiety, multiple modifications cannot accommodate in the enzyme pocket. On the other hand, SedUTP **4** shows a lower percentage of double modifications and a higher amount of multiple incorporations. Due to the smaller heterocyclic, multiple incorporated products can be well accommodated in the enzyme pocket.

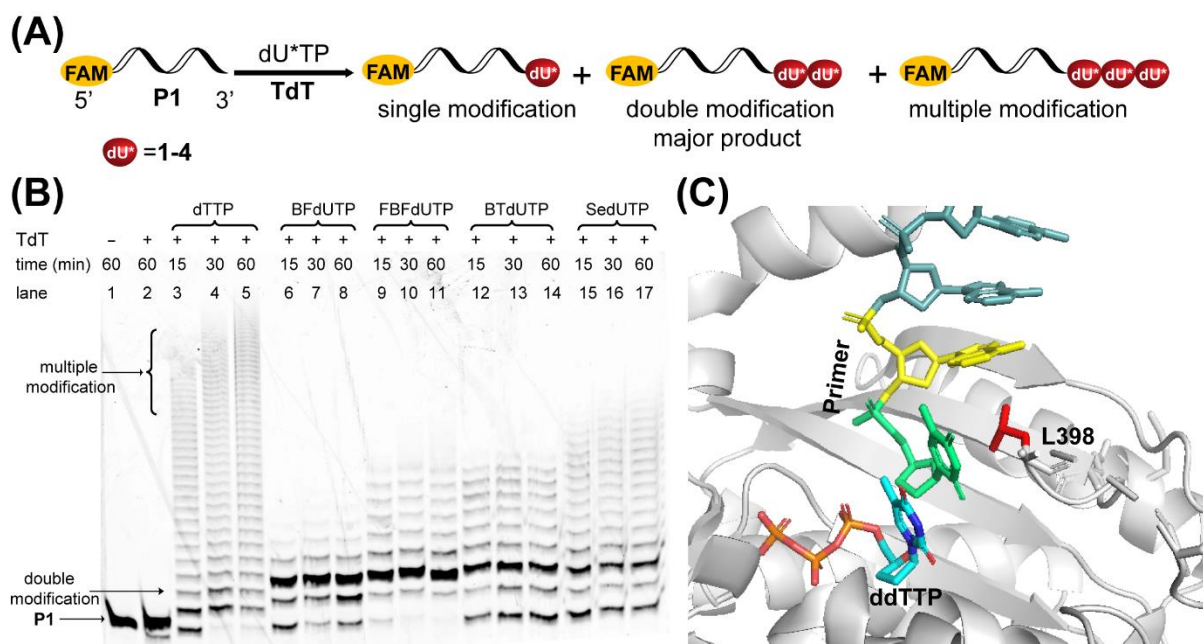


Figure 2. (A) Schematic diagram for the tailing reaction by TdT (B) Incorporation of modified nucleotides (1– 4) using primer P1 (C) crystal structure of the precatalytic state of TdT (PDB:4I27).

4.2.3 Incorporation of modified dideoxynucleotides depends on different initiator primers

The last two nucleotides at the 3'-end of the primer affect the TdT incorporation process, and the end base could potentially interact with the incoming nucleotide through pi-stacking interactions (Figure 2C). To investigate the role of the terminal bases, present in the initiator strands in incorporating modified nucleotides, we conducted a tailing reaction using a set of 5'-FAM-labeled primer strands (P2, P3, P4, P5, P6) terminated with various dinucleotides (CC/CT/CG/CA/CBFdU). As a nucleotide, we have utilized BFddUTP 5, due to the lack of a 3'-OH group, halts the elongation process and gives the single modified product exclusively, and TdT can process the dideoxynucleotides efficiently (Figure 3A). At different time points, reaction aliquots were heat-denatured and resolved by 20% polyacrylamide gel (Figure 3B, C). Quantifying the gel band corresponding to the labeled products shows that primers P4 and P5, terminated with purine bases, exhibit a higher incorporation rate. Even at 1 minute of reaction time, they produced more than 40% of the labeled product. Meanwhile, in 10 minutes, more than 75% of the product was formed. Specifically, the primer P4 containing guanine at the 3'-end base shows maximum labeled product formation (approx. 90%). On the other hand, primers P2 and P3 bearing pyrimidine bases (CC/CT) at the 3'-end, demonstrated less product formation with a lower reaction rate. Notably, the primer ending with a thymine base P3 showed better substrate efficacy and gave more than 60% product formation. The primer ending with a cytosine base P2 exhibited difficulty in accepting the nucleotide and yielded a maximum of 50% single-labeled product. Surprisingly, 3'-end BFdU-labeled primer P6 shows

a very poor rate of incorporation. Only 50% of the product formation occurs after 10 minutes of the reaction time (Figure 3D). Collectively, this data suggests that the pi-stacking interactions in the pre-catalytic steps between the 3'-end base of the primer and the incoming modified nucleotide (BFddUTP 5) control the rate of reactions. The purine bases offer a larger heterocyclic aromatic core for stacking interactions with the incoming nucleotides. Thus, it demonstrates a higher rate of incorporation and more product formation. On the other hand, due to the weaker base stacking interactions in pyrimidine bases, the overall rate of incorporation and the percentage of product formation are also lower compared to purine analog.

However, considering the 3'-end BFdU-modified primer **P6** (CBFdU), although it contains extended conjugated aromatic residues for stronger stacking interactions, it shows poor incorporation efficiency. The crystal structure of the pre-catalytic state revealed that the distance between the C5 methyl group of ddTMP at the 3'-end and the incoming nucleotide is only 4.5 Å away (Figure 2C). We can hypothesize that due to the insertion of the L398 residues in between the last two bases of the primer strand, the stacking interaction between the bases gets hampered⁶⁶ and also, heterocyclic modification or aromatic ring at the C5 position may cause steric clashes with the end base and the incoming nucleotide, which makes it difficult to accommodate in the enzyme pocket and hamper the stacking interactions.

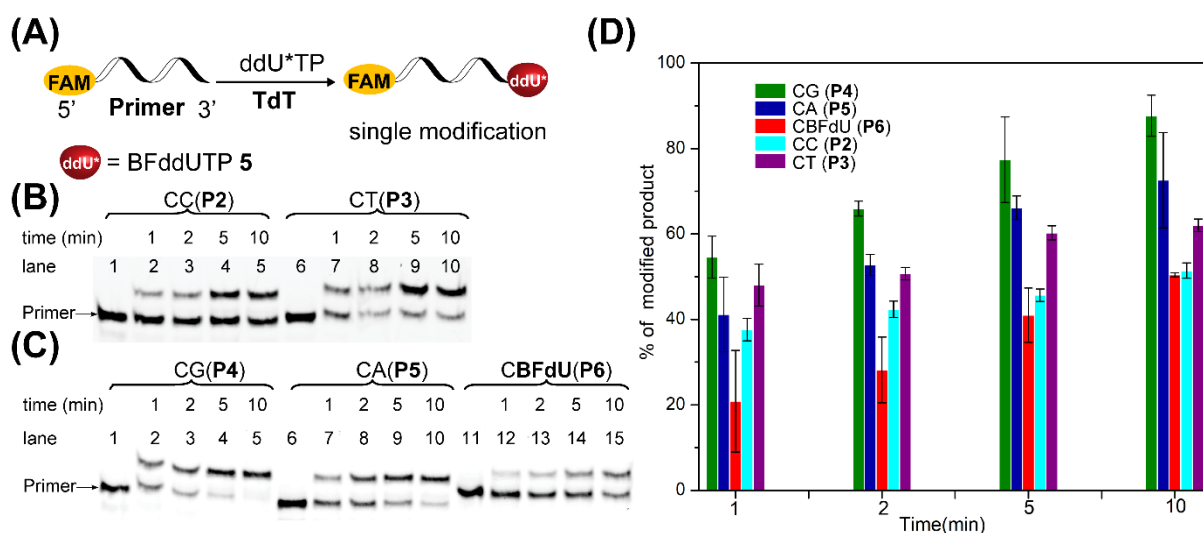


Figure 3. (A) Schematic illustrations of the incorporation of BFddUTP 5 in the presence of different initiators (**P2**, **P3**, **P4**, **P5**, and **P6**). (B, C) 20% PAGE analysis of the incorporation products. (D) Bar diagram of the percentage of modified products against time in the presence of different primers.

4.2.4 Competition between natural and modified nucleotide incorporation

To investigate the differences in the acceptance and incorporation efficiency of modified nucleotides compared to natural counterparts by TdT DNA polymerase, competitive nucleotide

incorporation studies using mixtures with different ratios of BFddUTP **5** and the natural analog ddTTP **6** have been conducted (Figure 4A). The experiments were carried out in the presence of the most efficient primer strand terminated with pyrimidine base **P3** (3'-dT) and the most reactive primer strand terminated with purine bases **P4** (3'-dG labeled). The ratio of unmodified versus modified nucleotide-incorporated products was analysed using 20% denaturing PAGE and visualized by scanning at FAM wavelength. The dideoxy nucleotides (**5** and **6**) efficiently halt the reaction at a single modification due to the lack of a 3'-OH and the 3'-BFddU labeled products migrate slower than the 3'-ddT labeled product in gel. The idea was to determine the relative ratio of the modified versus natural substrate required in order to achieve a 1:1 ratio of 3'-end labeled products containing modified and natural nucleotides. Interestingly, when studying the tailing reaction with **P3** terminating with a dT, we observed that TdT preferentially accepts the BFddUTP **5** than the natural ddTTP **6**. It incorporates the modified analog equally to the unmodified nucleotide when a ~7.5-fold excess of ddTTP was used (Figure 4C). Similarly, when the dG-modified primer **P4** was extended, only ~2.5-fold excess of ddTTP was required to exhibit equal incorporation efficiencies (Figure 4B). These results suggest that the benzofuran ring attached to the uridine base provides extra stability during the incorporation process by helping pi-stacking interaction with the end bases, reflecting its preferential incorporation over the natural analog (Figure 5A–D). In the case of the primer strand end with a pyrimidine base (**P3**), it provides lower extends of stacking interaction for the incoming nucleotides (Figure 5A, B). Thus, extra stability from BFddUTP analog, makes it much more selective than the natural one (approx. 7.5 times) (Figure 4C, 5B). Since the initiator strands end with a purine base (**P4**) and provide larger extents of stacking interactions for incoming nucleotides (Figure 5C, D), the extra stability from the heterocyclic ring of BFddUTP is not much dominated. TdT shows a lower selectivity for BFddUTP (approximately 2.5 times) over dTTP than the primer strand end with a pyrimidine base (Figure 4B, 5D). These data clearly suggest that in the first incorporation of modified nucleotides, the modified nucleotides show high incorporation efficiency due to the stacking interactions with the terminal bases. However, once the modified nucleotides are incorporated in the 3'-end, it demonstrates less substrate efficacy in accepting incoming nucleotides. It speculates the reason behind the low efficiency in incorporating the multiple C5-modified heterocyclic moiety.

To gain insight into the incorporation mechanism of TdT, we have tried to crystalize the ternary complex of the TdT with a primer (five consecutive dA residue (dA)₅) and BFddUTP in the active site (Figure 6A)^{60,63}. However, we were not able to get a ternary

complex of crystal structure. We obtained a high-resolution binary complex structure of TdT containing the primer only (resolution 1.9 Å, Figure 6B, C).

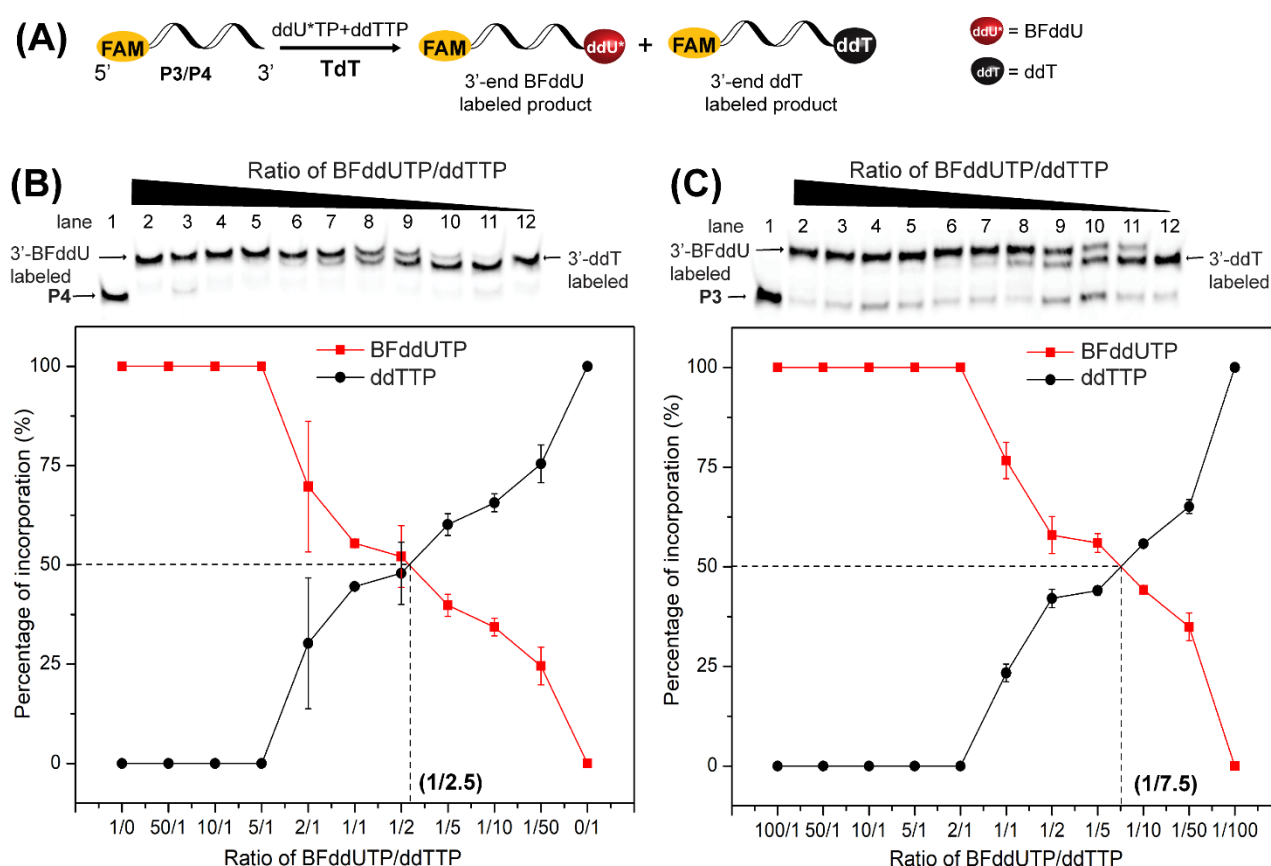


Figure 4. (A) Illustrations of the competition reactions assay. (B) PAGE analysis of the competition experiment using TdT for BFddUTP and ddTTP and corresponding graphical readout for primer **P4**. (C) PAGE analysis and graphical readout for competition between BFddUTP and ddTTP for primer **P3**. The point of 50 % incorporation is indicated with a dashed line. All the data was collected in triplicates.

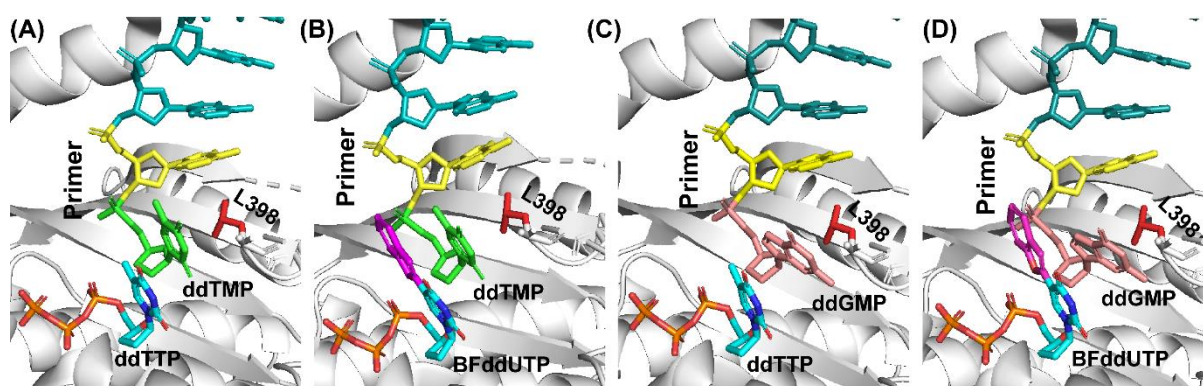


Figure 5. (A) Crystal structure of the pre-catalytic state of TdT when ddTTP is present (PDB:4I27). Model structures for (B) the pre-catalytic state of modified nucleotide BFddUTP 5, (C) the precatalytic state when ddTTP stacked with primer end with dG and (D) BFddUTP in the active site when the end base is dG. The model structures were made in Pymol software, where the crystal structure of the precatalytic state of TdT (PDB:4I27) was used as a reference.

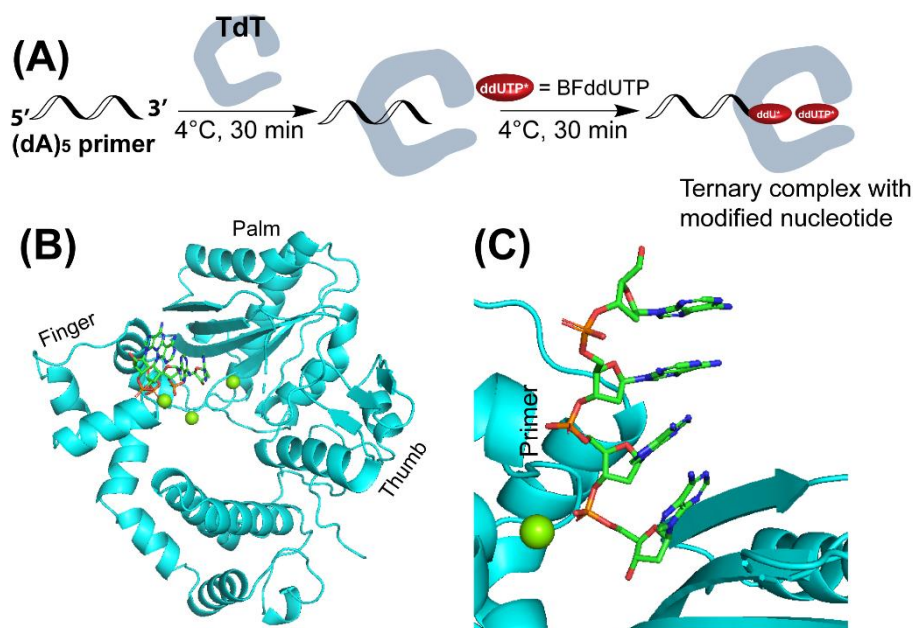


Figure 6. (A) Schematic illustrations of the crystallization set-up method for the BFddUTP-TdT-primer ternary complex (B) Crystal structure of the binary complex of TdT and primer (C) Close-view of the active-site.

4.2.5 Insight into the binary complex formation by fluorescence

Further, we explored whether the BFdU-labeled primer **P6*** can be used to sense the binding of ssDNA to the TdT enzyme (Figure 7A). For this purpose, primer **P6*** was incubated with varying concentrations of TdT for 30 min at 37°C. Fluorescence was recorded by exciting the sample at 330 nm. ssDNA exhibits a low intensity fluorescence band due to stacking interactions with the neighboring base, however, upon increasing TdT concentration, a dose-dependent enhancement in fluorescence intensity was observed (Figure 7B). BFdU acts as a molecular rotor element due to the rotatable aryl-aryl single bond, which connects the heterocyclic rings with the uridine base.^{67–70} Also, we have reported that increases in the viscosity of the surroundings restrict the free rotation of this single bond, which leads to an increase in fluorescence intensity. It has been shown earlier that BFdU exhibits an enhancing and red shift in fluorescence upon an increase in micro-polarity. In Chapters 2 and 3, we have shown that once this modification enters the enzyme's active sites, the free rotation of the single bond might get rigidified and show enhancement in fluorescence intensity. Notably, it has been reported that TdT binds the last three bases of the primer from the 3'-end to facilitate the elongation process, and our modification is also placed at the 3'-end. Therefore, these observations indicate that the enhancement in fluorescence intensity is due to the binding of the primer **P6*** with the TdT enzyme pocket, which led to the rigidification of the single bond

and minor changes in polarity. Upon fitting the normalized fluorescence changes to a Hill equation, we observed an apparent K_d of $0.34 \pm 0.06 \mu\text{M}$ (Figure 7C). To validate that TdT is indeed binding to the BFdU-labeled ssDNA **P6***, we titrated 5'-FAM, 3'-BFdU labeled DNA **P6** and unmodified 5'-FAM labeled primer **P3** with TdT and monitored the binary complex formation using microscale thermophoresis (MST) assay at FAM wavelength. The MST measurements revealed the formation of the binary complexes, wherein BFdU-modified primer **P6** ($K_d = 1.07 \mu\text{M}$) and natural primer **P3** ($K_d = 1.86 \mu\text{M}$) exhibited almost similar binding affinities for TdT (Figure 8). This indicates that the modification does not hamper the binary complex formation.

The next objective is to prove that the increase in fluorescence intensity results from the specific binding of BFdU in the ssDNA to the active site of the enzyme and not through a non-specific binding between the modification and TdT. As the BFdU is a thymidine analog, we conducted a competitive assay between a 3'-BFdU labeled primer **P6*** and a 3'-dT labeled primer **P3***. A binary complex of TdT and primer **P6*** was titrated with increasing concentrations of the unlabeled primer end with thymidine base **P3***, and changes in the fluorescence were recorded (Figure 9). A concentration-dependent decrease was observed upon increasing the unlabeled ssDNA **P3*** concentration. The results indicate that the unlabeled primer **P3*** competes with the BFdU-labeled DNA for binding into the catalytic active site of TdT. The decrease in fluorescence is due to the displacement of the modified DNA from the enzyme pocket by the unlabeled DNA. It is noted that the 20-fold concentration of the unlabeled ssDNA cannot completely displace the modified primer **P6*** from the enzyme pocket. This indicates that BFdU-labeled ssDNA binds more effectively to the active site than unlabeled DNA.

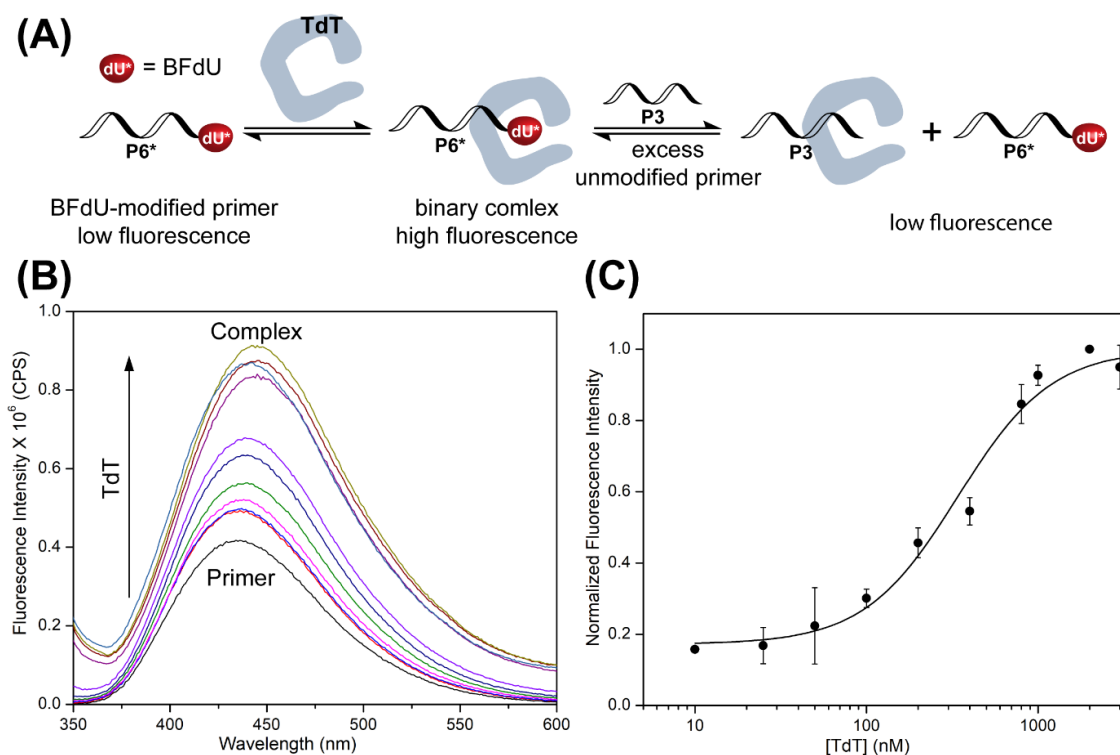


Figure 7. BFdU reports the binary complex formation by TdT **(A)** illustrations of the binary complex formation by fluorescence assay. **(B)** Fluorescence spectra for the primer **P6*** (50 nM) with increasing concentrations of TdT (0–2 μ M). Samples were excited at 330 nm with excitation and emission slit widths of 7 nm and 8 nm, respectively. **(C)** Curve fits for the binding of the TdT. Normalized fluorescence intensity λ_{em} (430 nm) versus the concentration of the TdT was plotted.

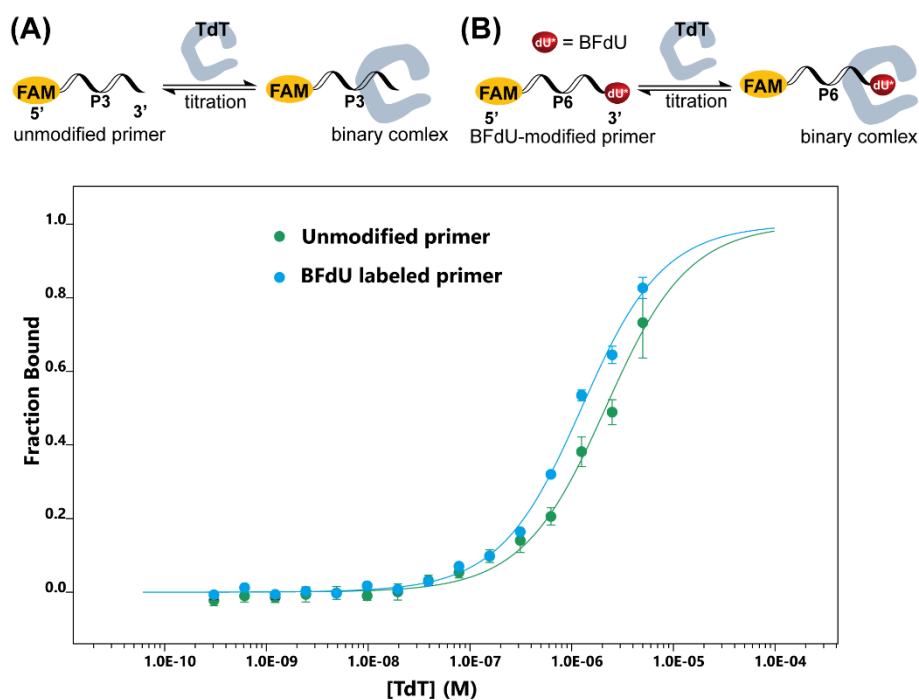


Figure 8. **(A, B)** The schematic diagram for the MST assay for 5'-FAM labeled unmodified and 3'-BFdU modified primer, **P3** and **P6**, respectively, with TdT. MST titration curve characterizing the interaction of TdT with natural primer (green color) and 3'-BFdU modified primer (cyan color). *We

want to acknowledge Prof. Souvik Maiti from Institute of Genomics and Integrative Biology, Delhi, for helping us with MST assay.

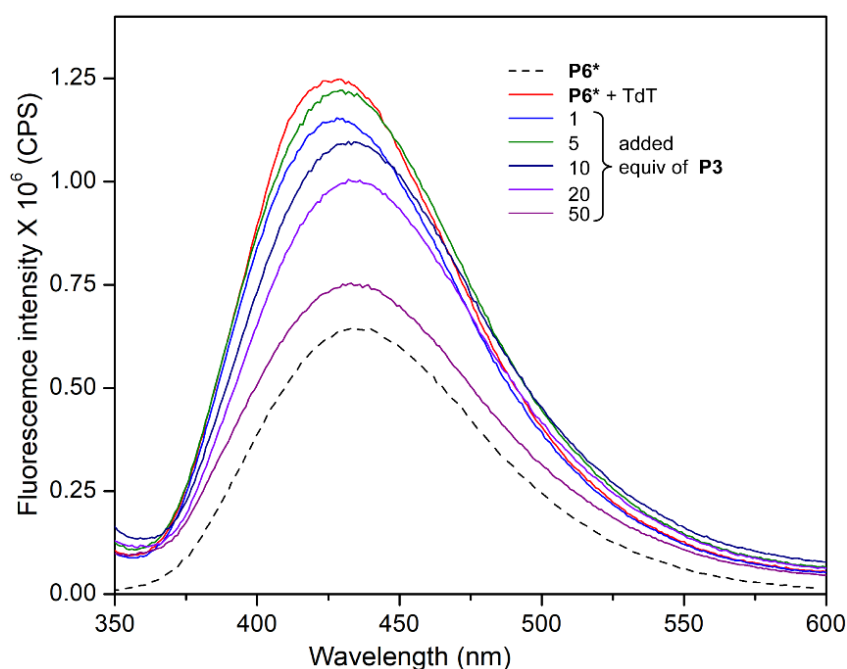


Figure 9. Fluorescence was employed in a competitive displacement assay. The BFdU-modified primer **P6*** (50 nM) exhibited minimal fluorescence (depicted by the black dashed line). A binary complex of **P6*** (50 nM) and TdT (1 μ M) shows very high fluorescence (red line). The introduction of increasing amounts of unlabeled primer **P3** led to the displacement of the complex, resulting in a gradual reduction in fluorescence (illustrated by the solid lines).

4.2.6 Detection of TdT activity by fluorescence

To achieve fluorescence detection of TdT activity, we utilized a microenvironment-sensitive dideoxy fluorescent nucleotide analog (BFddUTP **5**) (Figure 10A). We focused on incorporating BFddUTP, as it is well recognized by TdT, and TdT can process this nucleotide even at very low concentrations. Due to the absence of the 3'-OH group, it extensively provides a single-modified product. The BFddUTP **5** (100 nM) was incubated with an unlabelled primer **P4*** (terminating with dG, see Table 1) in the presence of varying concentrations of TdT in a reaction buffer at 37 °C. Samples were excited at 330 nm and fluorescence measurements were conducted at the emission maximum of BFddUTP **5** (447nm) and were measured every 10 seconds for 10 minutes. With increasing concentrations of TdT, the rate of decrease in fluorescence intensity increases (Figure 10B). Furthermore, the overall fluorescence decreased after 10 minutes (Figure 10C) and the steepness of the curves exhibited TdT concentration dependence (Figure 10B). Upon incorporation in the primer strand, a decrease in fluorescence emission was observed, which signifies the stacking interaction with the neighboring base (Figure 10C). Further, we have calculated the rate of incorporation for each enzyme

concentration by fitting it into the pseudo-first-order reaction kinetics. A plot of K_{obs} (sec^{-1}) versus TdT concentrations (nM) gives a gradual increase in the rate of reaction with increasing enzyme concentration (Figure 10D, E). Notably, the fluorescence did not decrease in the absence of TdT (depicted by the black line in Figure 10B, D and the black dotted line in Figure 10C). Further, to prove that the change in fluorescence is indeed happening due to the incorporation, we have performed an incorporation assay in the presence of 3'-dideoxycytidine modified oligonucleotide **P7**. Due to the absence of the 3'-OH group at the primer end, it is not able to incorporate modified nucleotide thus no change in fluorescence is observed (Figure 10B red line and Figure 10C red line).

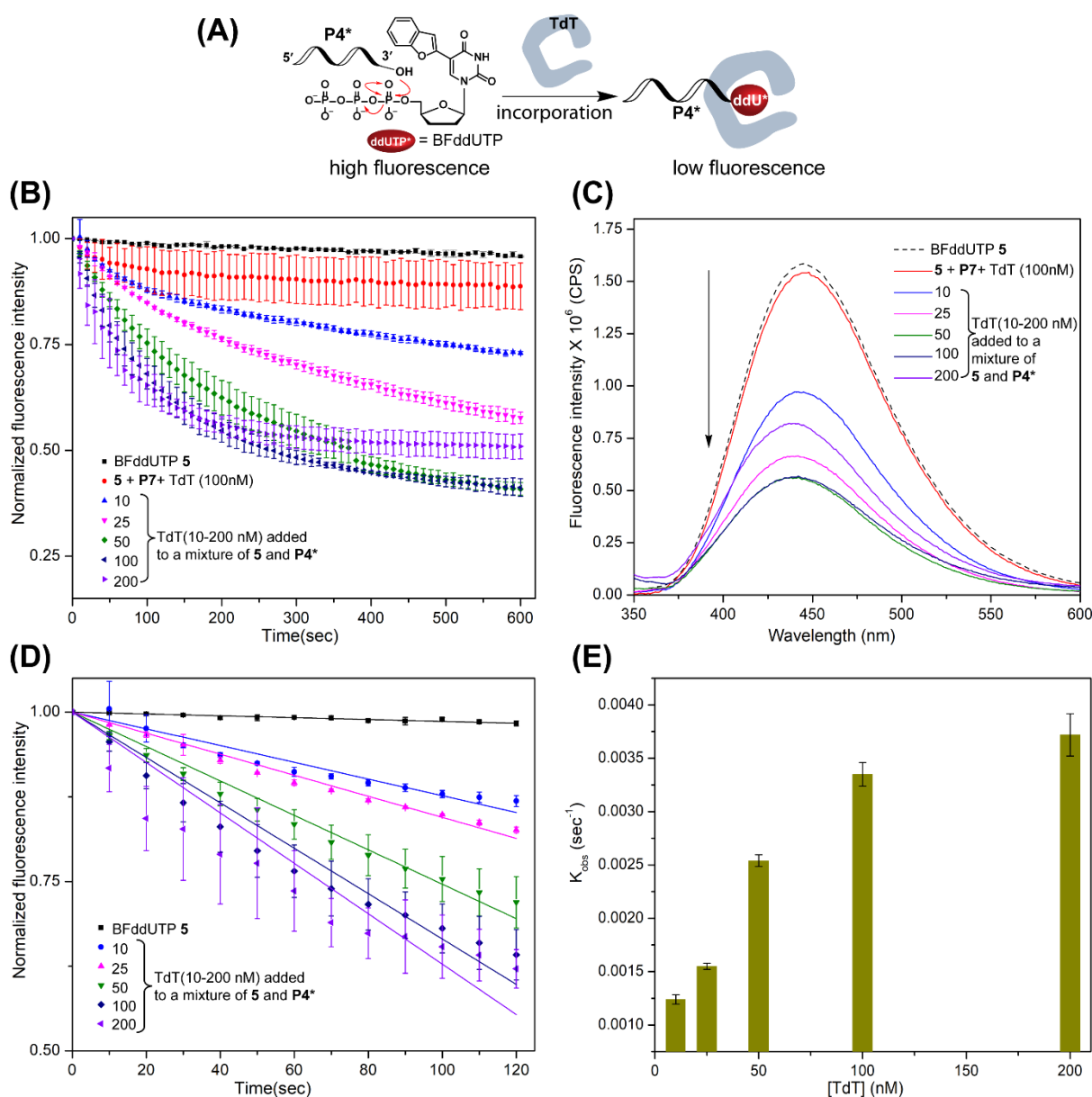


Figure 10. Tailing reaction by TdT **(A)** Schematic representation of TdT reaction in the presence of BFddUTP 5 **(B)** detection of TdT activity with BFddUTP 5 with varying concentrations of TdT (0–200

nM) were incubated with 100 nM of the BFddUTP **5** and 200 nM of primer **P4*** at 37 °C and the fluorescence at 447 nm was measured every 10 seconds for 10 minutes, using 330 nm as excitation wavelength. Excitation and emission slit widths were 7 and 8 nm, respectively. (C) Corresponding fluorescence spectra illustrate the decrease in fluorescence intensity of BFddUTP **5** (100 nM) with increasing TdT concentration (0–200 nM) when the **P4*** primer (200 nM) is present. Control experiments demonstrate minimal changes in fluorescence intensity when BFddUTP **5** (100 nM) alone or **P7** primer (**P7** contained a ddC at the 3'-end) and BFddUTP (100 nM) are present does not exhibit a reduction in fluorescence intensity. (D) The rate constant (k_{obs}) and the initial points till 150 seconds were taken and fitted into a pseudo-first-order kinetics equation. (E) A graphical representation of the increase in the rate constant with an increase in TdT concentration.

4.2.7 Monitoring the competition with ddTTP by fluorescence

To confirm that our detection system can report the inhibition or termination in the enzymatic incorporation process, we performed a competition assay in the presence of ddTTP **6** (Figure 11A). Dideoxy nucleotides are suitable substrates for TdT incorporation, and they also stop the polymerization process as they lack the 3'-OH group. Therefore, the incorporation of ddTTP will prevent the tailing reaction and the rate of decrease in fluorescence should be significantly lower than when the modified nucleotide BFddUTP is used alone. To check the efficacy of the detection system to report the competition assay, samples containing BFddUTP (100 nM) and TdT (100 nM) were incubated with increasing concentrations of the (0–50 equiv.) ddTTP at 37 °C. Reactions were monitored at 447 nm wavelength every 10 seconds for 10 minutes. With the increasing equivalent of the ddTTP, the rate of change in the fluorescence intensity substantially diminished (Figure 11B, C). Titration of the ddTTP resulted in the recovery of the fluorescence intensity from the quenched state. Since it has been proven previously that the decreased fluorescence intensity is due to the π -stacking interactions between BFddUTP and the neighboring base, the observed recovery in the fluorescence intensity indicates the direct competition of ddTTP over BFddUTP. Notably, a 50-fold ddTTP concentration was insufficient to recover the fluorescence, implying that BFddUTP is a better substrate than ddTTP. As we observed previously in the competition assay (section 4.2.4), the BFddUTP gets incorporated efficiently than the natural analogs, almost 2.5 times more due to extra stacking interaction. Thus, we can conclude that the real-time fluorescence data correlates with the incorporation profile as we observed by biochemical assay.

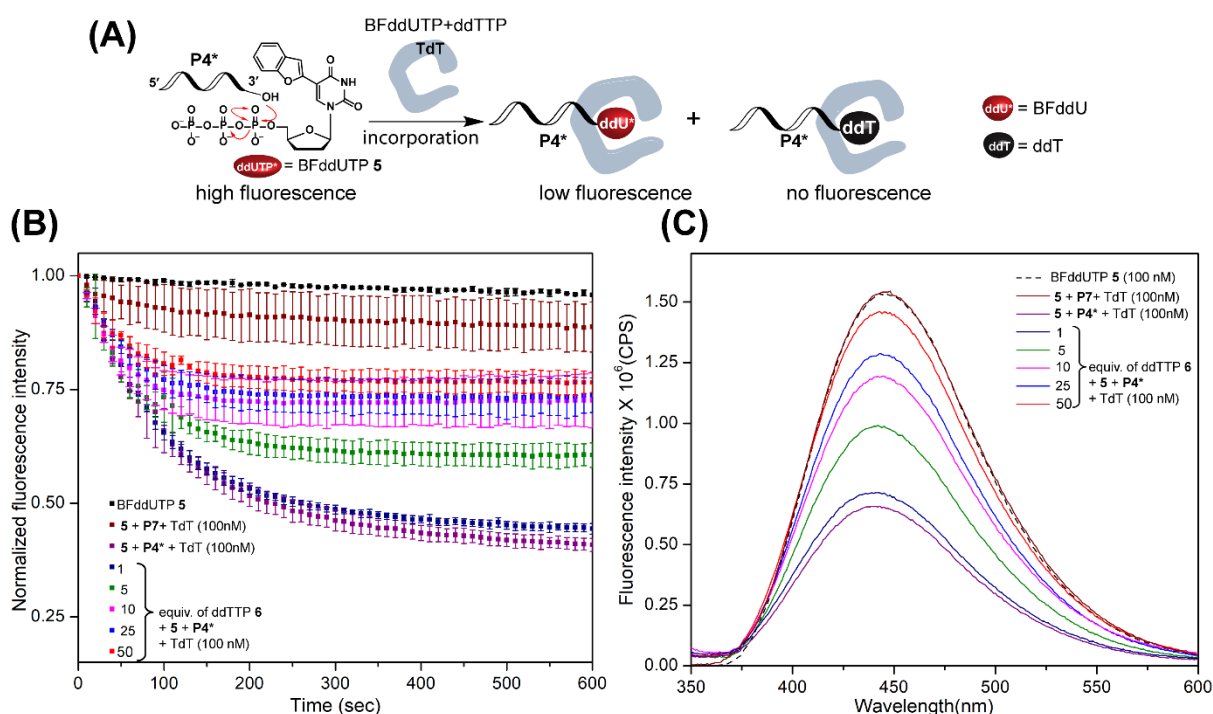


Figure 11. (A) Competition experiments with ddTTP **6**. (B) TdT (100 nM) was added to a mixture of BFddUTP **5** (100 nM) and varying concentrations of ddTTP (0–5 μ M) and primer **P4*** (200 nM). Reactions were monitored at 37 °C for 10 minutes. Subsequently, the fluorescence at 447 nm was measured every 10 seconds for 10 minutes. (C) corresponding fluorescence spectra for each ddTTP **6** concentration.

4.3 Conclusion

We have introduced useful microenvironment-sensitive nucleotide probes carrying different sizes of heterocyclic moiety, namely BFddUTP, FBFddUTP, BTdUTP and SedUTP, at the 3'-end of the ssDNA by utilizing terminal deoxynucleotide transferases (TdT). Larger heterocyclic modifications BFddUTP, BTdUTP, and FBFddUTP result in higher efficiency in distributive incorporation compared to the smaller heterocyclic modification SedUTP or natural thymidine analog. We hypothesized that stacking interaction between the incoming nucleotides with the terminal base of the primer in the pre-catalytic step is responsible for the discrepancies in the TdT activity. To elucidate the key interaction, we used benzofuran-modified-2',3'-dideoxyuridine triphosphate (BFddUTP) and 2',3'-dideoxy thymidine triphosphate (ddTTP) and compared their incorporation by TdT DNA polymerase. These dideoxy-modified nucleotides are suitable substrates for the TdT DNA polymerase and terminate the extension process at a single nucleotide-labeled product. We performed an incorporation reaction in the presence of primer strands terminated with either purine or pyrimidine bases. It was found that the purine bases (dG and dA) exhibit higher incorporation efficiency than pyrimidine analogs (dC and dT). Similarly, when we performed a competition reaction between the BFddUTP and ddTTP, it was observed that the BFddUTP exhibits a high

incorporation efficiency compared to the natural substrate. These observations align with the enhanced stacking interaction between the incoming nucleotide analog and the 3'-end nucleobase of the primer in the pre-catalytic state. Further, we have utilized the 3'- BFdU-labeled ssDNA to monitor the binary complex formation with the TdT enzyme by fluorescence and calculated the apparent dissociation constant in solution. We also describe the development of an assay to monitor the TdT activity using BFddUTP by fluorescence. Finally, the strategy of monitoring competition between BFddUTP and ddTTP will help in the set up of an assay for a therapeutic screening platform to find potential small molecule inhibitors of TdT DNA polymerases.

4.4 Experimental sections

4.4.1 Materials

5-methyl uridine and uridine were purchased from TCI chemicals. Bis (triphenylphosphine)-palladium (II) chloride, butyllithium, and tributyltin chloride were purchased from Sigma Aldrich. DNA templates were purified by denaturing (PAGE) and desalted using Sep-Pak Classic C18 cartridges (Waters Corporation). dNTPs were purchased from Thermo Fischer Scientific.

4.4.2. Instrumentation

NMR spectra were recorded on Bruker Ultrashield 400 MHz WB Plus spectrometers. Mass measurements were conducted using a Water Synapt G2 High Definition mass spectrometer. HPLC analysis was done using Agilent Technologies 1260 Infinity HPLC. Reverse-phase flash chromatographic (C18 RediSepRf column) purifications of nucleotides were carried out using Teledyne ISCO Combi Flash *Rf*. ESI-MS analysis of DNA ONs was performed using a Water Synapt G-Si Mass spectrometry instrument in negative mode. DNA and nucleotides were quantified by measuring absorbance in Shimadzu UV-2600 or NanoDrop™ 2000c spectrophotometer. Fluorescent DNA ONs were resolved by PAGE on an OWL S4S sequencing gel electrophoresis instrument and were imaged using Typhoon FLA 9500 Variable mode imager. Steady-state fluorescence spectra were recorded using a micro-fluorescence cuvette (Hellma, path length 1.0 cm) on a Horiba Fluoromax-4 spectrophotometer.

4.4.3 Synthesis of 5-benzofuran-2',3' -dideoxyuridine triphosphate (5)

Synthesis of 3',5' -Di-O-acetyl-2' -bromouridine (5a)

Uridine (1 g, 4.0 mmol) was dissolved in anhydrous acetonitrile (20 mL) and heated at reflux for 20 minutes. Acetyl bromide (4.2 g, 34.8 mmol) was added to this mixture dropwise and

refluxed for 0.5 h. The solution was then cooled at room temperature and concentrated in vacuo till it a yellowish oily product. The resulting compound was dissolved in DCM (20 mL) and washed with water (3 X 10 mL). The organic layer was collected and dried over Na₂SO₄. The product was concentrated and purified by column chromatography to yield yellow powder (1.2 g, 3.0 mmol, 75 %). TLC (100 % ETAC) *R_f*: 0.76; ¹H-NMR (400 MHz, *d*₆-DMSO): δ (ppm) 11.53 (s, 1H), 7.67 (d, *J* = 8.0 Hz, 1H), 6.13 (d, *J* = 7.2 Hz, 1H), 5.76 (dd, *J* = 8 Hz, 2.4 Hz, 1H), 5.23 (dd, *J* = 6 Hz, 3.6 Hz, 1H), 4.99 (dd, *J* = 7.6 Hz, 6 Hz, 1H), 4.25–4.33 (m, 3H), 2.14 (s, 3H), 2.07 (s, 3H); ¹³C NMR (100 MHz, *d*₆-DMSO): δ (ppm) 170.1, 169.3, 162.8, 150.5, 140.1, 102.9, 88.8, 79.7, 62.8, 47.5, 20.5; HRMS: *m/z* calculated for C₁₃H₁₅N₂O₇Br [M+H]⁺=391.0141, found [M+H]⁺ = 391.0141.

Synthesis of 5'-O-Acetyl-2',3'-didehydro-2', 3'-dideoxyuridine (5b)

A solution of compound **5a** (5 g, 12.8 mmol), freshly activated Zn powder (6.3 g, 95.8 mmol), NH₄Cl (5.1 g, 95.8 mmol) in anhydrous Acetonitrile was heated at 50 °C for 45 min. reaction progress was monitored by TLC. After completion of the reaction, the reaction mixture was filtered through a celite pad and evaporated at reduced pressure. Further, the product was purified by column chromatography to get a solid white powder (1.8 g, 7.1 mmol, 55%). TLC (100 % ETAC) *R_f*: 0.44; ¹H-NMR (400 MHz, *d*₆-DMSO): δ (ppm) 11.37 (s, 1H), 7.45 (d, *J* = 8.0 Hz, 1H), 6.80–6.82 (m, 1H), 6.42 (dt, *J* = 6 Hz, 2 Hz, 1H), 6.00 (dt, *J* = 6 Hz, 1.6 Hz, 1H), 5.66 (d, 1H, *J*=8 Hz), 4.97– 4.99 (m, 1H), 4.18 (d, *J* = 3.6 Hz, 1H), 2.01 (s, 3H). ¹³C NMR (100 MHz, *d*₆-DMSO): 170.8, 163.8, 150.9, 140.9, 102, 85.7, 78.4, 65.7, 31.3, 26, 21; HRMS: *m/z* calculated for C₁₁ H₁₂N₂O₅ [M-H]⁻=251.0673, found [M-H]⁻ =251.0672.

Synthesis of 5'-O-Acetyl-2',3',-dideoxyuridine (5c)

Compound **5b** (900 mg, 3.6 mmol) was dissolved in anhydrous methanol (15 mL) and hydrogenated in the presence of a catalyst (100 mg) (wet 10% palladium on carbon). The reaction mixture was stirred for 2 h at room temperature. After that, the catalyst was removed by filtration, and the filtrate was concentrated at reduced pressure. Compound **4** was obtained after column chromatography (800 mg, 3.1mmol, 87 %). TLC (100 % ETAC) *R_f*: 0.44; ¹H-NMR (400 MHz, *d*₆-DMSO): δ (ppm) 11.37 (s, 1H), 7.45 (d, *J* = 8.0 Hz, 1H), 6.80–6.82 (m, 1H), 6.42 (dt, *J* = 6 Hz, 2 Hz, 1H), 6.00 (dt, *J* = 6 Hz, 1.6 Hz, 1H), 5.66 (d, 1H, *J*=8 Hz), 4.97– 4.99 (m, 1H), 4.18 (d, *J* = 3.6 Hz, 1H), 2.01 (s, 3H). ¹³C NMR (100 MHz, *d*₆-DMSO): 170.7, 163.7, 150.9, 140.9, 102, 85.7, 78.4, 65.4, 31.3, 26, 21.1. HRMS: *m/z* calculated for C₁₁ H₁₄N₂O₅ [M-H]⁻=253.0830, found [M-H]⁻ = 253.0822.

Synthesis of 5'-O-Acetyl-5-iodo-2',3',-dideoxyuridine (5d)

ICl (2.10 gm, 12.9 mmol, 3.3 equiv) and compound **5c** (1.0 gm, 3.9 mmol, 1 equiv) were dissolved in anhydrous DCM and heated to reflux for 2 h and then cooled to room temperature. To reduce the excess amount of ICl, a minimum quantity of 2% NaHSO₃ aqueous solution was added. The organic phase was collected and washed with distilled water (3 x 50 mL) and dried over Na₂SO₄. The organic phase was evaporated, and further purification was performed by column chromatography to get the desired product compound **5d** (3.17 gm, 8.4 mmol, 65 % of yield). TLC (100 % ETAC) *R_f*: 0.44; ¹H NMR (400 MHz, DMSO-d₆) δ 11.68 (s, 1H), 8.00 (s, 1H), 5.90 (dd, *J* = 7.0, 3.8 Hz, 1H), 4.22–4.25 (m, 3H), 2.30–2.27 (m, 1H), 2.13 (s, 3H), 2.10–2.05 (m, 1H), 1.99–1.95 (m, 1H), 1.83–1.76 (m, 1H). ¹³C NMR (100 MHz, *d*₆-DMSO): 170.3, 160.6, 150.1, 144.5, 86, 78.4, 69.2, 64.5, 31.4, 25, 21; HRMS: *m/z* calculated for C₁₁H₁₃IN₂O₅ [M-H]⁺ = 378.9796, found [M-H]⁺ = 378.9851

Synthesis of 5'-O-Acetyl-5-benzofuran-2',3',-dideoxyuridine (5e)

Dry dioxane (25 mL) was taken in a round bottom flask and degassed for 20 min in the presence of nitrogen. Compound **5d** (1 gm, 2.6 mmol, 1 equiv) was suspended into dry dioxane and followed by palladium catalyst Pd(PPh₃)₂ Cl₂ (93 mg, 0.13 mmol, 0.05 equiv) was added in N₂ atmosphere. Into that solution, stannylated Benzofuran (2.71 g, 6.7 mmol, 2.5 equiv) was added and refluxed for 2 h at 90 °C. After the reaction was completed, the reaction mixture was evaporated and resuspended into MeOH. The product was precipitated along with a fluorescent impurity upon resuspension in MeOH. So, the product was further purified by column chromatography to get a white solid powder compound **6** (amount: 635 mg, 67 % of yield). TLC (MeOH: DCM 5:95) *R_f*: 0.64; ¹H-NMR (400 MHz, *d*₆-DMSO): δ (ppm) 8.25 (s, 1H), 7.64 (d, *J* = 4.0 Hz, 1H), 7.51 (d, *J* = 8.0 Hz, 1H), 7.37 (s, 1H), 7.29 (t, *J* = 8.0 Hz 1H), 7.23 (t, *J* = 8.0 Hz 1H), 6.02 (dd, *J* = 4 Hz, 2.8 Hz, 1H), 4.39–4.35 (m, 1H), 4.34–4.32 (m, 1H), 2.44–2.35 (m, 1H), 2.21–2.14 (m, 1H), 2.03 (s, 3H), 1.89–1.80 (m, 1H). ¹³C NMR (100 MHz, *d*₆-DMSO): δ (ppm) 175.5, 165.6, 158.1, 154.5, 154.2, 141.3, 134, 129.7, 128.3, 126.3, 115.7, 109.4, 100, 91.8, 83.8, 69.9, 37, 30.4, 25.7. HRMS: *m/z* calculated for C₁₉ H₁₇ N₂ O₆ [M-H]⁺ = 369.3535, found [M-H]⁺ = 369.1075.

Synthesis of 5-benzofuran-2',3',-dideoxyuridine (5f)

Compound **5e** (230 mg, 0.63 mmol) and sodium methoxide (400 mg, 9.42 mmol) were dissolved in a solution of dry MeOH: DCM (1:1) (15 mL) and the mixture was stirred for 2.5 hours at room temperature. The progress of the reaction was monitored by TLC (5% MeOH: DCM). The reaction was quenched by adding Amberlite IRN-120 H⁺ resin, and the resulting

mixture was filtered using a Whatman filter paper. The organic phase was then dried over Na_2SO_4 and evaporated to dryness. The compound was purified through column chromatography, yielding a white solid powder (153 mg, 75% yield). TLC (MeOH: DCM 5:95) R_f : 0.54; ^1H -NMR (400 MHz, d_6 -DMSO): δ (ppm) 11.71 (s, 1H), 8.88 (s, 1H), 7.62 (d, J = 7.6 Hz, 1H), 7.56 (d, J = 8.0 Hz, 1H), 7.32 (s, 1H), 7.29 (t, J = 6.8 Hz, 1H), 7.23 (t, J = 7.2 Hz, 1H), 6.05 (dd, J = 6 Hz, 2.4 Hz, 1H), 5.26 (t, J = 4.8 Hz, 1H), 4.17–4.12 (m, 1H), 3.88 (td, J = 10.6 Hz, 4.4 Hz, 1H), 3.67 (td, J = 12 Hz, 3.6 Hz, 1H), 2.41–2.31 (m, 1H), 2.15–2.11 (m, 1H), 1.97–1.90 (m, 2H). ^{13}C NMR (100 MHz, d_6 -DMSO): δ (ppm) 160.4, 153, 149.3, 137.4, 128.8, 124.1, 123, 120.9, 110.8, 103.9, 103.5, 86.2, 82.4, 61.4, 32.7, 24.2. HRMS: m/z calculated for $\text{C}_{19}\text{H}_{17}\text{N}_2\text{O}_6$ $[\text{M}-\text{H}]^-$ = 327.3165, found $[\text{M}-\text{H}]^-$ = 327.0970

Synthesis of 5-benzofuran-2',3',-dideoxyuridine triphosphate (5)

The nucleotide analog **5** was synthesized using the procedure as mentioned in Chapter 2 (section 2.4.3) from Compound **5f** (70 mg, 0.21 mmol, 1 equiv.). Yield (32 mg, 26%). ^1H -NMR (400 MHz, D_2O): δ (ppm) 7.85 (s, 1H), 7.44 (s, 1H), 7.42 (s, 1H), 7.20 (t, J = 7.2 Hz, 1H), 7.1 (t, J = 7.2 Hz, 1H), 7.03 (s, 1H), 5.85–5.83 (m, 1H), 4.32–4.28 (m, 1H), 4.20–4.15 (m, 1H), 4.05–3.99 (m, 1H), 2.88–2.80 (m, 1H), 2.40–2.33 (m, 1H), 2.12–2.05 (m, 2H); ^{13}C NMR (100 MHz, D_2O): δ (ppm) 180.8, 161.6, 153.2, 150.1, 147.8, 136.6, 128.9, 125, 123.2, 121, 111.2, 105.6, 104.8, 86.9, 67.6, 30.7, 25.6; ^{31}P NMR (162 MHz, D_2O): δ (ppm) -10.65 (br, P_α) -11.32 (d, J = 17.82, P_γ), -23.16 (br, P_β); HRMS: m/z calculated for $\text{C}_{17}\text{H}_{18}\text{N}_2\text{O}_{14}\text{P}_3$ $[\text{M}-\text{H}]^-$ = 567.2528, found $[\text{M}-\text{H}]^-$ = 566.9977.

Synthesis of 3',5'-Di-O-acetyl-2'-bromo 5 methyluridine (6a)

5-methyl uridine (1 g, 4.0 mmol) was dissolved in anhydrous acetonitrile (20 mL) and heated at reflux for 20 minutes. Acetyl bromide (4.2 g, 34.8 mmol) was added to this mixture dropwise and refluxed for 0.5 h. The solution was then cooled at room temperature and concentrated in vacuo till it was a yellowish oily product. The resulting compound was dissolved in DCM (20 mL) and washed with water (3 x 10 mL). The organic layer was collected and dried over Na_2SO_4 . The product was concentrated and purified by column chromatography to yield yellow powder (1.2 g, 3.0 mmol, 75 %). TLC (100 % EtOAc) R_f : 0.76; ^1H -NMR (400 MHz, d_6 -DMSO): δ (ppm) 11.53 (s, 1H), 7.52 (d, J = 1.2 Hz, 1H), 6.13 (d, J = 8 Hz, 1H), 5.25–5.27 (m, 1H), 4.96–4.99 (m, 1H), 4.26–4.32 (m, 3H), 2.14 (s, 3H), 2.08 (s, 3H), 1.80 (s, 3H). ^{13}C NMR (100 MHz, d_6 -DMSO): δ (ppm) 170.7, 169.8, 164, 151.1, 135.8, 111.1, 88.7, 80.1, 71.7, 63.4, 47.6, 21, 12.6; HRMS: m/z calculated for $\text{C}_{14}\text{H}_{17}\text{N}_2\text{O}_7\text{Br}$ $[\text{M}+\text{H}]^+$ = 405.0292, found $[\text{M}+\text{H}]^+$ = 405.0298.

Synthesis of 5'-O-Acetyl-2', 3', didehydro-2', 3' – dideoxy-5-methyluridine (6b)

A solution of compound **6a** (3.5 g, 8.6 mmol), freshly activated Zn powder (3.5 g, 65 mmol), NH₄Cl (4.2 g, 65 mmol) in anhydrous Acetonitrile (25 mL) was heated at 50 °C for 45 min. reaction progress was monitored by TLC. After completion of the reaction, the reaction mixture was filtered through a celite pad and evaporated at reduced pressure. Further, the product was purified by column chromatography to get a solid white powder (1.1 g, 3.8 mmol, 45%). TLC (100 % ETAC) *R_f*: 0.60; ¹H-NMR (400 MHz, *d*₆-DMSO): δ (ppm) 11.38 (s, 1H), 7.26 (d, *J* = 1.2 Hz, 1H), 6.80–6.81(m, 1H), 6.40 (dt, *J* = 6 Hz, 2 Hz, 1H), 5.99– 6.02 (m, 1H), 4.95– 4.99 (m, 1H), 4.16 –4.25 (m, 1H), 2.02 (s, 3H), 1.77 (s, 3H). ¹³C NMR (100 MHz, *d*₆-DMSO): 170.6, 164.2, 151.2, 136.3, 134.1, 127.1, 110, 89.7, 84.1, 65, 21.1, 12.6. HRMS: *m/z* calculated for C₁₂H₁₄N₂O₅ [M-H]⁻=265.0829, found [M-H]⁻ = 265.0826.

Synthesis of 2',3', didehydro-2', 3'– dideoxy-5-methyluridine (6c)

Compound **6b** (250 mg, 0.94 mmol) and sodium methoxide (400 mg, 9.42 mmol) were dissolved in a solution of dry MeOH: DCM (1:1) (15 mL), and the mixture was stirred for 2.5 hours at room temperature. The progress of the reaction was monitored by TLC (5% MeOH: DCM). The reaction was quenched by adding Amberlite IRN-120 H⁺ resin, and the resulting mixture was filtered using a Whatman filter paper. The organic phase was then dried over Na₂SO₄ and evaporated to dryness. The compound was purified through column chromatography, yielding a white solid powder (143 mg, 75% yield). TLC (MeOH: DCM 5:95) *R_f*:0.44; ¹H NMR (400 MHz, *d*₆-DMSO): δ (ppm) 11.29 (s, 1H), 7.64 (d, *J* = 1.3 Hz, 1H), 6.82 (dt, *J* = 3.3, 1.7 Hz, 1H), 6.39 (dt, *J* = 6.0, 1.7 Hz, 1H), 5.91 (ddd, *J* = 6.0, 2.4, 1.4 Hz, 1H), 5.01 (t, *J* = 5.3 Hz, 1H), 4.84 – 4.66 (m, 1H), 3.60 (dd, *J* = 5.3, 3.4 Hz, 2H), 2.50 (s, 1H), 1.72 (s, 3H). ¹³C NMR (100 MHz, *d*₆-DMSO):164.3, 151.2, 137.2, 135.4, 126.4, 109.4, 89.3, 87.8, 62.7, 12.6. HRMS: *m/z* calculated for C₁₀H₁₂N₂O₄ [M+H]⁺=225.0870, found [M+H]⁺= 225.0877

Synthesis of 2', 3' –dideoxy-5-methyluridine (6d)

Compound **6c** (900 mg, 3.6 mmol) was dissolved in 15 mL of anhydrous methanol and subjected to hydrogenation in the presence of a catalyst (100 mg) (wet 10% palladium on carbon). The reaction mixture was stirred at room temperature for 2 hours. Subsequently, the catalyst was removed by filtration, and the filtrate was concentrated under reduced pressure. Compound **4** was purified by column chromatography (800 mg, 3.1 mmol, 87 %). TLC (100 % ETAC) *R_f*: 0.44; ¹H NMR (400 MHz, *d*₆-DMSO) δ 7.79 (s, 1H), 5.96 (dd, *J* = 6.4, 3.6 Hz, 1H), 4.00 (d, *J* = 12.9 Hz, 1H), 3.67 (d, *J* = 11.8 Hz, 1H), 3.53 (d, *J* = 8.3 Hz, 1H), 2.31 – 2.17 (m, 1H), 2.00–1.81 (m, 5H), 1.77 (s, 3H). ¹³C NMR (100 MHz, *d*₆-DMSO): 163.9,150.5,136.4,

108.8, 84.8, 81.3, 62.2, 31.6, 25, 12.3 HRMS: m/z calculated for $C_{10}H_{14}N_2O_4$ $[M+H]^+ = 227.1032$, found $[M+H]^+ = 227.1070$.

Synthesis of 2', 3' –dideoxy-5-methyluridine triphosphate (6)

The nucleotide analog **6** was synthesized by a similar procedure as mentioned earlier (Chapter 2, section 2.4.3) using compound **6d** (50 mg, 0.22 mmol, 1 equiv.). we obtained desired product **6** as a triethylammonium salt. (30 mg, 26%). ^1H -NMR (400 MHz, D_2O): δ (ppm) 7.60 (s, 1H), 5.96 (dd, $J = 6.4, 3.6$ Hz, 1H), 4.20 (br, 1H), 4.11– 4.09 (m, 1H), 3.99– 3.95 (m, 1H), 2.00– 1.94 (m, 2H), 1.76 (s, 3H), 1.74 (s, 2H); ^{13}C NMR (100 MHz, D_2O): 166.7, 137.7, 111.2, 86, 80.2, 66.9, 31, 24.9, 11.6; ^{31}P NMR (162 MHz, D_2O): δ (ppm) -10.94 (d, $J = 17.82$ Hz, P_α), -11.48 (d, $J = 21.06$ Hz, P_γ), -23.4 (d, $J = 17.82$ Hz, P_β); HRMS: m/z calculated for $C_{10}H_{16}N_2O_{13}P_3$ $[M-H]^- = 464.9871$, found $[M-H]^- = 464.9862$.

4.4.4 Solid-phase synthesis of BFdU-modified oligonucleotides

BFdU-modified phosphoramidite was synthesized following a previously reported procedure.⁵³ 3'-BFdU modified oligonucleotides (**P3*** and **P6***) were synthesized as mentioned in Chapter 3 (section 3.4.4). The purity and integrity of the ONs were assessed through reversed-phase high-performance liquid chromatography (RP-HPLC) and electrospray ionization mass spectrometry (ESI-MS), respectively (see appendix-III).

4.4.5 Expression and purification of TdT enzyme

Terminal deoxynucleotidyl transferase was cloned into an expression plasmid pET-MCN and transformed into BL21-DE3 *E. coli* competent cells. The suspension culture was grown overnight with Luria Bertani medium containing 50 μ g/ml of antibiotic Kanamycin sulfate at 37°C with 200 rpm shaking. The primary culture was diluted to O.D.600=0.2 in a 3 L secondary culture. The cells were grown at 37°C with 200 rpm shaking till the O.D.600 reached 0.4-0.5. The culture was incubated at room temperature (RT) for 45 minutes without shaking. The culture was pre-cooled for 45 minutes at 16°C with 200 rpm shaking. 1mM IPTG was added to the culture flasks and incubated overnight at 16°C with 200 rpm shaking. The cells were harvested after overnight incubation using a fixed-angle centrifuge at 11,000 rpm for 15 minutes. All steps involved in the protein purification were performed at 4°C. The pellets harvested were resuspended with lysis buffer (50 mM Tris–HCl pH 8, 500 mM NaCl, 5 mM imidazole) (buffer A). Followed by lysozyme (0.5 mg/mL) and PMSF (1mM) were added and the pellet suspension was kept in ice for 15 minutes and then sonicated. The lysate was centrifuged at 4°C, 15000rpm for 30 minutes and filtered using a 0.45 μ m filter. The lysate was subjected to Ni-NTA column purification (5mL His Trap HP, GE Healthcare). The protein was eluted with an elution buffer of pH 8 (50 mM Tris–HCl pH 8, 500 mM NaCl, and 500 mM

imidazole) (buffer B). Protein was eluted using a stepwise gradient of the elution buffer (5%, 10%, 15%, 20%, 25%, 30%, 50% and 100% of buffer B) with a flow rate of 5mL/min. The protein fractions showing purity were pooled and concentrated to 1ml volume using an Amicon Ultra10K filter and centrifugal concentrator (Membrane: 10,000 MWCO). The concentrated fractions were exchanged with storage buffer (100 mM NaCl, 200 mM K₂HPO₄, 5 mM DTT, pH 7.4). The samples were then loaded on an SEC 650 size-exclusion column (ENrich™ SEC 650 10 x 300 Column) pre-incubated with the storage buffer (100 mM NaCl, 200 mM K₂HPO₄, 5 mM DTT, pH 7.4) and purified using an isocratic method with a flow rate of 1mL/min. The fractions corresponding to the peak observed at 280 nm, were checked on 12% SDS PAGE gel. The purest fractions were collected and concentrated and stored in the storage buffer (100 mM NaCl, 200 mM K₂HPO₄, 5 mM DTT, pH 7.4) containing 50% glycerol at -20°C.

4.4.6 TdT reactions with natural and modified nucleotides (1–4)

Tailing reactions were performed in the presence of a 5'- FAM labeled 19-nt primer **P1**. Reactions mixtures (total volume 20 µL) were prepared in 1X TdT reaction buffer (50 mM KOAc, 20 mM Tris-acetate, 10 mM Mg(OAc)₂, pH 7.9, 0.25 mM CoCl₂) containing primer **P1** (2.5 µM), natural dTTP or modified nucleotides (**1–4**) (50 µM). Reactions were started with the addition of TdT enzyme (0.45 µM) and incubated at 37 °C. 5 µL of the reaction aliquot was collected after 15, 30, and 60 minutes and mixed with 55 µL of denaturing loading buffer (8.3 M urea in 10 mM Tris-HCl, 100 mM EDTA, 0.05% bromophenol blue, pH 8), followed by heat denatured at 75 °C for 5 minutes. 5 µL (1 pmol) of each sample was loaded onto an analytical 20% denaturing PAGE and electrophoresed at 18 W in 1X TBE buffer for 4 h. The gel was imaged using a Typhoon gel scanner at the FAM wavelength.

4.4.7 Tailing reactions in the presence of BFddUTP 5 and different initiators

A 5'-FAM labeled primer (**P2**, **P3**, **P4**, **P5**, **P6**) (2.5 µM) was incubated with 20 eq. of BFddUTP 5 (50 µM) in a TdT reaction buffer (50 mM KOAc, 20 mM Tris-acetate, 10 mM Mg(OAc)₂, pH 7.9) containing 0.25 mM of CoCl₂. Reactions were initiated by adding the TdT enzyme (0.30 µM) at 37 °C. After 1, 2, 5, and 10 minutes, 5 µL of the reaction aliquots were mixed with 55 µL of denaturing loading buffer and heated at 75 °C for 5 minutes. Reactions were analyzed by loading 5µL (1 pmol) of the samples to a 20% denaturing polyacrylamide gel and electrophoresing it for 18 W in 1 X TBE running buffer for 4 h. The gels were visualized by scanning at FAM wavelength using a Typhoon gel scanner. Gel bands were quantified by Fiji image analysis software. All the experiments were performed in triplicate to calculate the standard deviation and the average yield.

4.4.8 Competition reaction between BFddUTP 5 and ddTTP 6

20 µL of the reactions contained 2.5 µM of primer (**P3** or **P4**), 50 µM of ddNTP mixture of BFddUTP 5/ddTTP 6 in TdT reaction buffer (50 mM KOAc, 20 mM Tris-acetate, 10 mM Mg(OAc)₂, 0.25 mM CoCl₂, pH 7.9), and 0.30 µM of TdT. The ratio of ddTTP 6 to BFddUTP 5 varied from 0:1 to 1:0 (0/1, 1/50, 1/10, 1/5, 1/2, 1/1, 2/1, 5/1, 10/1, 50/1 and 1/0). Reactions were incubated at 37 °C for 20 minutes. Reactions were quenched by the addition of 180 µL of denaturing loading buffer and followed by heat-denatured at 75 °C for 5 minutes. Loaded 5 µL (1.25 pmol) of samples onto a 20% denaturing PAGE. Electrophoresed at 18 W for 4 h in 1X TBE buffer. The gel images at FAM emission wavelength were taken by a Typhoon gel scanner. The relative band intensity was calculated using Fiji image analysis software. All the experiments were repeated in triplicate to determine the standard deviation and average values of incorporation efficiency.

4.4.9 Steady-state fluorescence binding assay for BFdU-labeled DNA and TdT

By fluorescence:

A series of samples containing BFdU-labeled primer **P6*** (50 nM) were mixed with increasing concentrations of TdT enzyme (0– 2 µM) in 1 X TdT reaction buffer in a final volume of 200 µL maintaining an overall ~1% glycerol. The samples were kept at 37°C for 30 minutes. Subsequently, steady-state fluorescence analysis was conducted by exciting each sample at a 330 nm wavelength. The fluorescence readings were recorded within the 345 to 600 nm wavelength range, with excitation and emission slit widths maintained at 7 nm and 8 nm, respectively. The control sample lacking **P6*** exhibited no fluorescence. The fluorescence experiments were conducted in triplicate, and the standard deviation was calculated. The increase in fluorescence intensity at 430 nm for BFdU-labeled primer **P6*** against TdT concentration was analyzed using the Hill equation (Origin 8.5). The data, depicting normalized fluorescence intensity (F_N) against the logarithm of TdT concentration, revealed a Hill coefficient (n) of 1.59 for the binding of BFdU-labeled ssDNA **P6***.

$$F_N = \frac{F_i - F_0}{F_S - F_0}$$

F_i represents the measured fluorescence intensity at individual protein concentrations. F_0 , F_S = fluorescence intensity in the absence of protein and at the saturation point, respectively. The Hill coefficient (n) signifies the level of cooperativity in binding, while K_d represents the dissociation constant.

$$F_N = F_0 + (F_S - F_0) \left(\frac{[\text{Polymerase}]^n}{[K_d]^n + [\text{Polymerase}]^n} \right)$$

By Microscale thermophoresis (MST) assay:

To investigate the binary complex formation between the primers **P6/P3** and TdT was determined by means of a Monolith NT.115 system (NanoTemper Technologies) using standard capillaries (MonolithTM NT.115 Standard Treated Capillaries). Samples containing (10 μ L) a primers **P6/P3** (0.5 μ M) and varying concentrations of TdT (0.001–5 μ M) in a buffer (50 mM Tris–HCl, pH 8.0, 9% of glycerol and 1 mM CoCl₂) were incubated at 37°C. MST was recorded at 37°C and the dissociation constants were calculated by fitting it into the single-stage binding model using the DynaFit software (BioKin).

4.4.10 Detection of TdT activity in real-time by fluorescence

The real-time detection of TdT activity was performed using a micro-fluorescence cuvette (Hellma, path length 1.0 cm) on a Horiba Fluoromax-4 spectrophotometer. In the reaction mixtures (total volume 200 μ L) containing primer **P4*** (200 nM) and BFddUTP **5** (100 nM) in 1 X TdT reaction buffer, TdT was added at varying concentrations (0–200 nM). The measurements were started after the addition of TdT and pre-incubation for 30 seconds. The conversion of the BFddUTP **5** was detected by measuring emission at 447 nm (excitation wavelength 330 nm, excitation and emission slit width 7 and 8 nm, respectively) every 10 seconds for 10 minutes at 37°C. **P7** (3'-ddC modified) was used as a primer for the control experiment. After 10 minutes of reaction time, corresponding fluorescence spectra were measured. Each set of experiments was performed in triplicate. The data was processed and plotted using Origin 8.5. To calculate the rate constant (k_{obs}), the initial points till 150 seconds were taken and fitted into a pseudo-first-order kinetics equation.

4.4.11 Competition assay with ddTTP by real-time fluorescence

Real-time competition assay was conducted in similar conditions to those used to monitor the TdT activity. For this purpose, in a mixture of primer **P4*** (200 nM), BFddUTP **5** (100 nM), and varying concentrations of ddTTP **6** (0–5 μ M) in 1 X TdT reaction buffer, 100 nM of TdT was added. The data was recorded after the addition of TdT and pre-incubation for 30 seconds. The conversion of BFddUTP **5** was detected by monitoring emission at 447 nm (excitation wavelength 330 nm, excitation and emission slit width 7 and 8 nm, respectively) every 10 seconds over a 10-minute period. The reactions (200 μ L) temperatures were maintained at 37°C during the measurement. For control experiments, **P7** (3'-ddC modified) served as a primer, and fluorescence spectra were measured after 10 minutes of reaction time. Each set of experiments was conducted in triplicate, and the data were processed and plotted using Origin 8.5.

4.5 References

1. Bollum, F. J. Calf thymus polymerase. *J. Biol. Chem.* **1960**, 235, 2399–2403.
2. Fowler, J. D.; Suo, Z. Biochemical, structural, and physiological characterization of Terminal deoxynucleotidyl transferase. *Chem. Rev.* **2006**, 106, 2092–2110.
3. Moteaa, E. A.; Berdis, A. J. Terminal deoxynucleotidyl transferase: the story of a misguided DNA polymerase. *Biochim Biophys Acta.* **2010**, 1804, 1151–1166.
4. Keller, S.; Marx, A. The use of enzymes for construction of DNA-based objects and assemblies *Chem. Soc. Rev.* **2011**, 40, 5690–5697.
5. Xiang, B.; He, K.; Zhu, R.; Liu, Z.; Zeng, S.; Huang, Y.; Nie, Z.; Yao, S. Self-Assembled DNA hydrogel based on enzymatically polymerized DNA for protein encapsulation and enzyme/DNAzyme hybrid cascade reaction. *ACS Appl. Mater. Interfaces.* **2016**, 8, 22801–22807.
6. Yang, Y.; Lu, Q.; Huang, C. M.; Qian, H.; Zhang, Y.; Deshpande, S.; Arya, G.; Ke, Y.; Zauscher, S. Programmable site-specific functionalization of DNA origami with polynucleotide brushes *Angew. Chem. Int. Ed.* **2021**, 60, 23241–23247.
7. Hao, Y.; Li, Q.; Fan, C.; Wang, F. Data Storage Based on DNA. *Small Structures.* **2021**, 2, 2000046.
8. Lim, C. K.; Nirantar, S.; Yew, W. S.; Poh, C. L. Novel modalities in DNA data storage. *Trends Biotechnol.* **2021**, 39, 990–1003.
9. Lee, H. H.; Kalhor, R.; Goela, N.; Bolot, J.; Church, G. M. Terminator-free template-independent enzymatic DNA synthesis for digital information storage. *Nat. Commun.* **2019**, 10, 1–12.
10. Lee, H.; Wiegand, D. J.; Griswold, K.; Punthambaker, S.; Chun, H.; Kohman, R. E.; Church, G. M. Photon-directed multiplexed enzymatic DNA synthesis for molecular digital data storage. *Nat. Commun.* **2020**, 11, 1–9.
11. Barthel, S.; Palluk, S.; Hillson, N. J.; Keasling, J. D.; Arlow, D. H. Enhancing Terminal deoxynucleotidyl transferase activity on substrates with 3' terminal structures for enzymatic De Novo DNA synthesis. *Genes* **2020**, 11, 102.
12. Schaudy, E.; Lietard, J.; Somoza, M. M. Sequence preference and initiator promiscuity for De Novo DNA synthesis by Terminal deoxynucleotidyl transferase. *ACS Synth. Biol.* **2021**, 10, 1750–1760.
13. Simonova, A.; Havran, L.; Pohl, R.; Fojta, M.; Hocek, M. Phenothiazine-linked nucleosides and nucleotides for redox labelling of DNA. *Org. Biomol. Chem.* **2017**, 15, 6984–6996.
14. Horáková, P.; Macíková-Cahová, H.; Pivoková, H.; Paek, J.; Havran, L.; Hocek, M.; Fojta, M. Tail-labelling of DNA probes using modified deoxynucleotide triphosphates and terminal deoxynucleotidyl transferase. Application in electrochemical DNA hybridization and protein-DNA binding assays. *Org. Biomol. Chem.* **2011**, 9, 1366–1371.
15. Adamopoulos, P. G.; Tsiakanikas, P.; Stolidi, I.; Scorilas, A. A versatile 5' RACE-Seq methodology for the accurate identification of the 5' termini of mRNAs. *BMC Genomics* **2022**, 23, 1–14.
16. Kyrylkova, K.; Kyryachenko, S.; Leid, M.; Kiousi, C. Detection of apoptosis by TUNEL assay. *Methods Mol. Biol.* **2012**, 887, 41–47.
17. Du, Y. C.; Wang, S. Y.; Wang, Y. X.; Ma, J. Y.; Wang, D. X.; Tang, A. N.; Kong, D. M. Terminal deoxynucleotidyl transferase combined CRISPR-Cas12a amplification strategy for ultrasensitive detection of uracil-DNA glycosylase with zero background. *Biosens. Bioelectron.* **2021**, 171, 112734.

18. Chen, X.; Cao, G.; Wang, X.; Ji, Z.; Xu, F.; Huo, D.; Luo, X.; Hou, C. Terminal deoxynucleotidyl transferase induced activators to unlock the trans-cleavage of CRISPR/Cpf 1 (TdT-IU-CRISPR/Cpf 1): An ultrasensitive biosensor for Dam MTase activity detection. *Biosens. Bioelectron.* **2020**, *163*, 11227.
19. Du, Y. C.; Wang, S. Y.; Li, X. Y.; Wang, Y. X.; Tang, A. N.; Kong, D. M. Label-free detection of miRNA cancer markers based on terminal deoxynucleotidyl transferase-induced copper nanoclusters. *Biosens. Bioelectron.* **2019**, *145*, 1117.
20. Deng, Y.; Peng, Y.; Wang, L.; Wang, M.; Zhou, T.; Xiang, L.; Li, J.; Yang, J.; Li, G. Target-triggered cascade signal amplification for sensitive electrochemical detection of SARS-CoV-2 with clinical application. *Anal. Chim. Acta* **2022**, *1208*, 339846.
21. Peng, Y.; Pan, Y.; Sun, Z.; Li, J.; Yi, Y.; Yang, J.; Li, G. An electrochemical biosensor for sensitive analysis of the SARS-CoV-2 RNA. *Biosens. Bioelectron.* **2021**, *186*, 113309.
22. Abnous, K.; Danesh, N. M.; Nameghi, M. A.; Ramezani, M.; Alibolandi, M.; Lavaee, P.; Taghdisi, S. M. *Biosens. Bioelectron.* **2019**, *144*, 111674.
23. Ashley, J.; Shahbazi, M. A.; Kant, K.; Chidambara, V. A.; Wolff, A.; Bang, D. D.; Sun, Y. Molecularly imprinted polymers for sample preparation and biosensing in food analysis: Progress and perspectives. *Biosens. Bioelectron.* **2017**, *91*, 606–615.
24. Vacacela, J.; Schaap-Johansen, A. L.; Manikova, P.; Marcatili, P.; Prado, M.; Sun, Y.; Ashley, J. Templated synthesis of enzyme-generated aptamers. *Angew. Chem. Int. Ed.* **2022**, *61*, e202201061.
25. Klingler, C.; Ashley, J.; Shi, K.; Stiefvater, A.; Kyba, M.; Sinnreich, M.; Aihara, H.; Kinter, J. *FASEB J.* **2020**, *34*, 4573–4590.
26. Benedict, C. L.; Gilfillan, S.; Thai, T. H.; Kearney, J. F. Terminal deoxynucleotidyl transferase and repertoire development. *Immunol. Rev.* **2000**, *175*, 150–157.
27. Baltimore, D. Is terminal deoxynucleotidyl transferase a somatic mutagen in lymphocytes? *Nature* **1974**, *248*, 409–411.
28. Desiderio, S. V.; Yancopoulos, G. D.; Paskind, M.; Thomas, E.; Boss, M. A.; Landau, N.; Alt, F. W.; Baltimore, D. Insertion of N regions into heavy-chain genes is correlated with expression of terminal deoxytransferase in B cells. *Nature* **1984**, *311*, 752–755.
29. Greaves, M.; Paxton, A.; Janossy, G.; Pain, C.; Johnson, S.; Lister, T. A. Acute lymphoblastic leukaemia associated antigen. III Alterations in expression during treatment and in relapse. *Leuk. Res.* **1980**, *4*, 1–14.
30. Hoffbrand, A. V.; Ganeshaguru, K.; Janossy, G.; Greaves, M. F.; Catovsky, D.; Woodruff, R. K. Terminal deoxynucleotidyl-transferase levels and membrane phenotypes in diagnosis of acute leukaemia. *Lancet* **1977**, *310*, 520–523.
31. Kung, P. C.; Long, J. C.; McCaffrey, R. P.; Ratliff, R. L.; Harrison, T. A.; Baltimore, D. Terminal deoxynucleotidyl transferase in the diagnosis of leukemia and malignant lymphoma. *Am. J. Med.* **1978**, *64*, 788–794.
32. Billum, F. Terminal deoxynucleotidyl transferase as a hematopoietic cell marker. *Blood*, **1979**, *54*, 1203–1215.
33. Kranz, B. R.; Thierfelder, S. Improved detection of terminal transferase (TdT): The use of detergents on glutaraldehyde-fixed non-dehydrated cells prevents denaturation and diffusion artifacts. *Leuk. Res.* **1986**, *10*, 1041–1049.
34. McCaffrey, R.; Lillquist, A.; Sallan, S.; Cohen, E.; Osband, M. Clinical utility of leukemia cell terminal transferase measurements. *Cancer Res.* **1981**, *41*, 4814–4820.
35. Kung, P. C.; Gottlieb, P. D.; Baltimore, D. Terminal deoxynucleotidyl transferase. serological studies and radioimmunoassay. *J. Biol. Chem.* **1976**, *251*, 2399–2404.

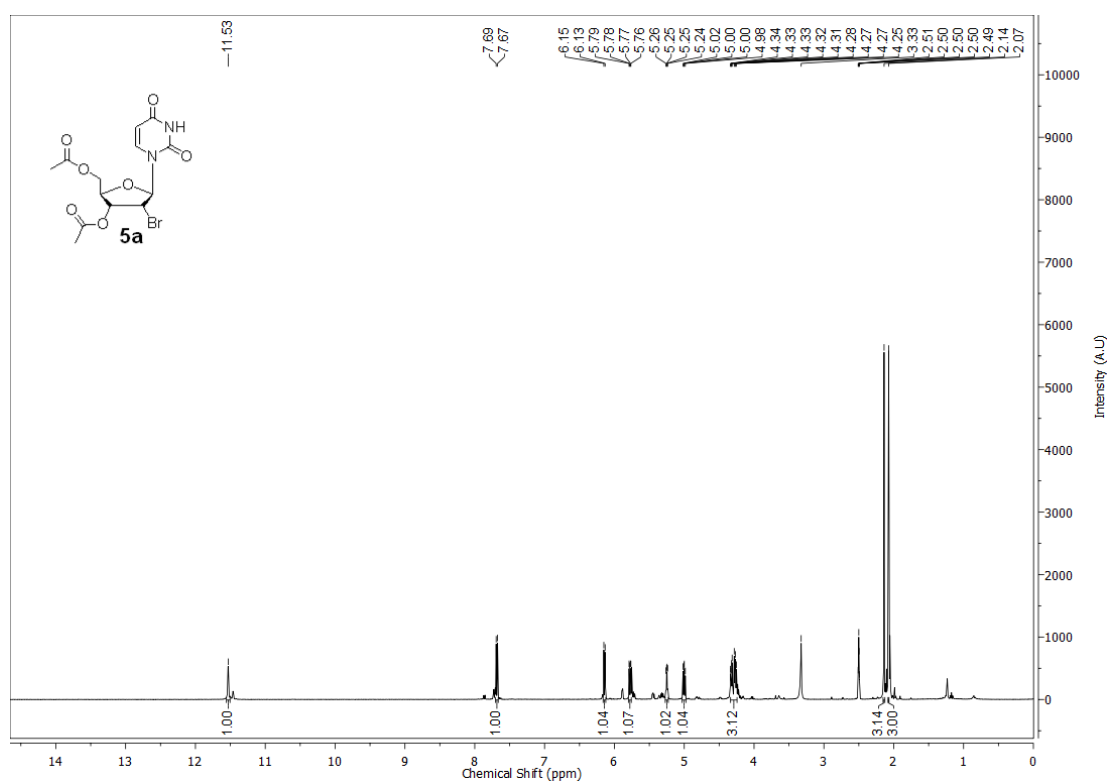
36. Mathewson, R. C.; Kjeldsberg, C. R.; Perkins, S. L. Detection of Terminal deoxynucleotidyl transferase (TdT) in nonhematopoietic small round cell tumors of children. *Fetal Pediatr. Pathol.* **1997**, *17*, 835–844.
37. Mertelsmann, R.; Mertelsmann, I.; Koziner, B.; Moore, M. A.; Clarkson, B. D. Improved biochemical assay for terminal deoxynucleotidyl transferase in human blood cells: results in 89 adult patients with lymphoid leukemias and malignant lymphomas in leukemic phase *Leuk. Res.* **1978**, *2*, 57–69.
38. Hu, Y.; Zhang, Q.; Xu, L.; Wang, J.; Rao, J.; Guo, Z.; Wang, S. Signal-on electrochemical assay for label-free detection of TdT and BamHI activity based on grown DNA nanowire-templated copper nanoclusters. *Anal. Bioanal. Chem.* **2017**, *409*, 6677–6688.
39. Hu, Y.; Zhang, Q.; Guo, Z.; Wang, S.; Du, C.; Zhai, C. In situ grown DNA nanotail-templated silver nanoclusters enabling label-free electrochemical sensing of terminal deoxynucleotidyl transferase activity. *Biosens. Bioelectron.* **2017**, *98*, 91–99.
40. Zhou, F.; Cui, X.; Shang, A.; Lian, J.; Yang, L.; Jin, Y.; Li, B. Fluorometric determination of the activity and inhibition of terminal deoxynucleotidyl transferase via in-situ formation of copper nanoclusters using enzymatically generated DNA as template. *Microchim. Acta* **2017**, *184*, 773–779.
41. Peng, F.; Liu, Z.; Li, W.; Huang, Y.; Nie, Z.; Yao, S. Enzymatically generated long polyT-templated copper nanoparticles for versatile biosensing assay of DNA-related enzyme activity. *Anal. Methods* **2015**, *7*, 4355–4361.
42. Liu, Z.; Li, W.; Nie, Z.; Peng, F.; Huang, Y.; Yao, S. Randomly arrayed G-quadruplexes for label-free and real-time assay of enzyme activity. *Chem. Commun.* **2014**, *50*, 6875–6878.
43. Lu, L.; Wang, M.; Liu, L. J.; Wong, C. Y.; Leung, C. H.; Ma, D. L. A luminescence switch-on probe for terminal deoxynucleotidyl transferase (TdT) activity detection by using an iridium(III)-based i-motif probe. *Chem. Commun.* **2015**, *51*, 9953–9956.
44. Wang, L. J.; Luo, M. L.; Zhang, Q.; Tang, B.; Zhang, C. Y. Single quantum dot-based nanosensor for rapid and sensitive detection of terminal deoxynucleotidyl transferase. *Chem. Commun.* **2017**, *53*, 11016–11019.
45. Liu, X.; Wang, H.; Deng, K.; Kwee, S.; Huang, H.; Tang, L. Single Primer Based Multisite Strand Displacement Reaction Amplification Strategy for Rapid Detection of Terminal deoxynucleotidyl transferase Activity. *Anal. Chem.* **2019**, *91*, 7482–7486.
46. A quenched size-expanded nucleotide reports activity of the leukemia biomarker Terminal deoxynucleotidyl transferase (TdT) *Angew. Chem. Int. Ed.* **2023**, *62*, e202302796.
47. Tanpure, A. A.; Srivatsan, S. G. A microenvironment-sensitive fluorescent pyrimidine ribonucleoside analogue: synthesis, enzymatic incorporation, and fluorescence detection of a DNA abasic site. *Chem. Eur. J.* **2011**, *17*, 12820–12827.
48. Dziuba, D.; Pospíšil, P.; Matyašovský, J.; Brynda, J.; Nachtigallová, D.; Rulíšek, L.; Pohl, R.; Hof, M.; Hocek, M. Solvatochromic fluorene-linked nucleoside and DNA as color-changing fluorescent probes for sensing interactions. *Chem. Sci.* **2016**, *7*, 5775–5785.
49. Li, Y.; Fin, A.; McCoy, L.; Tor, Y. Polymerase-mediated site-specific incorporation of a synthetic fluorescent isomorphous G surrogate into RNA. *Angew. Chem. Int. Ed.* **2017**, *56*, 1303–1307.
50. Sandin, P.; Stengel, G.; Ljungdahl, T.; Börjesson, K.; Macao, B.; Wilhelmsson, L. M. Highly efficient incorporation of the fluorescent nucleotide analogs tC and tC^O by Klenow fragment. *Nucleic Acids Res.* **2009**, *37*, 3924–3933.

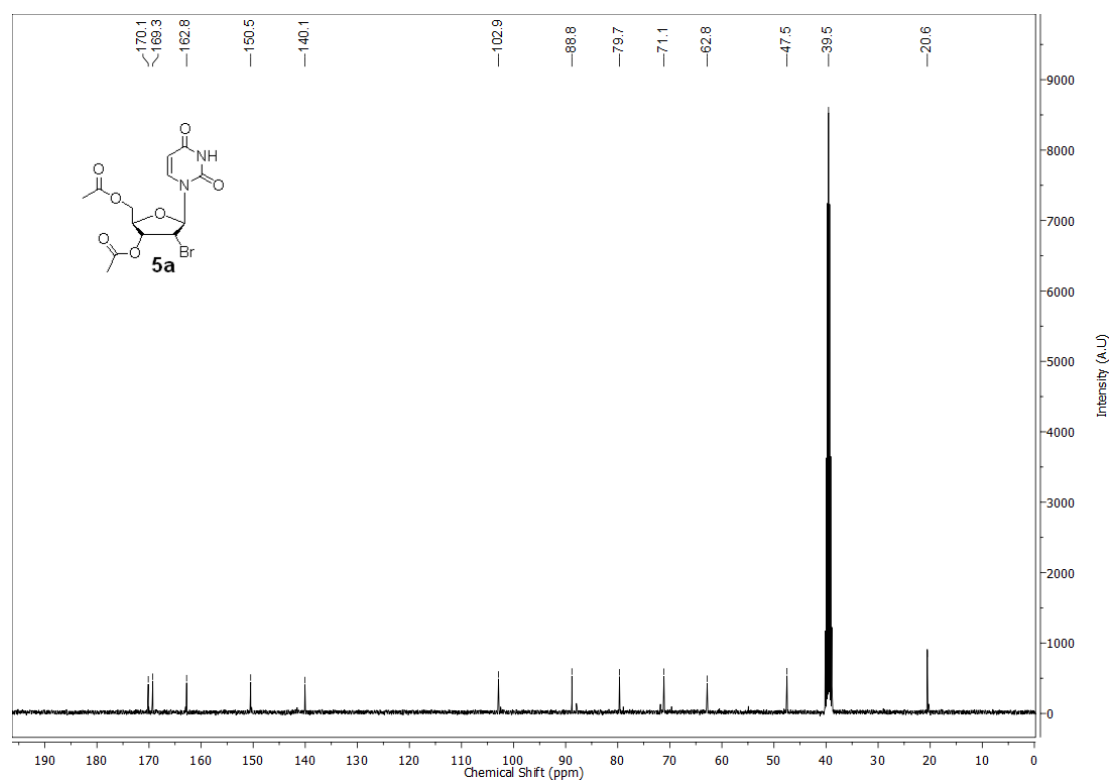
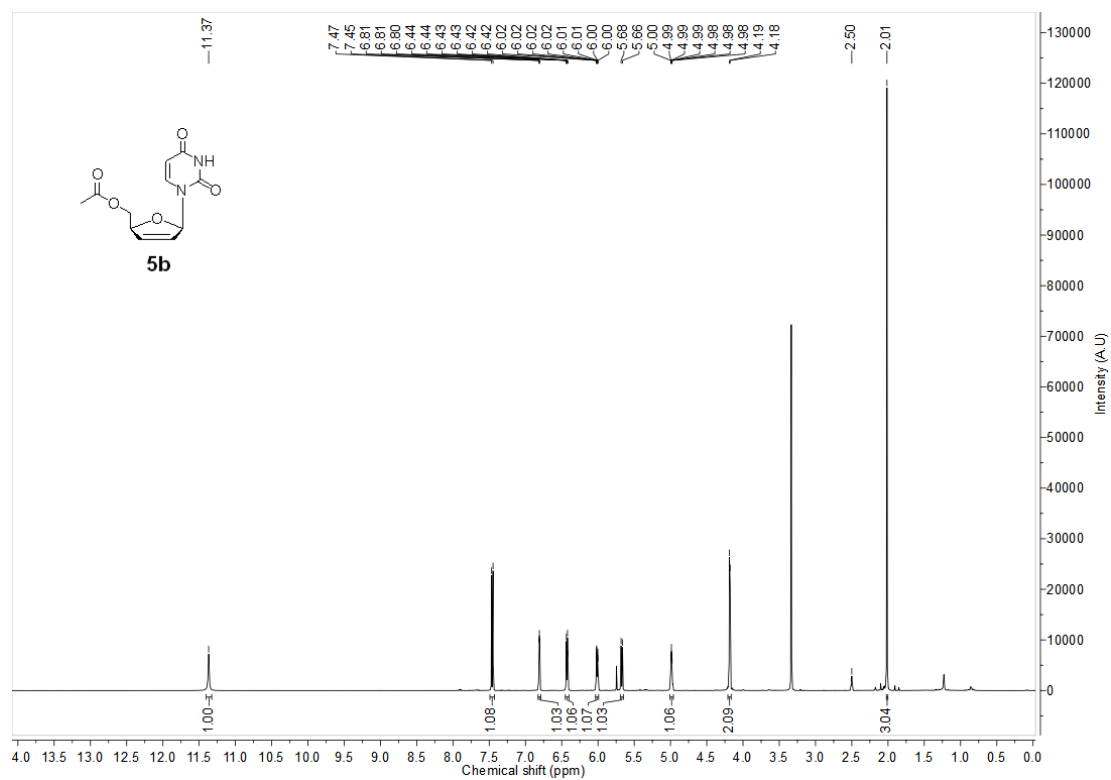
51. Nuthanakanti, A.; Ahmed, I.; Khatik, S. Y.; Saikrishnan, K.; Srivatsan S. G. Probing G-quadruplex topologies and recognition concurrently in real time and 3D using a dual-app nucleoside probe. *Nucleic Acids Res.* **2019**, *47*, 6059–6072.
52. Nuthanakanti, A.; Boerneke, M. A.; Hermann, T.; Srivatsan S. G. Structure of the ribosomal decoding site RNA containing a selenium-modified responsive fluorescent ribonucleoside probe. *Angew. Chem. Int. Ed.* **2017**, *56*, 2640–2644.
53. Tanpure, A. A.; Srivatsan, S. G. Conformation-sensitive nucleoside analogues as topology-specific fluorescence turn-on probes for DNA and RNA G-quadruplexes. *Nucleic Acids Res.* **2015**, *43*, e149.
54. Manna, S.; Sarkar, D.; Srivatsan, S. G. A dual-app nucleoside probe provides structural insights into the human telomeric overhang in live cells. *J. Am. Chem. Soc.* **2018**, *140*, 12622–12633.
55. Hansen, J. Characterization of rainbow trout terminal deoxynucleotidyl transferase structure and expression. TdT and RAG1 co-expression define the trout primary lymphoid tissues. *Immunogenetics* **1997**, *46*, 367–375.
56. Lee, A.; Hsu, E. Isolation and characterization of the *Xenopus* terminal deoxynucleotidyl transferase. *J Immunol.* **1994**, *152*, 4500–4507.
57. Yang, B.; Gathy, K. N.; Coleman, M. S. T-cell specific avian TdT: characterization of the cDNA and recombinant enzyme. *Nucleic Acids Res.* **1995**, *23*, 2041–2048.
58. Koiwai, O.; Yokota, T.; Kageyama, T.; Hirose, T.; Yoshida, S.; Arai, K. Isolation and characterization of bovine and mouse terminal deoxynucleotidyltransferase cDNAs expressible in mammalian cells. *Nucleic Acids Res.* **1986**, *14*, 5777–5792.
59. Peterson, R. C.; Cheung, L. C.; Mattaliano, R. J.; Chang, L. M.; Bollum, F. J. Molecular cloning of human terminal deoxynucleotidyl transferase. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 4363–4367.
60. Delarue, M.; Boule, J. B.; Lescar, J.; Expert-Bezancon, N.; Jourdan, N.; Sukumar, N.; Rougeon, F.; Papanicolaou, C. Crystal structures of a template-independent DNA polymerase: murine terminal deoxynucleotidyl transferase. *EMBO J.* **2002**, *21*, 427–439.
61. Bartl, S.; Miracle, A. L.; Rumfelt, L. L.; Kepler, T. B.; Mochon, E.; Litman, G. W.; Flajnik, M. F. Terminal deoxynucleotidyl transferases from elasmobranchs reveal structural conservation within vertebrates. *Immunogenetics* **2003**, *55*, 594–604.
62. Kuznetsova A. A.; Tyugashev T. E.; Alekseeva I. V. Timofeyeva N.A.; Fedorova O.S.; Kuznetsov N. A. Insight into the mechanism of DNA synthesis by human terminal deoxynucleotidyltransferase. *Life Sci Alliance.* **2022**, *5*, e202201428.
63. Gouge, J.; Rosario, S.; Romain, F.; Beguin, P.; Delarue, M. Structures of Intermediates along the Catalytic Cycle of Terminal Deoxynucleotidyltransferase: Dynamical Aspects of the Two-Metal Ion Mechanism. *J Mol Biol*, **2013**, *425*, 4334–4352.
64. Palluk, S.; Arlow, D. H.; de Rond, T.; Barthel, S.; Kang, J. S.; Bector, R.; Baghdassarian, H. M.; Truong, A. N.; Kim, P. W.; Singh, A. K.; Hillson, N. J.; Keasling, J. D. De Novo DNA Synthesis Using Polymerase-Nucleotide Conjugates. *Nat. Biotechnol.* **2018**, *36*, 645–650.
65. Rachofsky, E. L.; Osman, R.; Ross, J. B. Probing structure and dynamics of DNA with 2-aminopurine: effects of local environment on fluorescence. *Biochemistry.* **2001**, *40*, 946–956.
66. Lu, X.; Li, J.; Li, C.; Lou, Q.; Peng, K.; Cai, B.; Liu, Y.; Yao, Y.; Lu, L.; Tian, Z.; et al. enzymatic DNA synthesis by engineering terminal deoxynucleotidyl transferase. *ACS Catal.* **2022**, *12*, 2988–2997.
67. Haidekker, M. A.; Theodorakis, E. A. Molecular rotors—fluorescent biosensors for viscosity and flow. *Org. Biomol. Chem.* **2007**, *5*, 1669–1678.

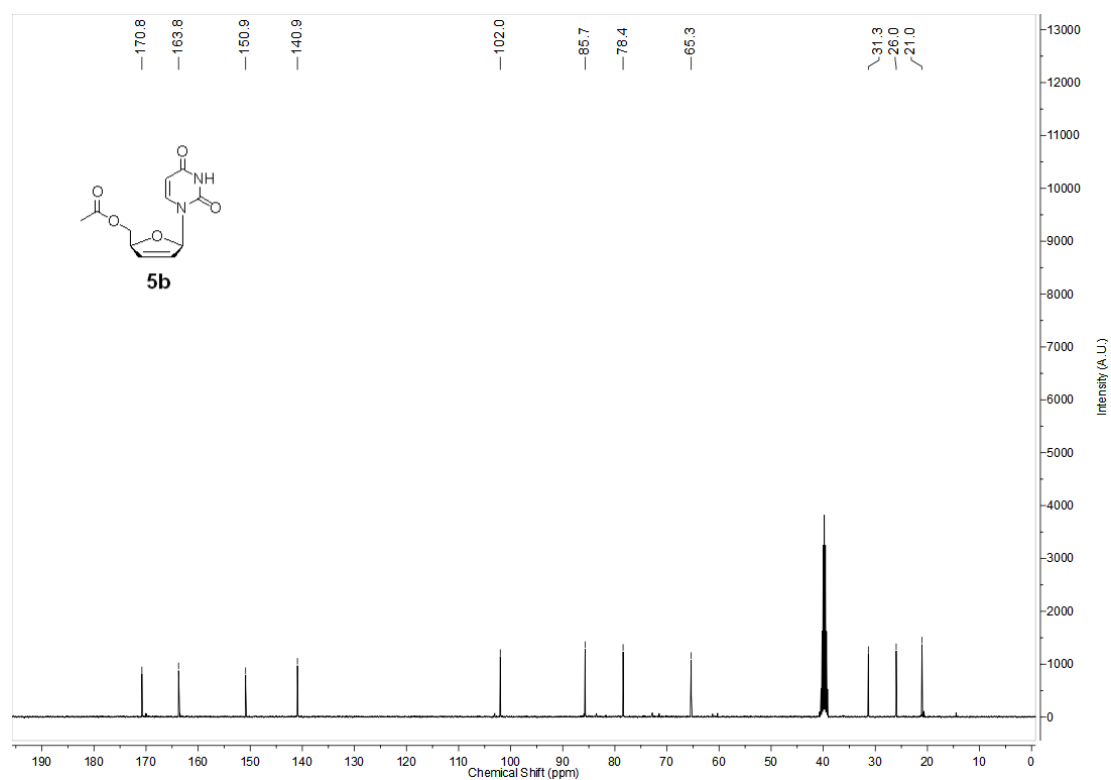
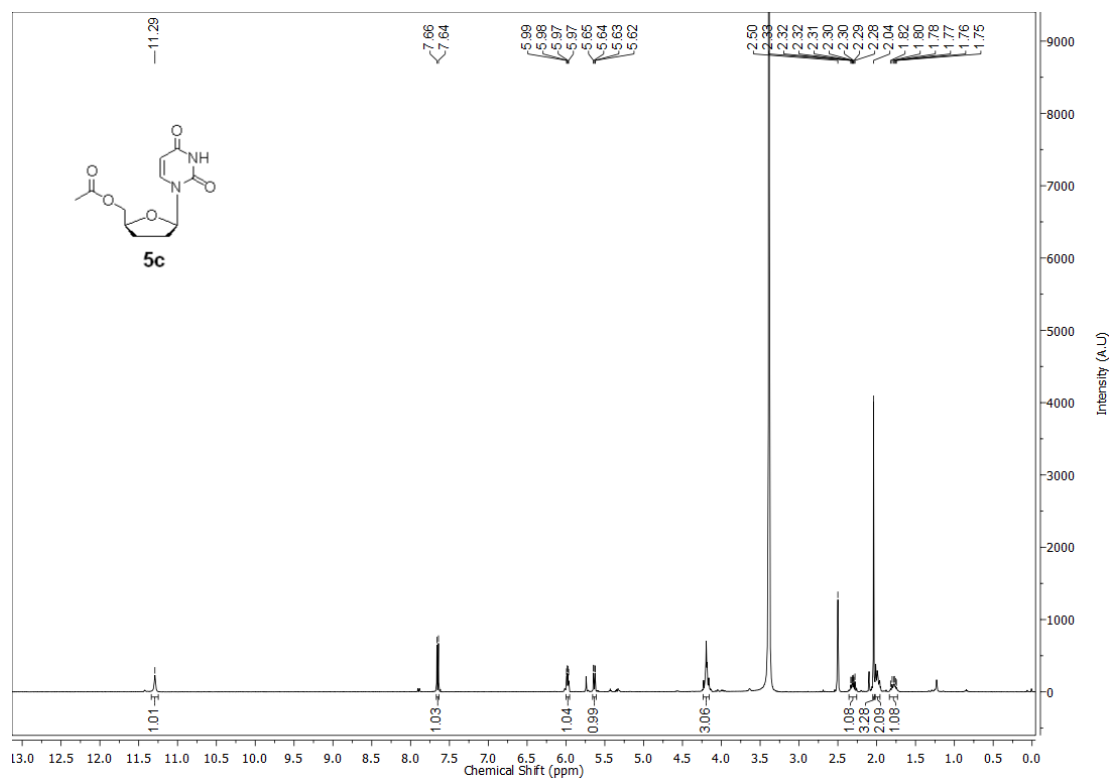
68. Förster, T.; Hoffmann, G. Die viskositätsabhängigkeit der fluoreszenz quantena usbeuten einiger farbstoffsysteme. *Z. Phys. Chem.* **1971**, *75*, 63–76.
69. Sinkeldam, R. W.; Wheat, A. J.; Boyaci, H.; Tor, Y. Emissive nucleosides as molecular rotors. *ChemPhysChem* **2011**, *12*, 567–570.
70. Noé, M. S.; Sinkeldam, R. W.; Tor, Y. Oligodeoxynucleotides containing multiple thiophene-modified isomorphpic fluorescent nucleosides. *J. Org. Chem.* **2013**, *78*, 8123–8128.

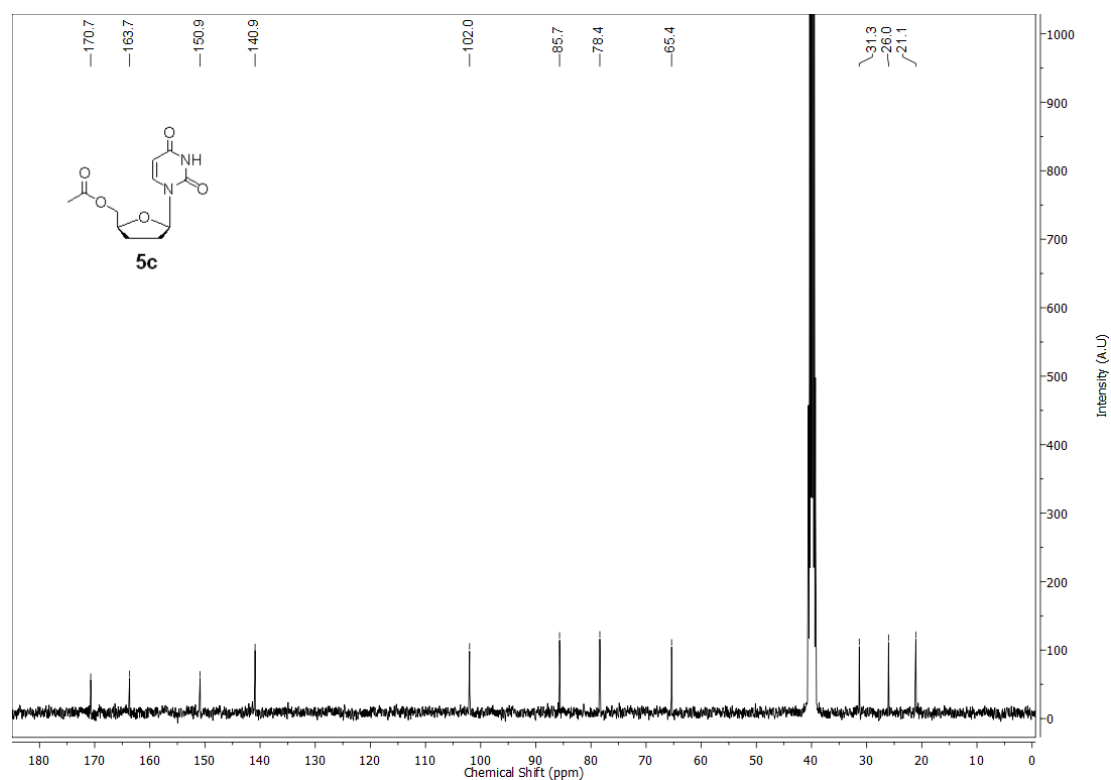
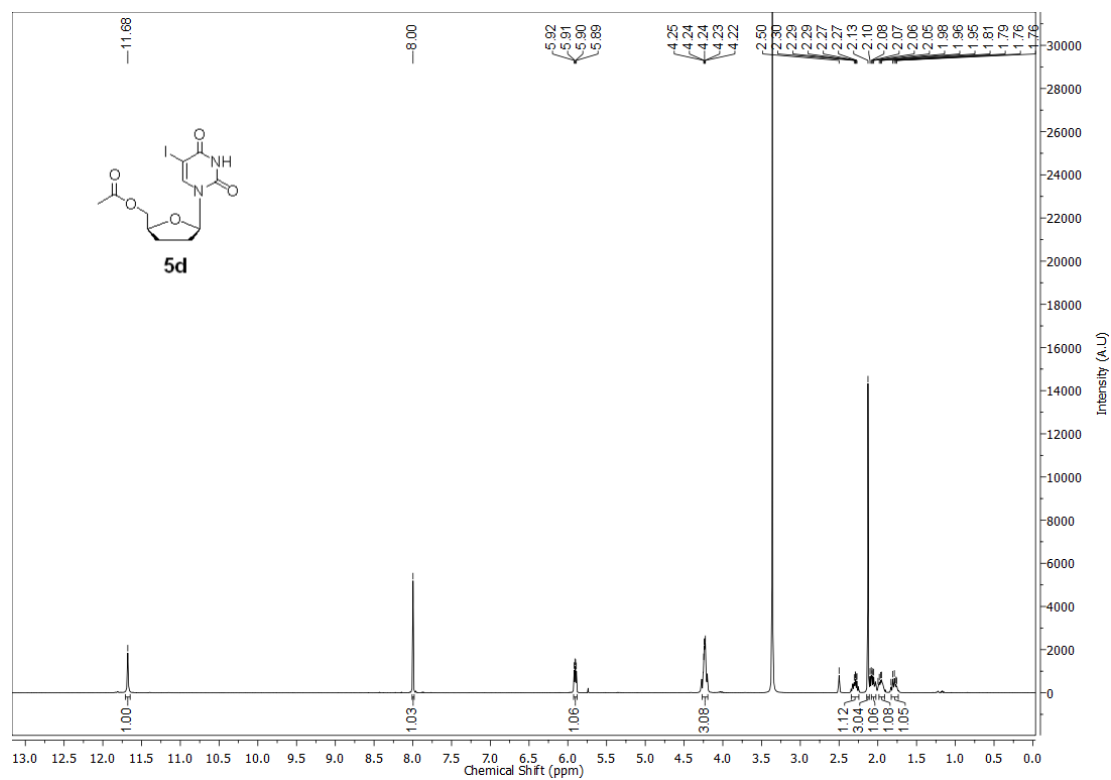
4.6 Appendix-III: NMR and mass spectra of synthesized compounds

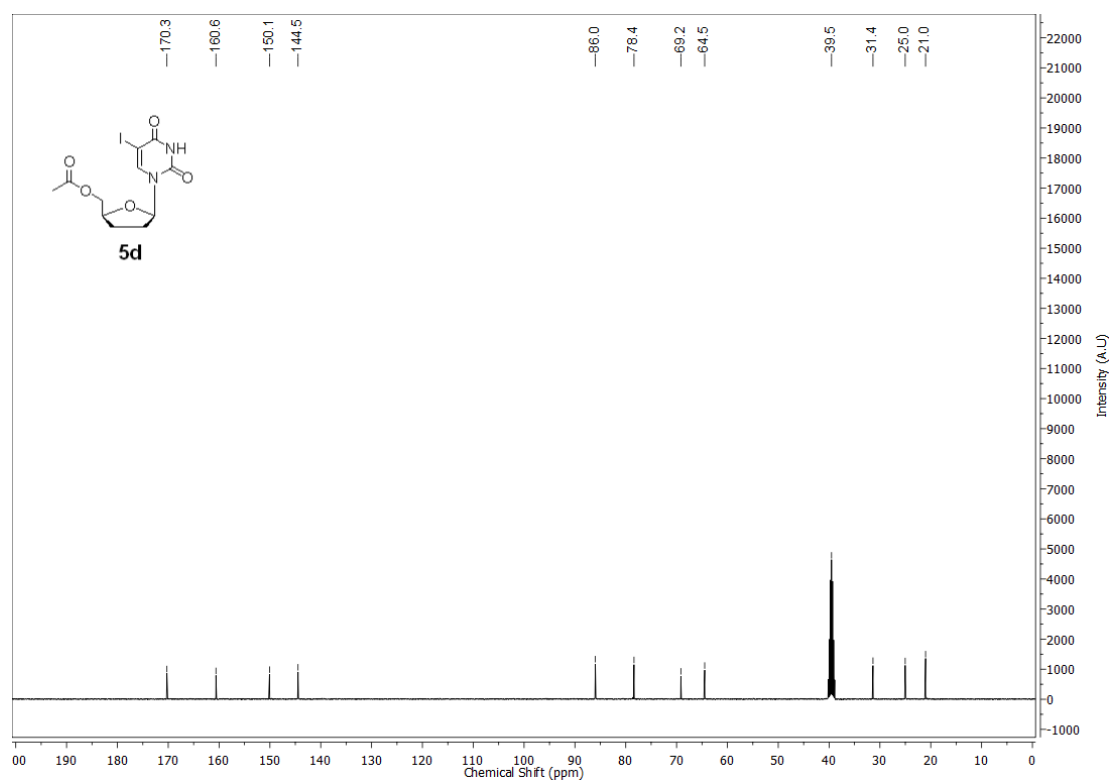
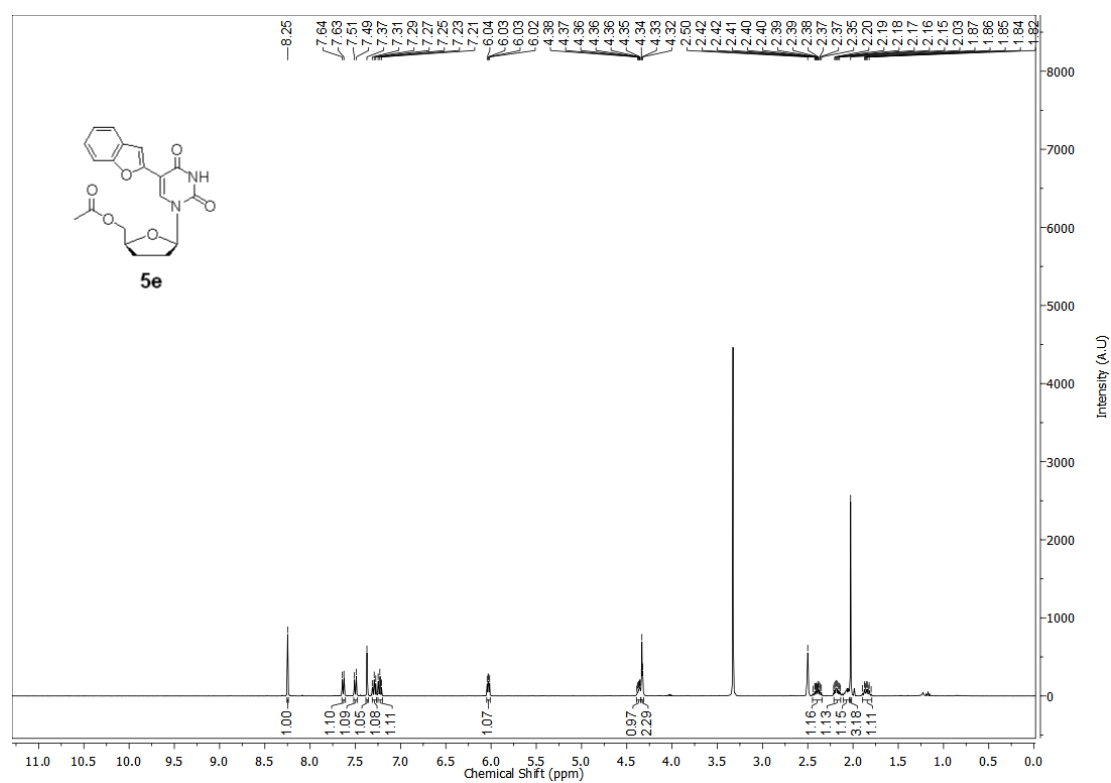
^1H -NMR of **5a** (400 MHz, *d*₆-DMSO)

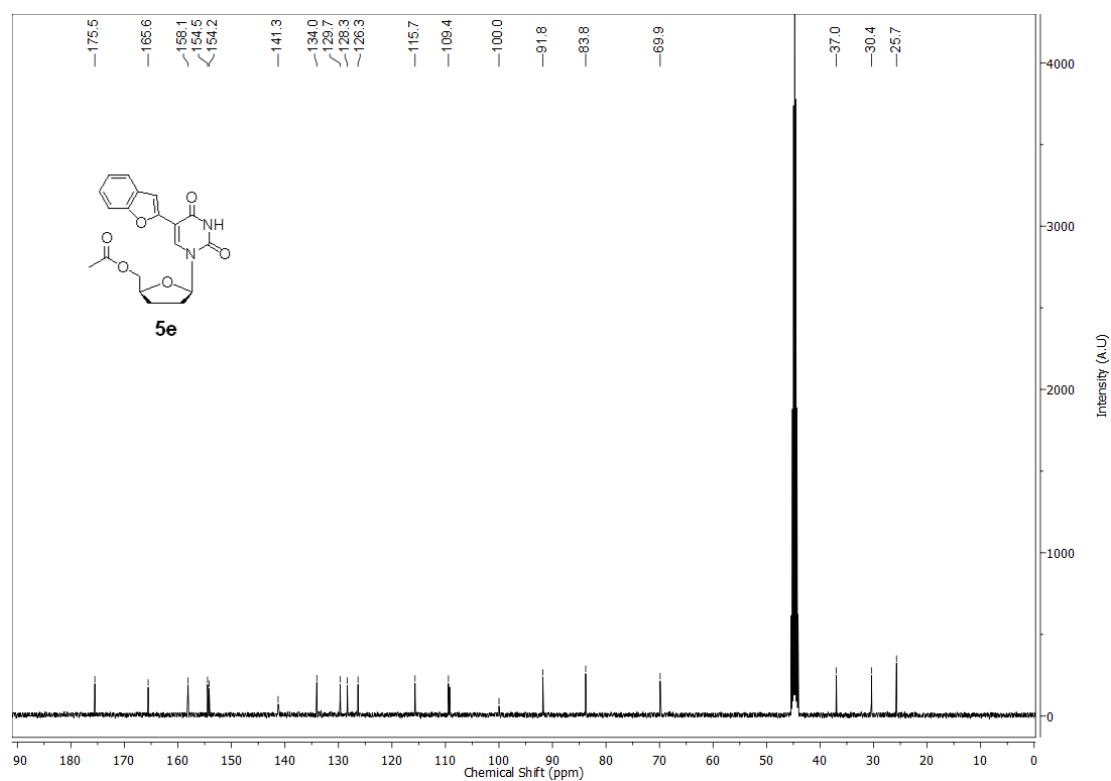
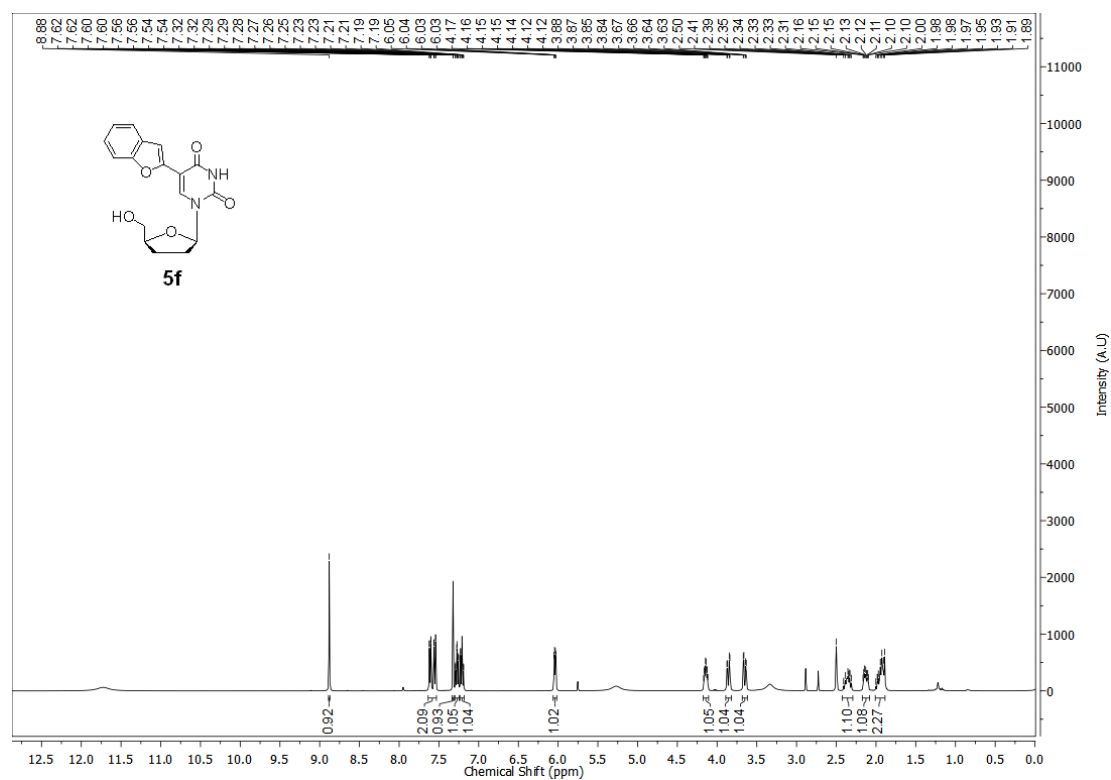


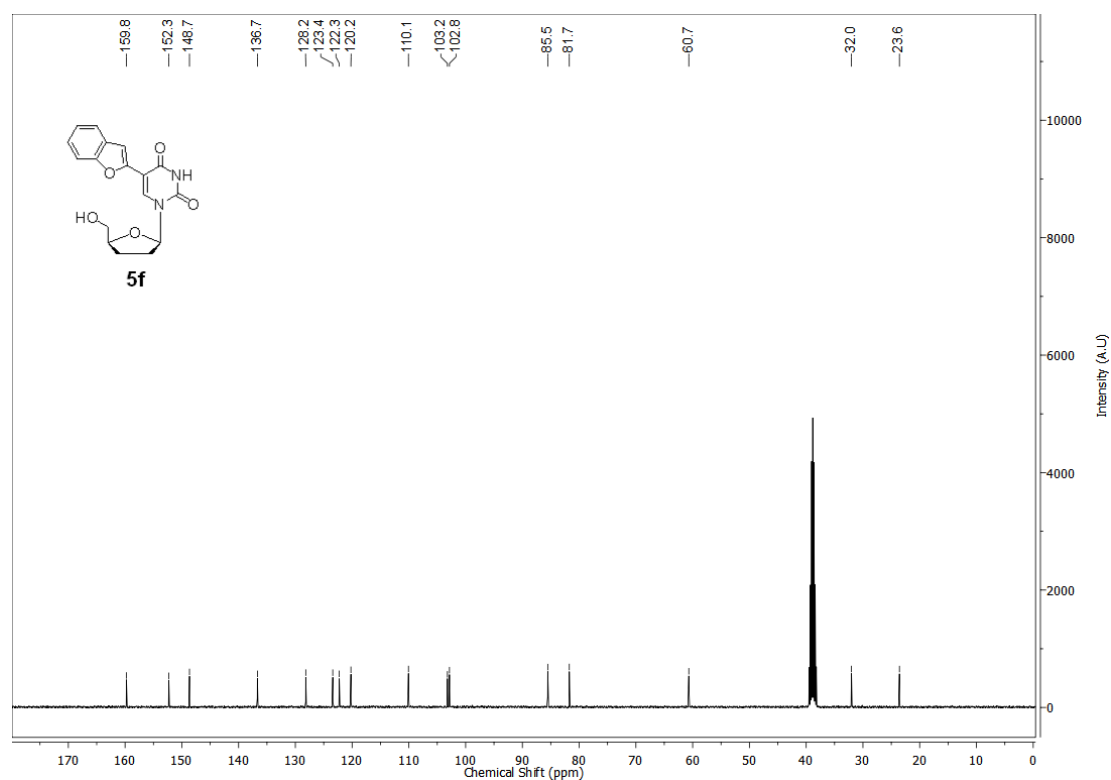
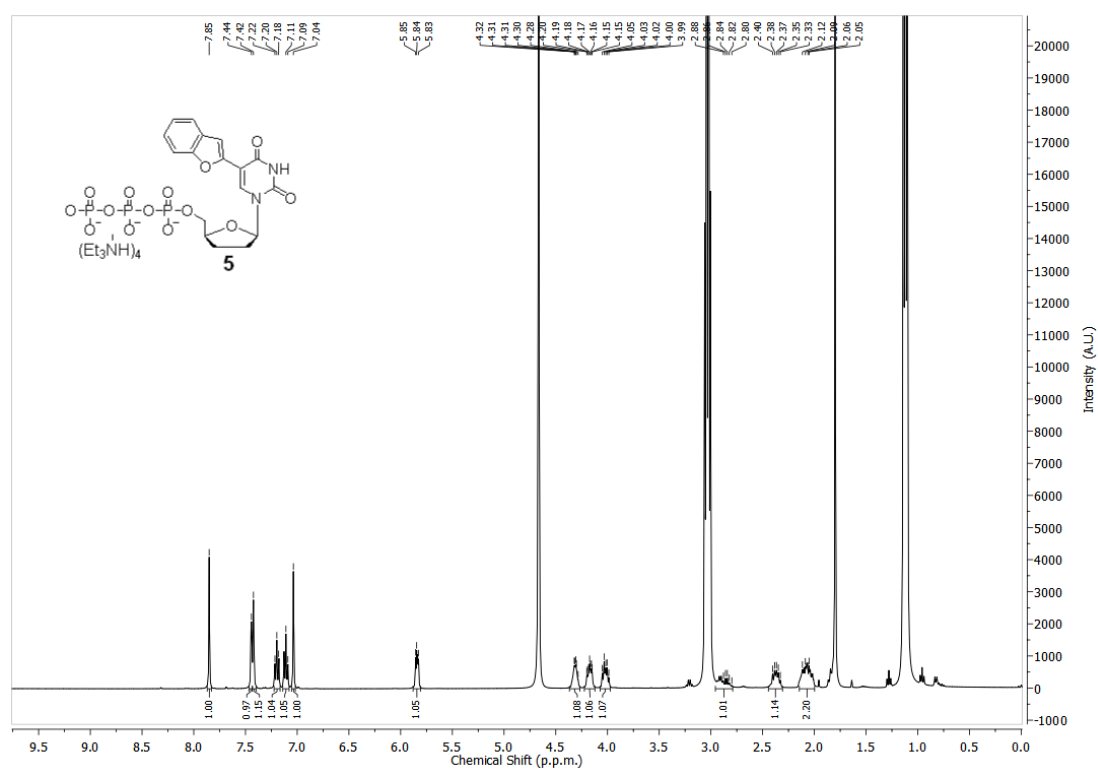
^{13}C -NMR of **5a** (100 MHz, d_6 -DMSO) ^1H -NMR of **5b** (400 MHz, d_6 -DMSO)

^{13}C -NMR of **5b** (100 MHz, d_6 -DMSO) ^1H -NMR of **5c** (400 MHz, d_6 -DMSO)

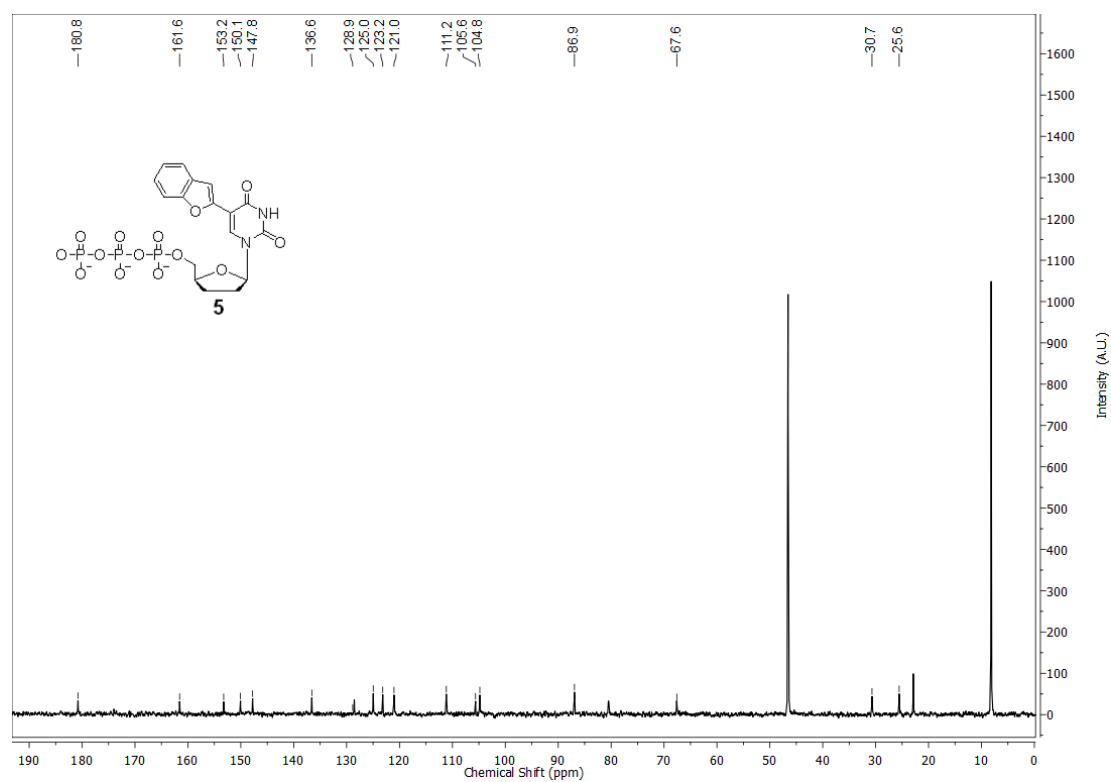
^{13}C -NMR of **5c** (100 MHz, d_6 -DMSO) ^1H -NMR of **5d** (400 MHz, d_6 -DMSO)

^{13}C -NMR of **5d** (100 MHz, d_6 -DMSO) ^1H -NMR of **5e** (400 MHz, d_6 -DMSO)

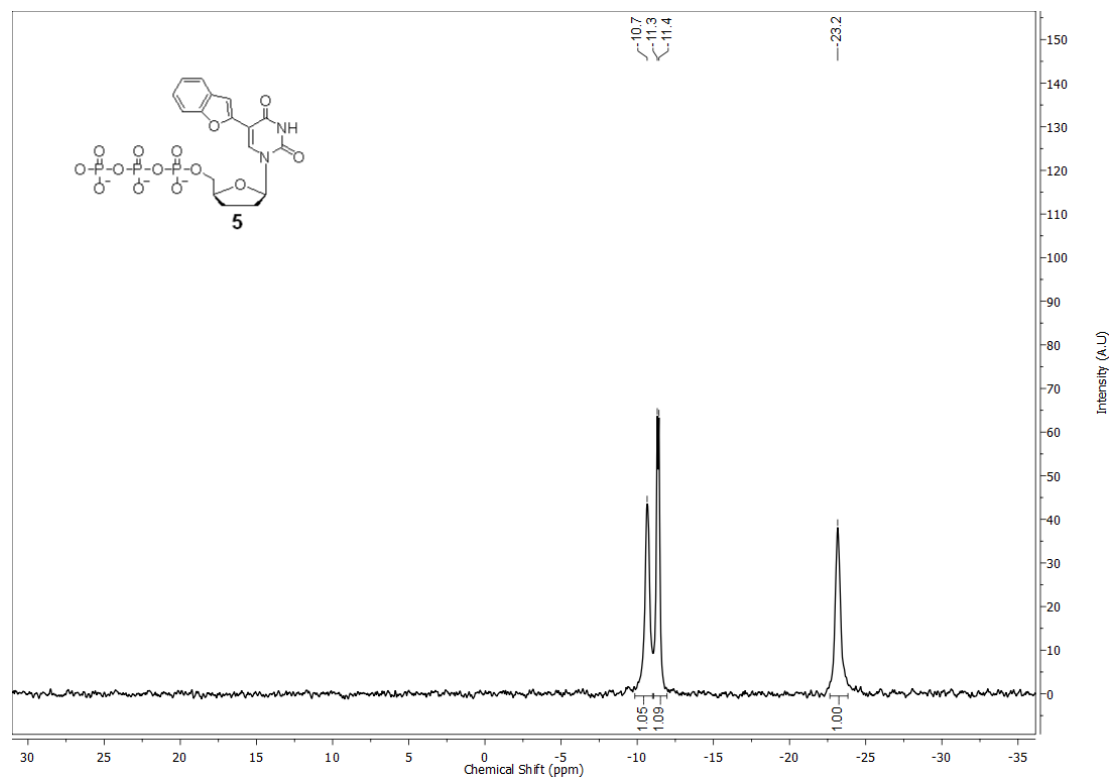
^{13}C -NMR of **5e** (100 MHz, *d*₆-DMSO) ^1H -NMR of **5f** (400 MHz, *d*₆-DMSO)

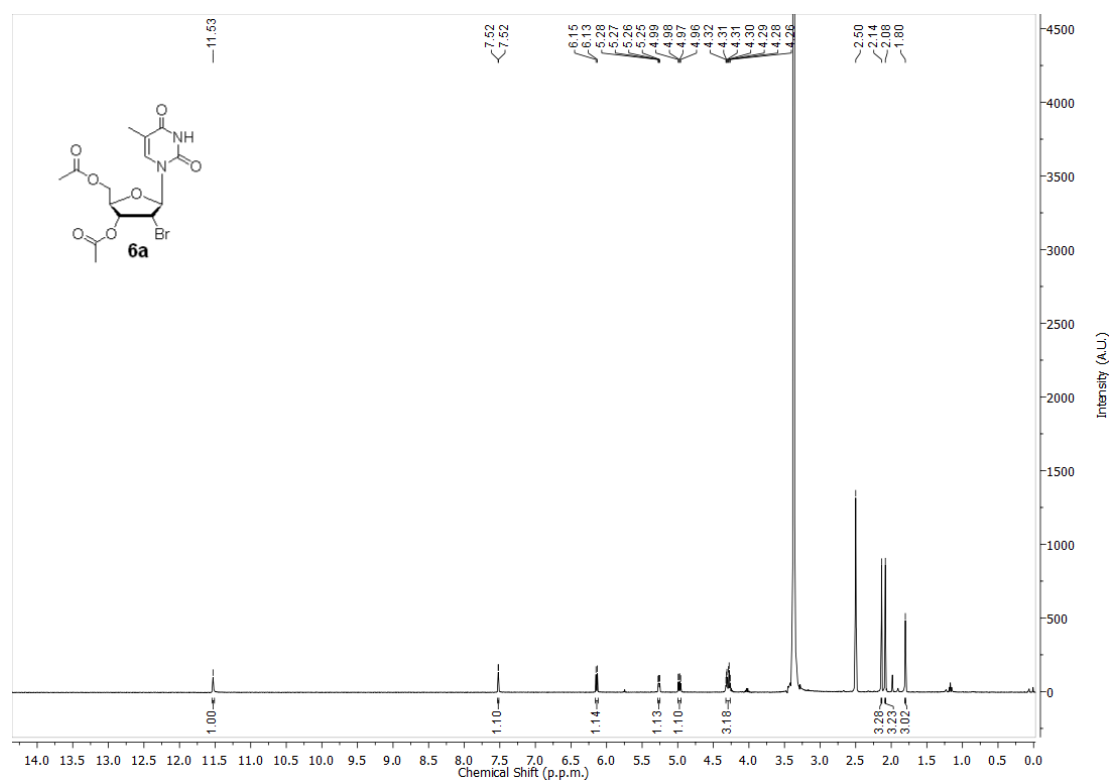
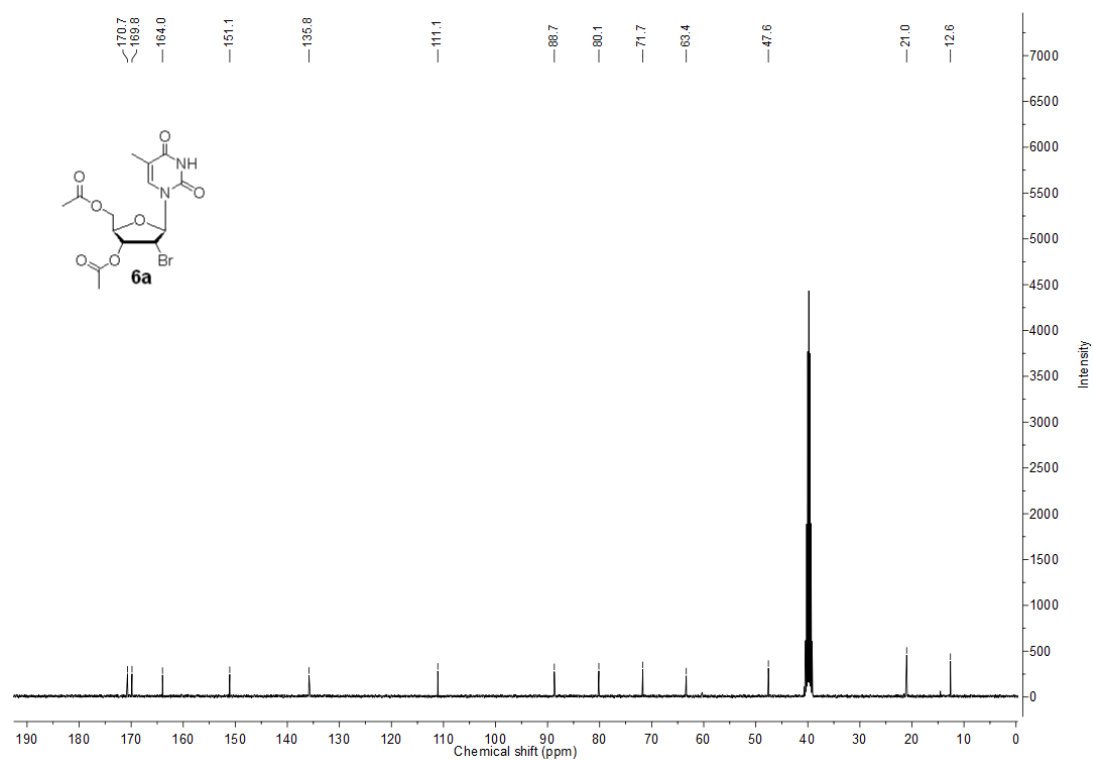
^{13}C -NMR of **5f** (100 MHz, d_6 -DMSO) ^1H -NMR of **5** (400 MHz, D_2O), a trace amount of triethylammonium acetate salt is present.

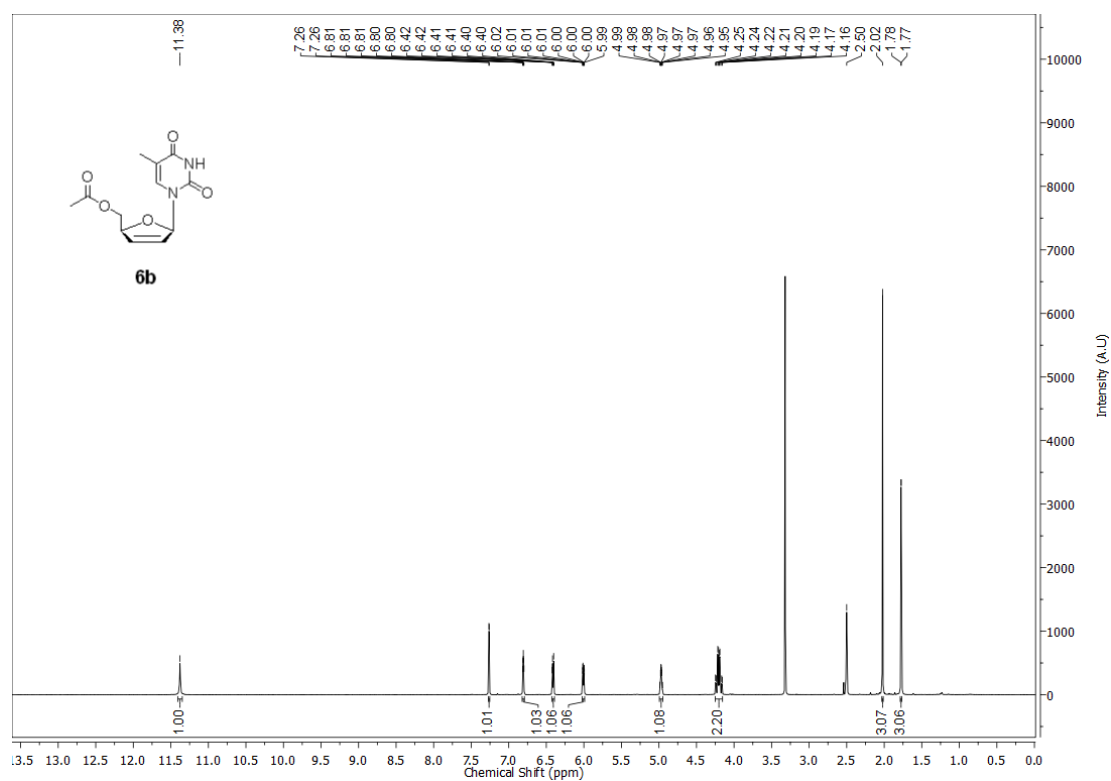
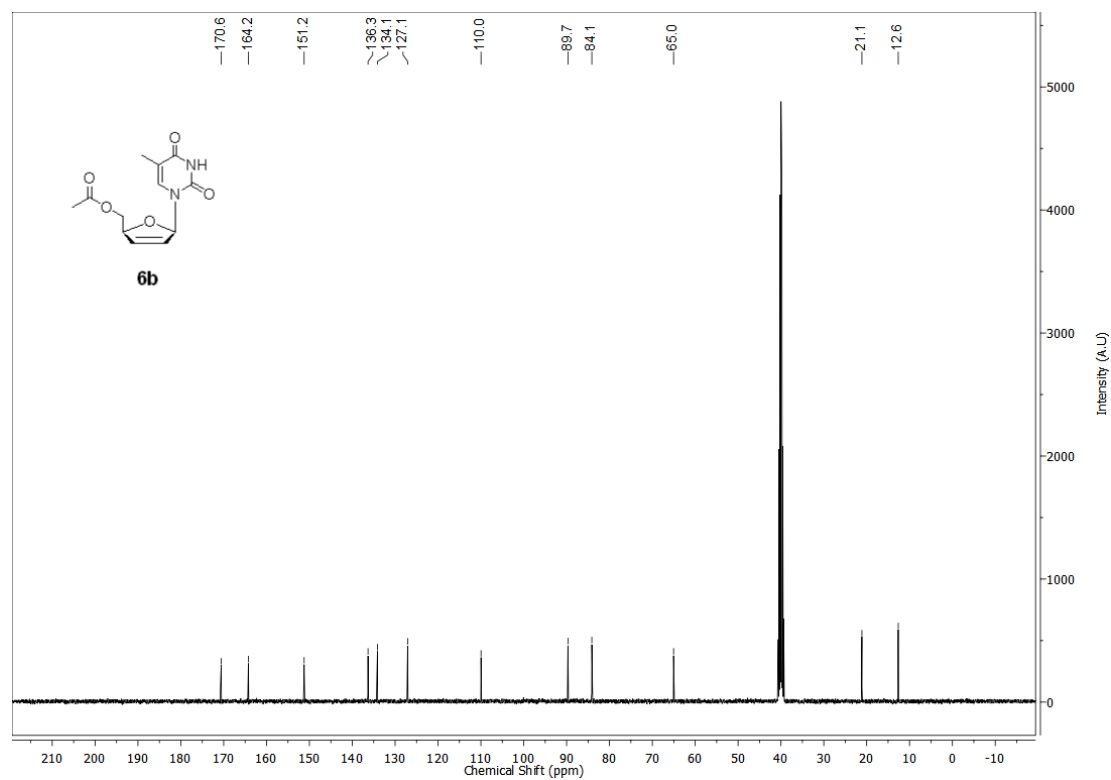
^{13}C -NMR of **5** (100 MHz, D_2O), a trace amount of triethylammonium acetate salt is present.

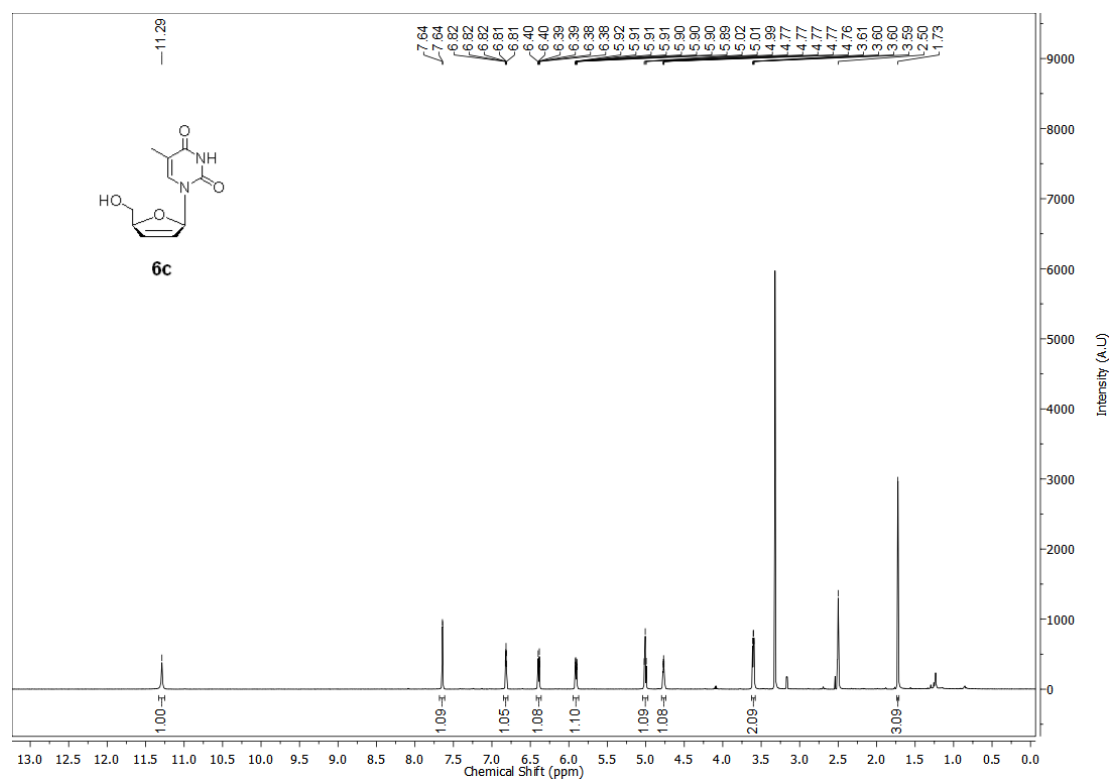
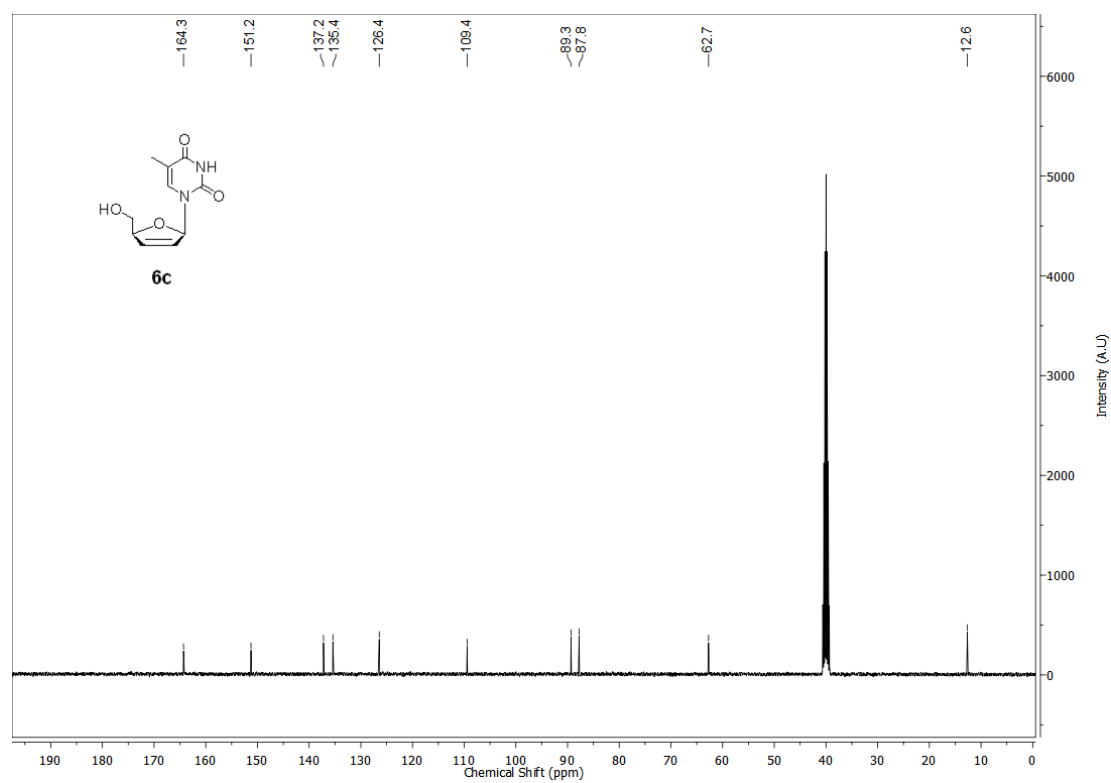


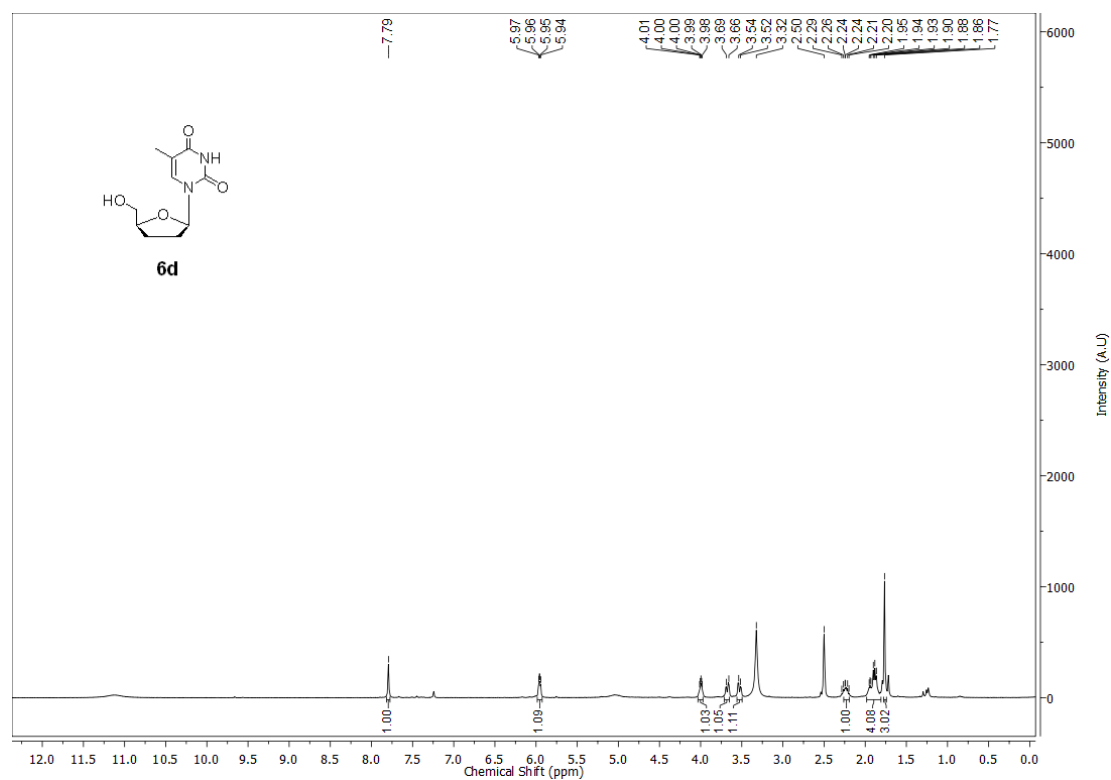
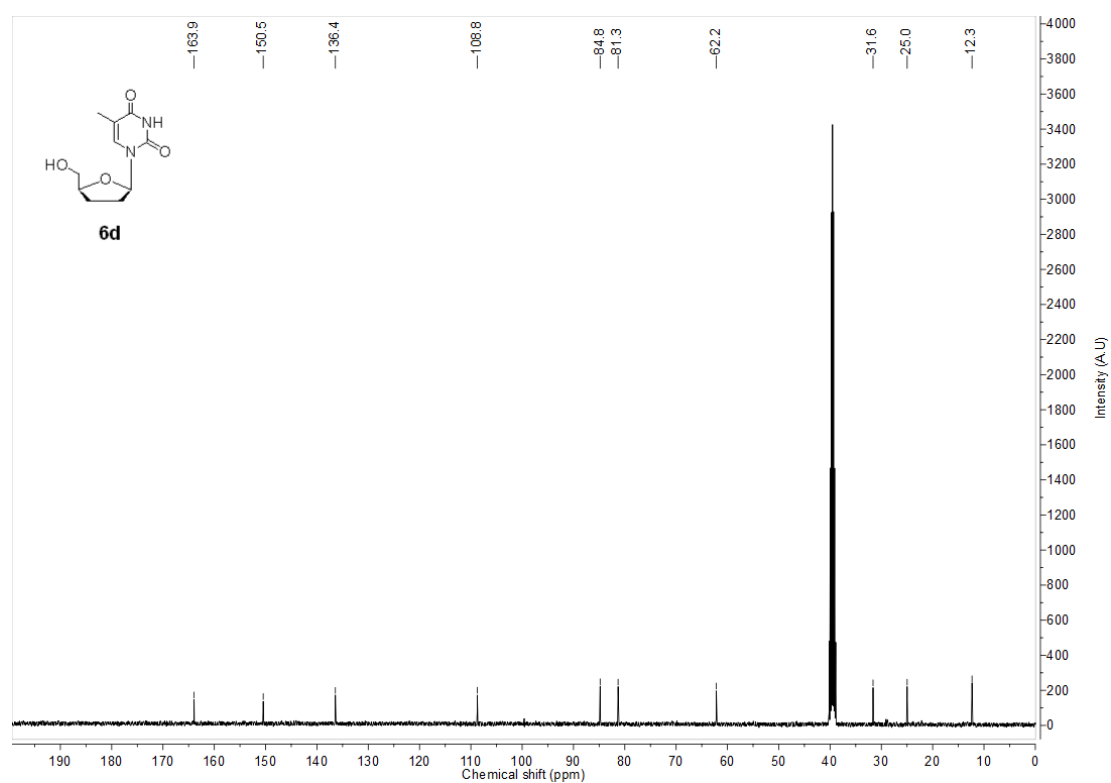
^{31}P -NMR of **5** (162 MHz, D_2O)



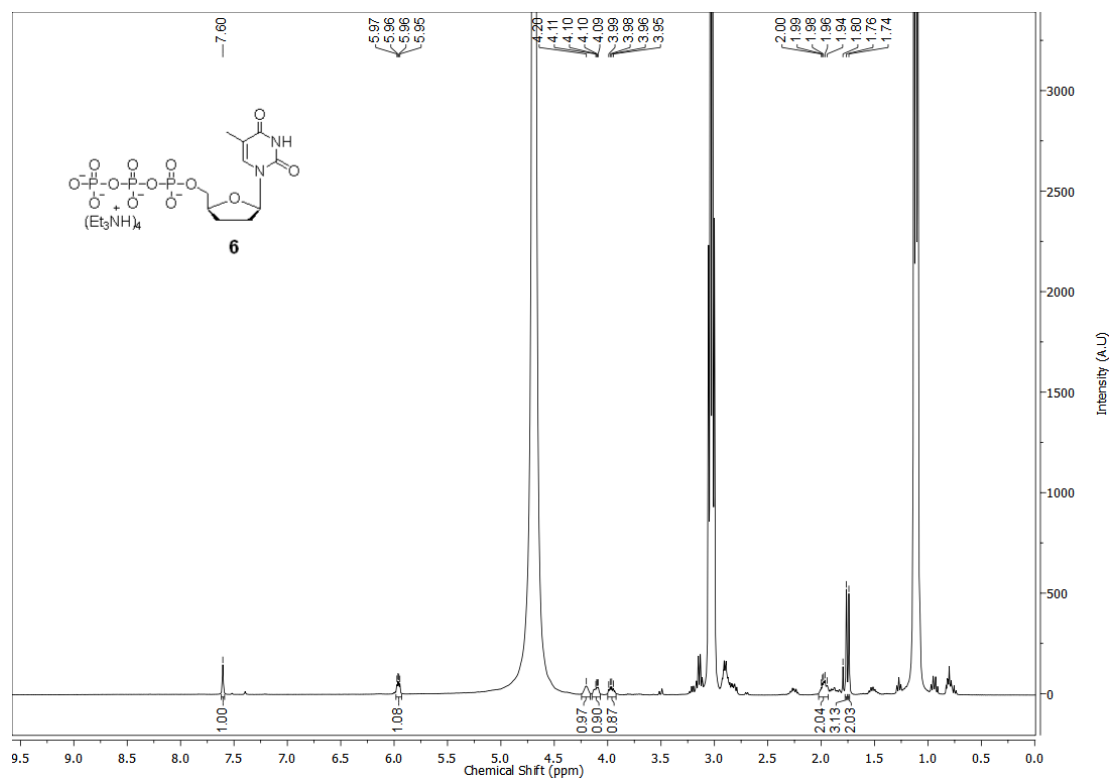
^1H -NMR of **6a** (400 MHz, d_6 -DMSO) ^{13}C -NMR of **6a** (100 MHz, d_6 -DMSO)

^1H -NMR of **6b** (400 MHz, d_6 -DMSO) ^{13}C -NMR of **6b** (100 MHz, d_6 -DMSO)

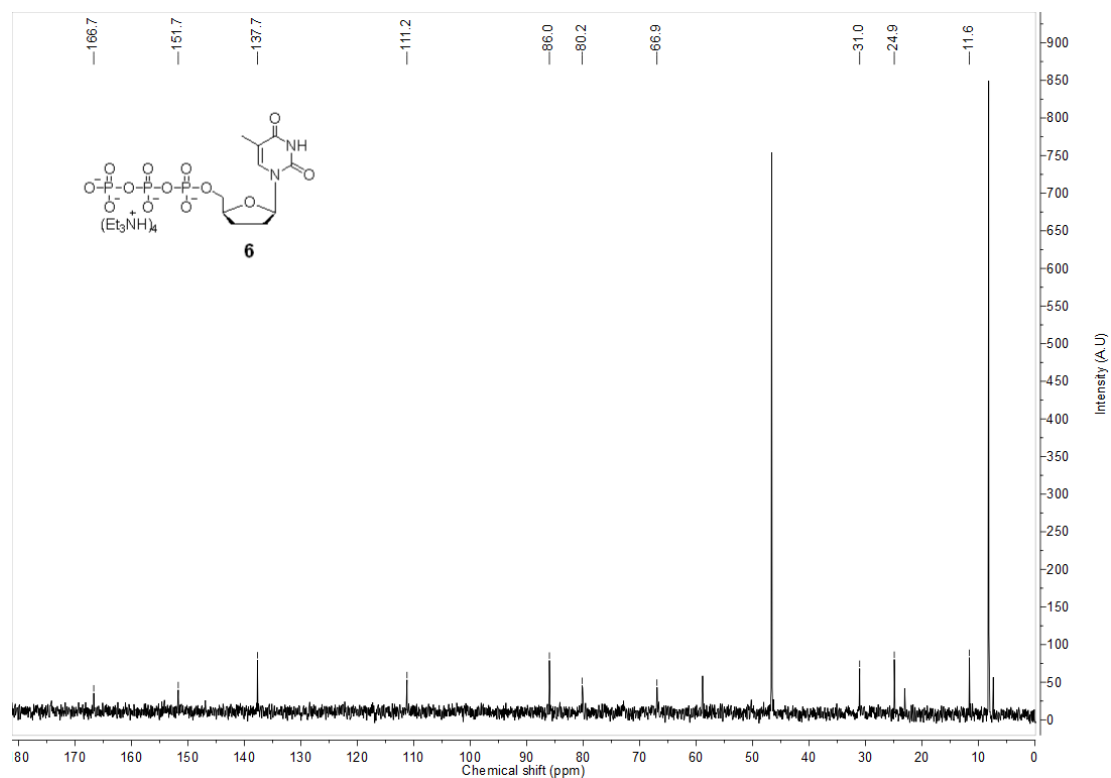
¹H-NMR of **6c** (400 MHz, *d*₆-DMSO)¹³C-NMR of **6c** (100 MHz, *d*₆-DMSO)

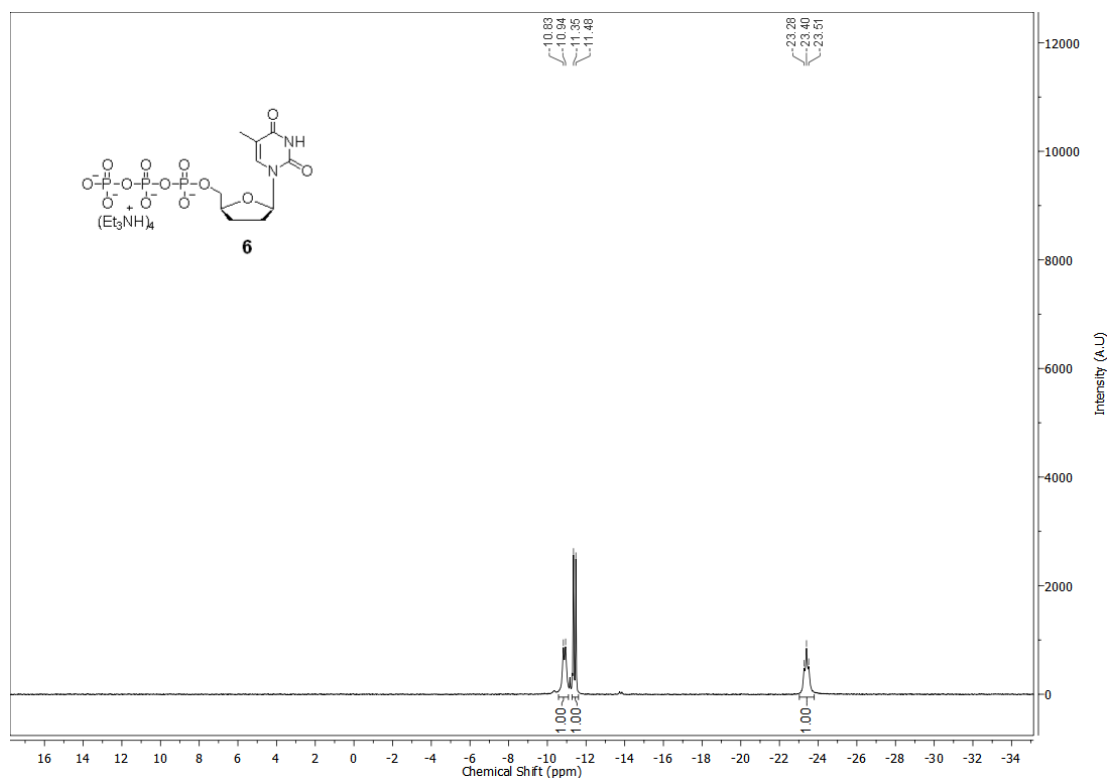
¹H-NMR of **6d** (400 MHz, *d*₆-DMSO)¹³C-NMR of **6d** (100 MHz, *d*₆-DMSO)

^1H -NMR of **6** (400 MHz, D_2O), a trace amount of triethylammonium acetate salt is present.

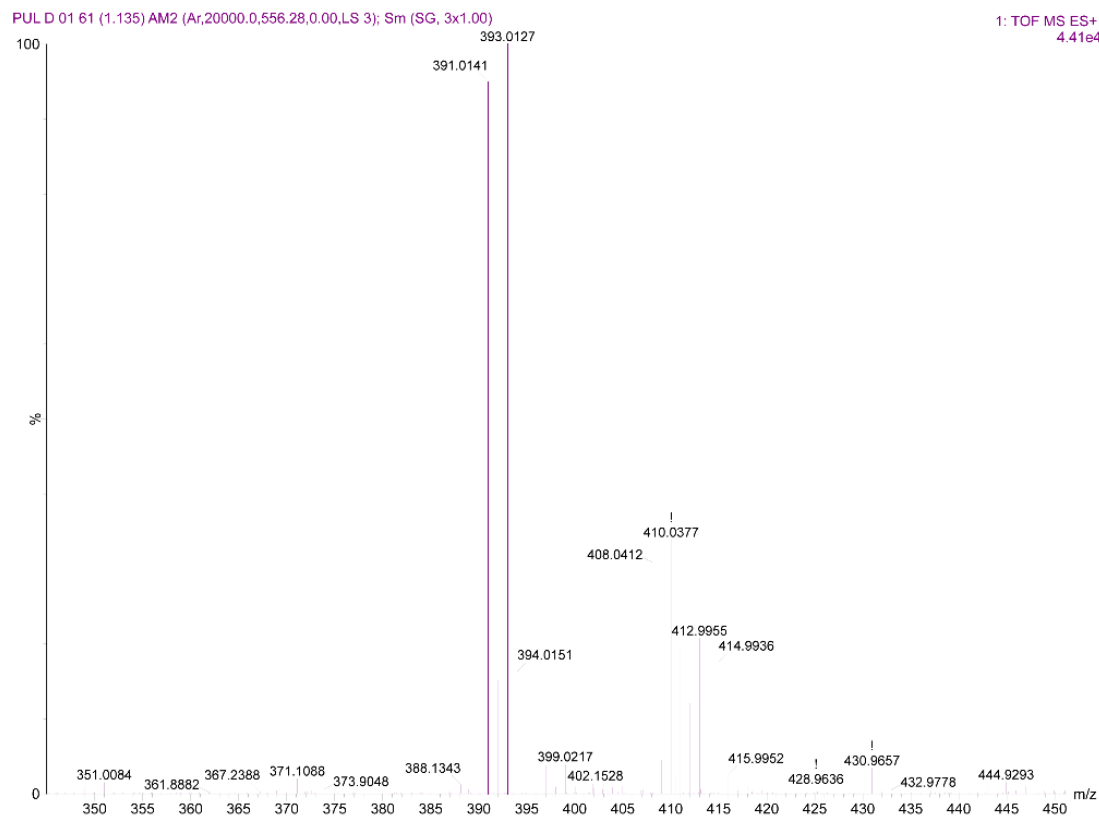


^{13}C -NMR of **5** (100 MHz, D_2O), a trace amount of triethylammonium acetate buffer is present.

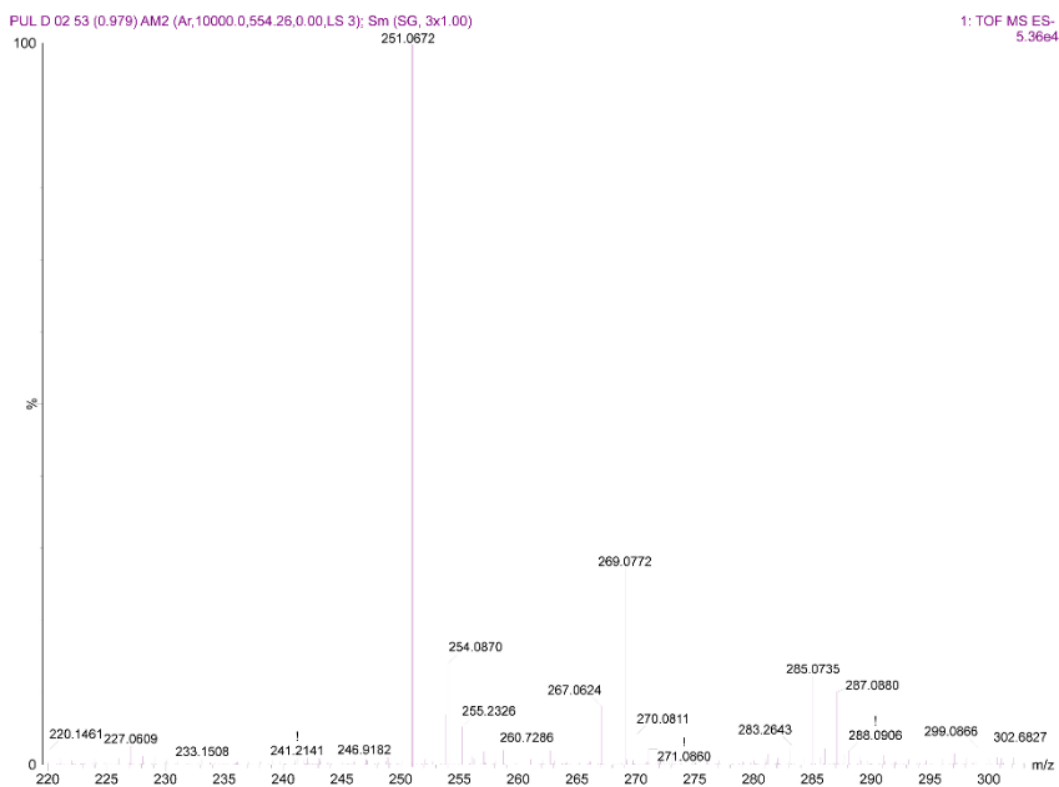


^{31}P -NMR of **6** (162 MHz, D_2O)

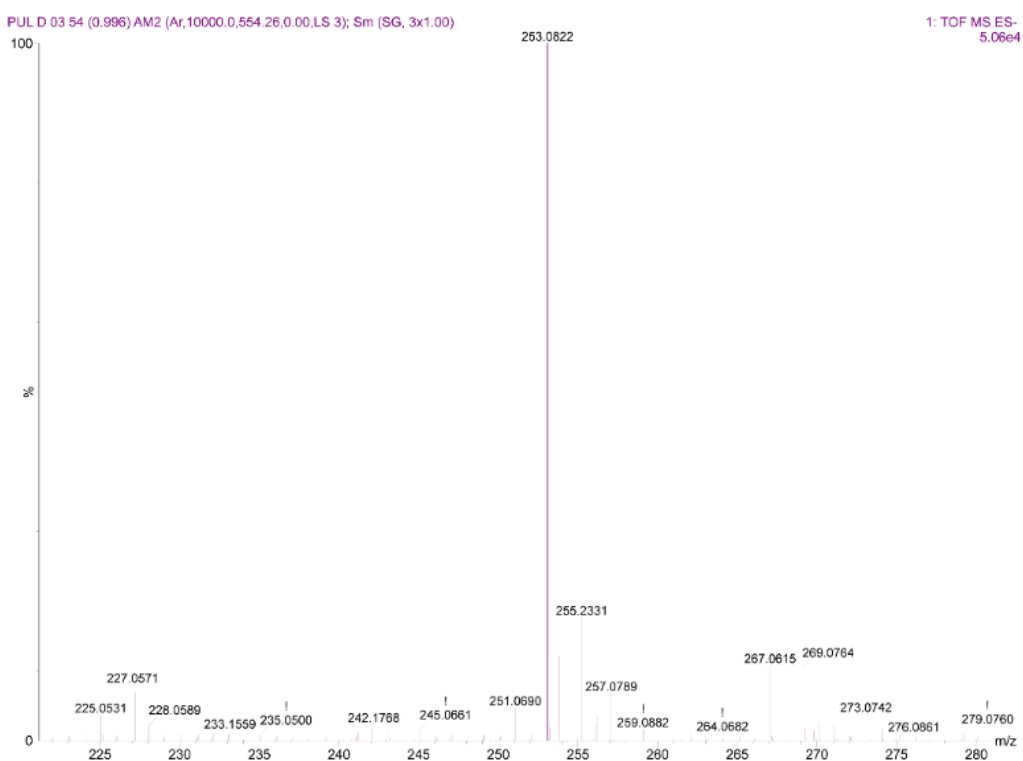
HRMS spectrum of **5a** in positive mode. Calculated for $\text{C}_{13}\text{H}_{15}\text{N}_2\text{O}_7\text{Br}$ $[\text{M}+\text{H}]^+ = 391.0141$, found $[\text{M}+\text{H}]^+ = 391.0141$.



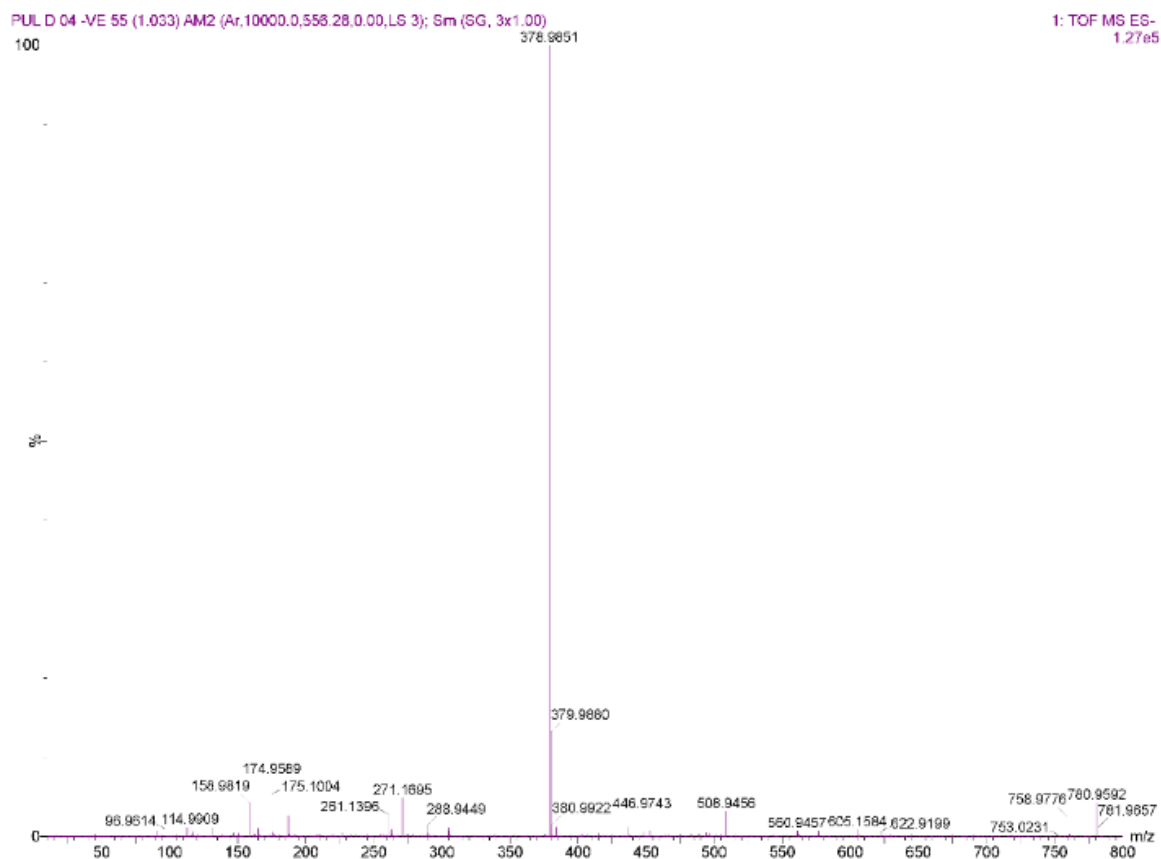
HRMS spectrum of **5b** in negative mode. Calculated for $C_{11}H_{12}N_2O_5$ $[M-H]^-$ = 251.0673, found $[M-H]^-$ = 251.0672.



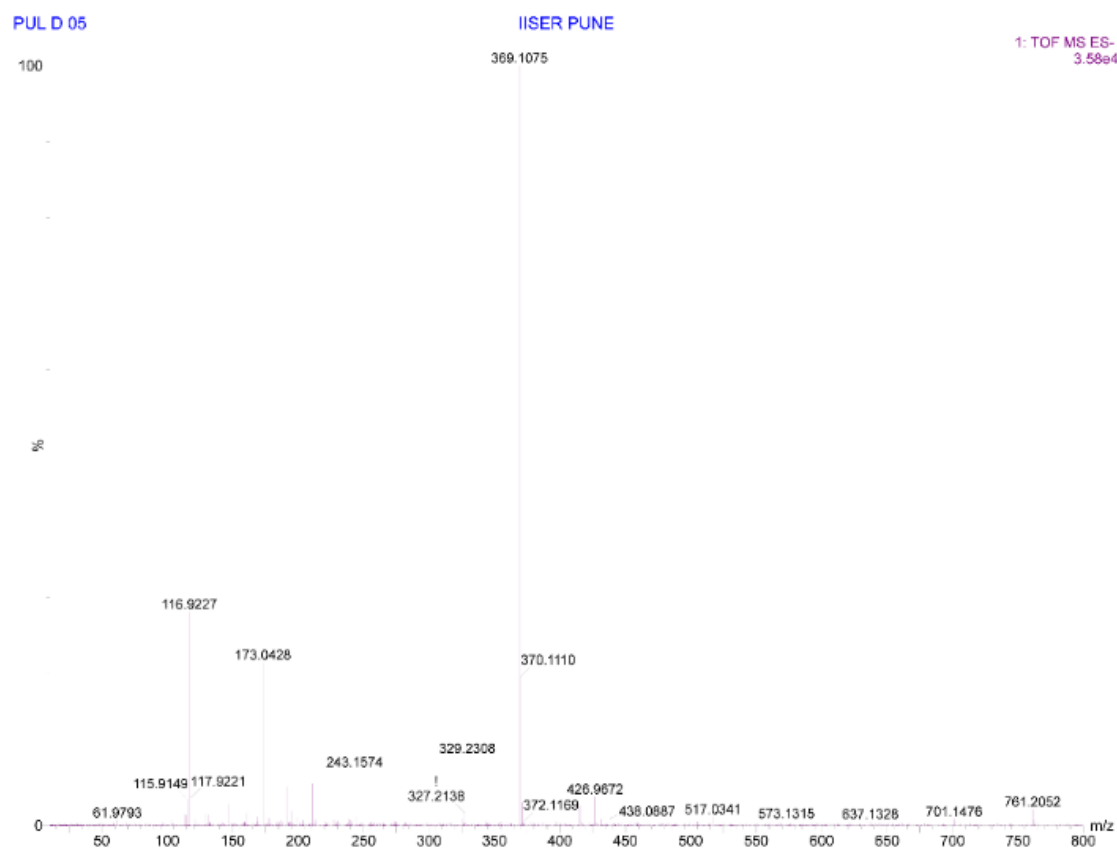
HRMS spectrum of **5c** in negative mode. Calculated for $C_{11}H_{14}N_2O_5$ $[M-H]^-$ = 253.0830, found $[M-H]^-$ = 253.0822.



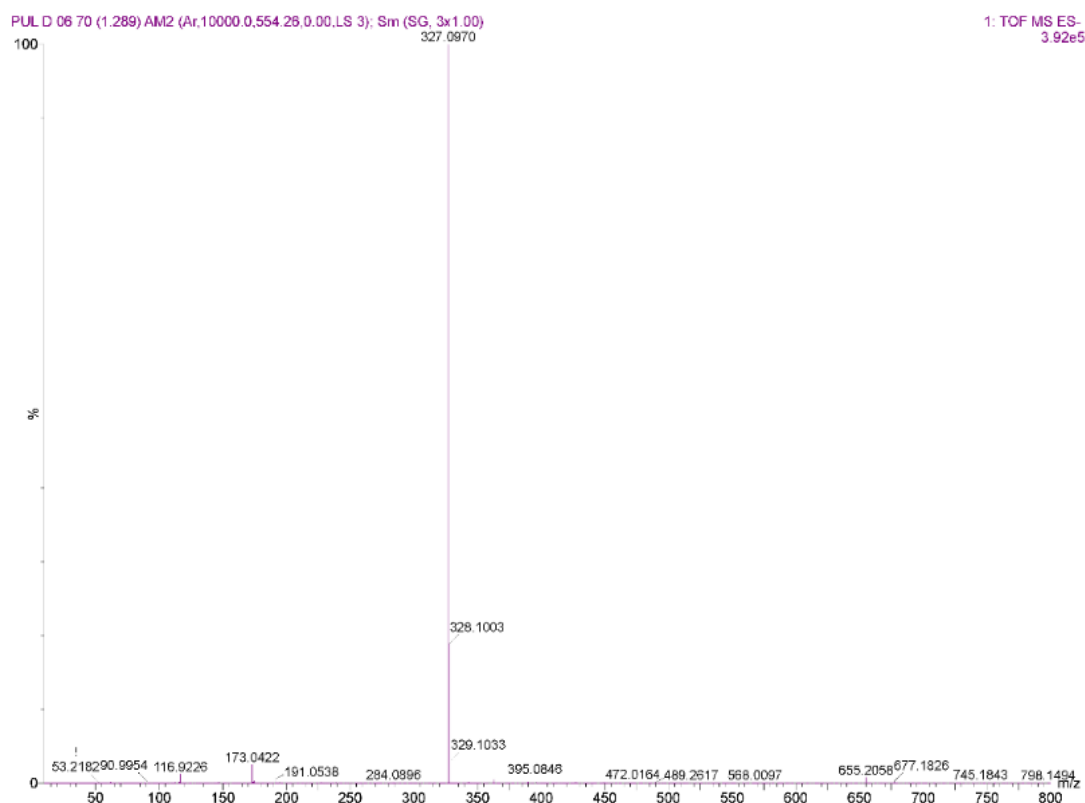
HRMS spectrum of **5d** in negative mode. Calculated for $C_{11}H_{13}IN_2O_5$ $[M-H]^-$ = 378.9796, found $[M-H]^-$ = 378.9851



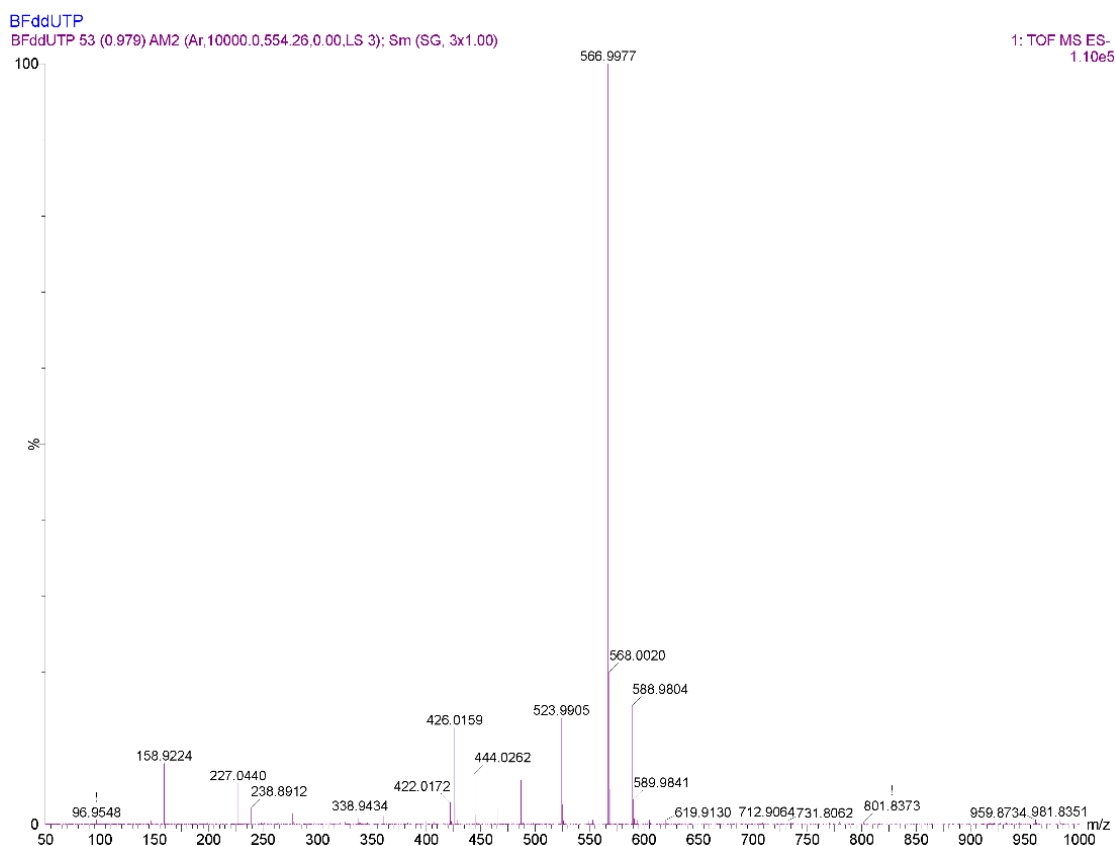
HRMS spectrum of **5e** in negative mode. Calculated for $C_{19}H_{17}N_2O_6$ $[M-H]^-$ = 369.3535, found $[M-H]^-$ = 369.1075.



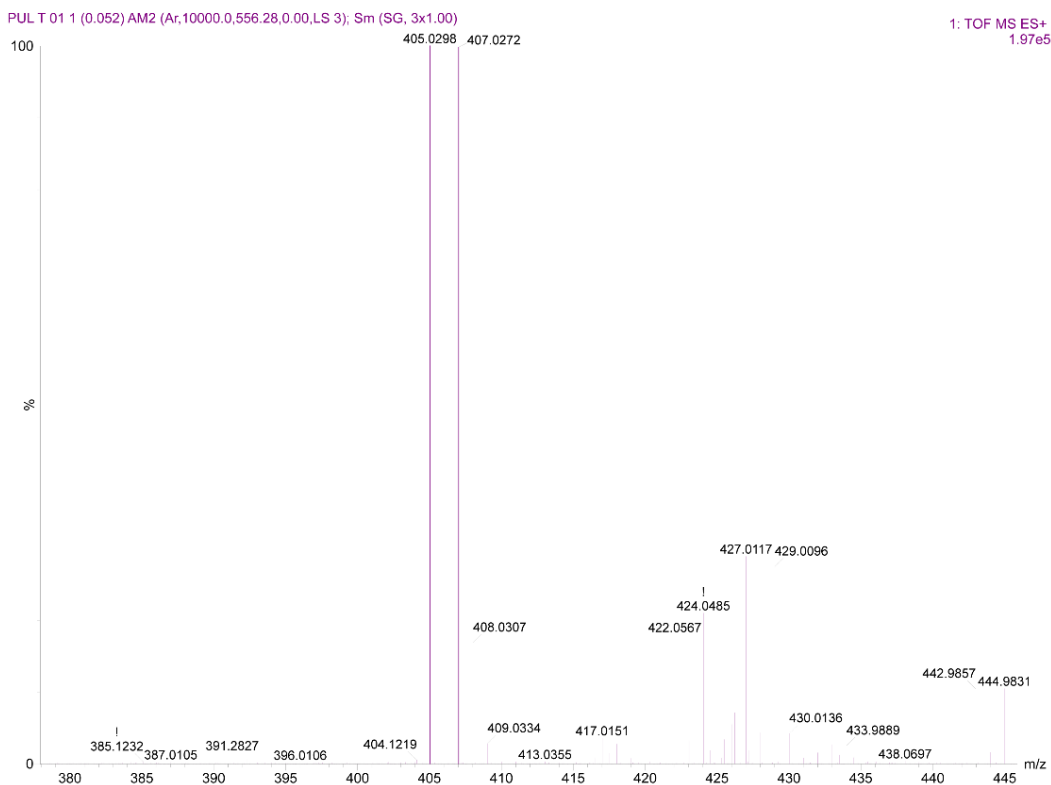
HRMS spectrum of **5f** in negative mode. Calculated for $C_{19}H_{17}N_2O_6$ $[M-H]^- = 327.3165$, found $[M-H]^- = 327.0970$



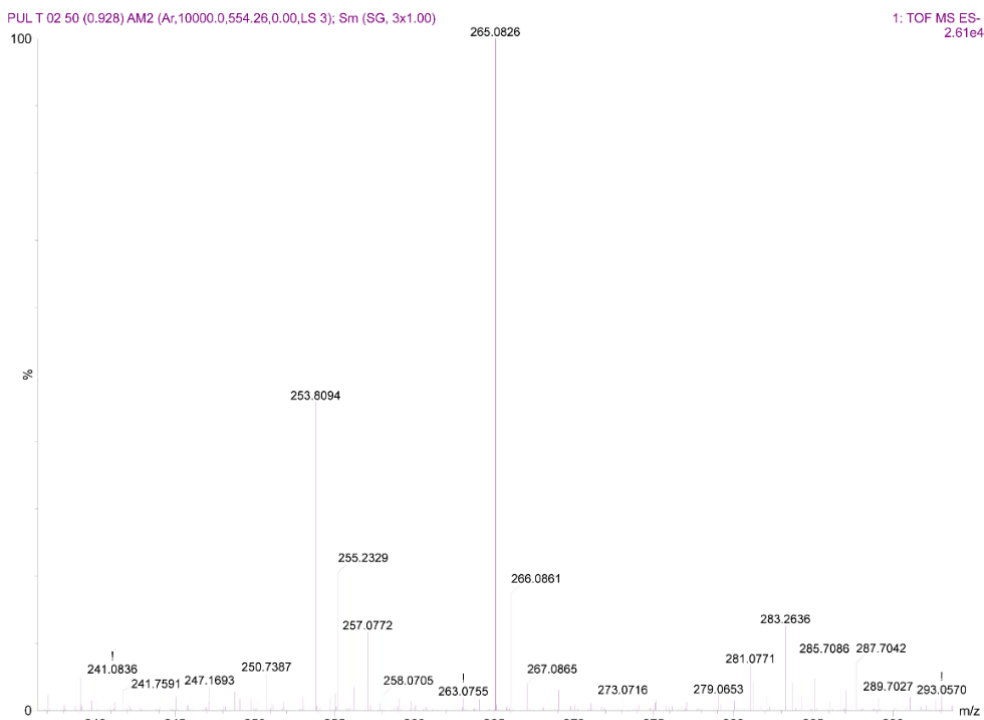
HRMS spectrum of **5** in negative mode. Calculated for $C_{17}H_{18}N_2O_{14}P_3$ $[M-H]^- = 567.2528$, found $[M-H]^- = 566.9977$.



HRMS spectrum of **6a** in negative mode. Calculated for $C_{14}H_{17}N_2O_7Br$ $[M+H]^+ = 405.0292$, found $[M+H]^+ = 405.0298$.



HRMS spectrum of **6b** in negative mode. Calculated for $C_{12}H_{14}N_2O_5$ $[M-H]^- = 265.0829$, found $[M-H]^- = 265.0826$.

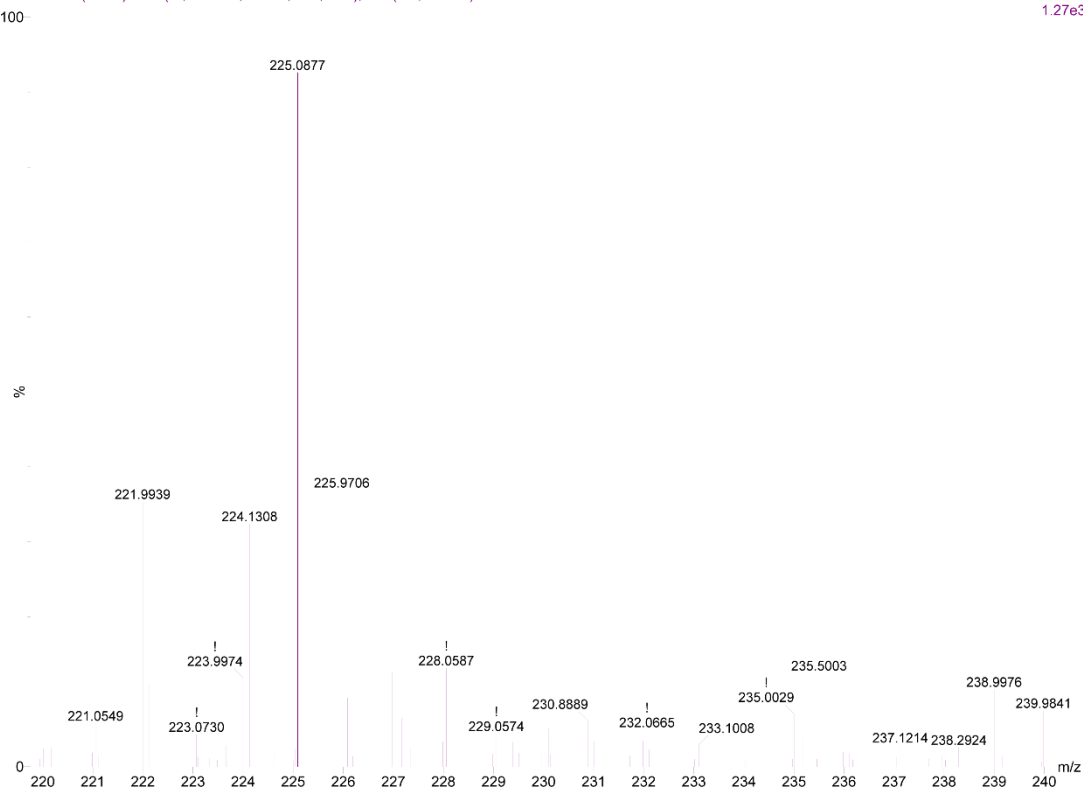


HRMS spectrum of **6c** in positive mode. Calculated for $\text{C}_{10}\text{H}_{12}\text{N}_2\text{O}_4$ $[\text{M}+\text{H}]^+ = 225.0875$, found $[\text{M}+\text{H}]^+ = 225.0877$

PUL T 04

PUL T 04 32 (0.601) AM2 (Ar,20000.0,556.28,0.00,LS 3); Sm (SG, 3x1.00)

1: TOF MS ES+
1.27e3



HRMS spectrum of **6c** in negative mode. Calculated for $\text{C}_{10}\text{H}_{14}\text{N}_2\text{O}_4$ $[\text{M}+\text{H}]^+ = 227.1032$, found $[\text{M}+\text{H}]^+ = 227.1070$

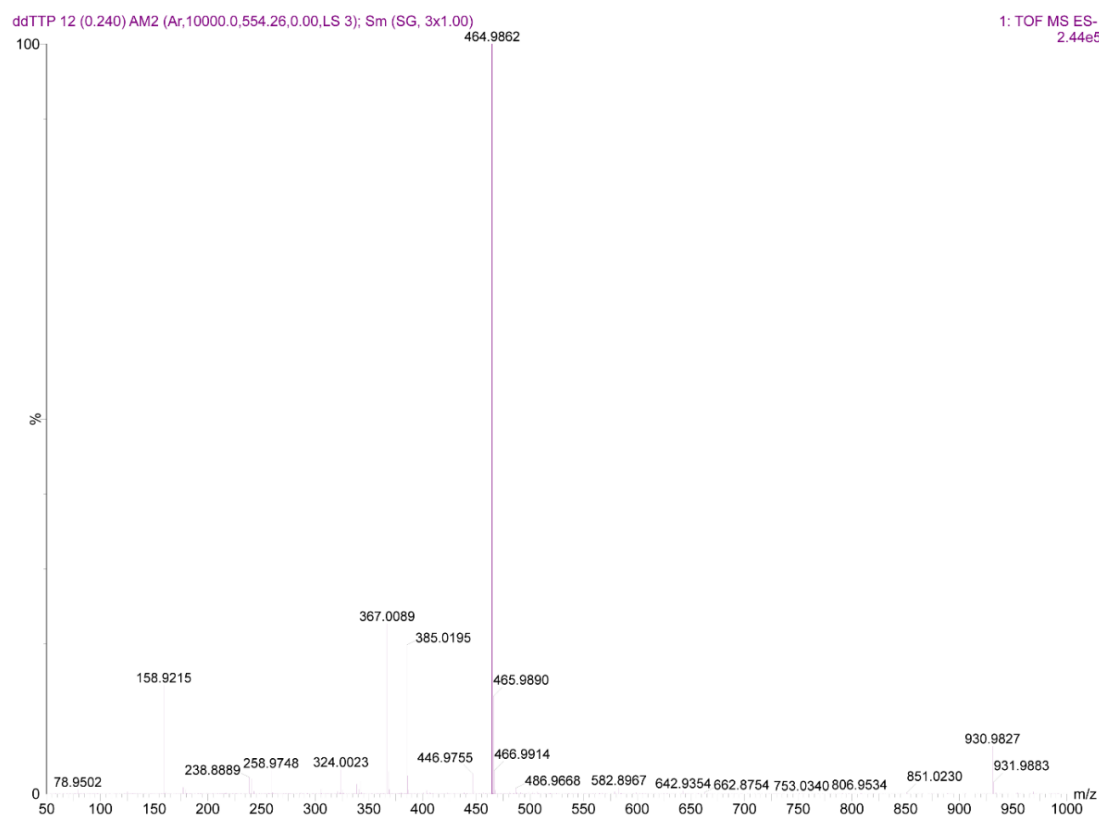
PUL T 05

PUL T 05 49 (0.911) AM2 (Ar,20000.0,556.28,0.00,LS 3); Sm (SG, 3x1.00)

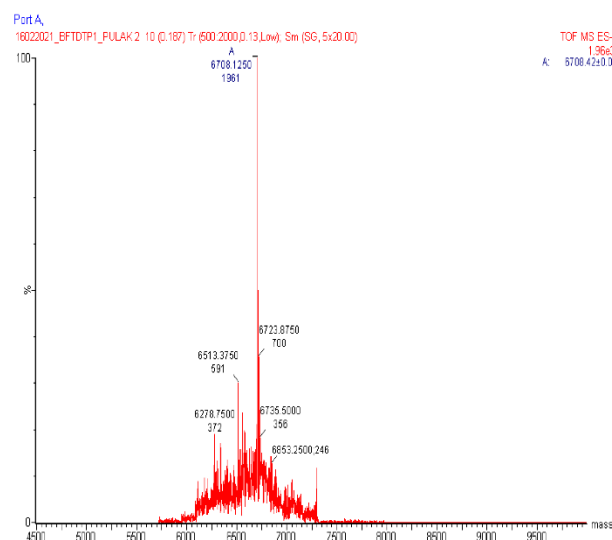
1: TOF MS ES+
663



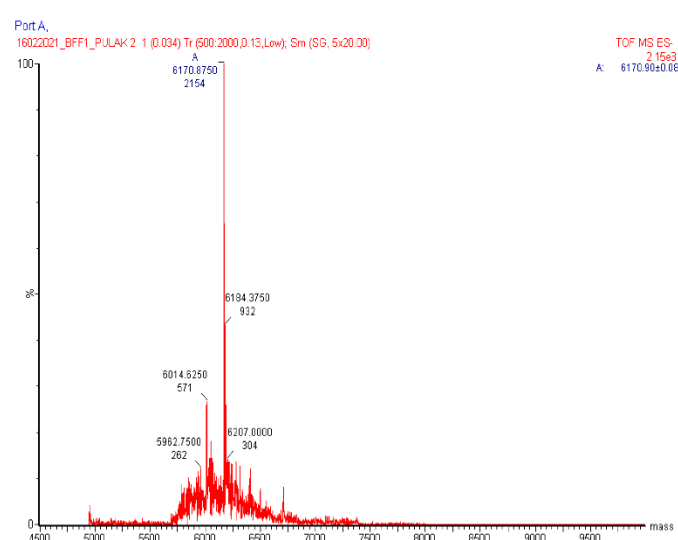
HRMS spectrum of **6** in negative mode. Calculated for $C_{17}H_{18}N_2O_{14}P_3$ $[M-H]^- = 464.9871$, found $[M-H]^- = 464.9862$.



(A)



(B)



ESI-MS spectra for BFdU-labeled DNA ON products (A) **P6** and (B) **P6***. For the sequence please see Table 1. Calculated mass for **P6** 6708.48 Da, found 6708.12. Calculated mass for **P6*** 6170.95, found 6170.87.

Overall conclusion and outlook

In this thesis, we have utilized a series of environment-sensitive nucleotide analogs to study the processing efficiency of nucleic acid processing enzymes, namely DNA polymerases and terminal nucleotide transferase (TdT). The synthesized C5-heterocyclic-modified dUTP analogs serve as good substrates for these enzymes, albeit with varying efficiency. Reactions with different templates made to guide the incorporation of single and multiple modifications by family A and B DNA polymerases reveal a size-dependent and neighboring base interaction-dependent incorporation pattern. Family B DNA polymerase exhibits efficient processing ability for all the modifications. On the other hand, family A DNA polymerase demonstrates stringency in incorporating bicyclic modifications in multiple positions. Whereas, a major limitation of the template-dependent nucleic-acid labeling techniques relies on introducing a biophysical tag in the site-specific position, where the template strand guides the incorporation of multiple modification sites. To tackle this challenge, we developed a two-step site-specific DNA labeling method by utilizing the differential processing ability of KlenTaq and KOD DNA polymerases. This technique proves valuable for incorporating fluorescent tags or post-synthetic modifications for future applications.

Further, by exploiting the responsiveness of the environment-sensitive nucleotide analogs, we investigated the catalytic events of enzymatic incorporation, such as binary complex formation and the enzymatic incorporation in solution state using fluorescence and ^{19}F NMR spectroscopy techniques. The fluorescence study sheds light on the differences in binding affinity between KlenTaq and KOD DNA polymerase towards the 3'-end modified primer-template duplex. Furthermore, crystal structures of three ternary complexes of KlenTaq containing BFdUTP, FBFdUTP, and SedUTP in the nucleotide-binding site provide unprecedented insights into the processing of these nucleotide analogs. This study highlighted important factors like the amino-acid conformational changes and stacking interaction between incoming nucleotides and primer strands that play crucial roles in enzymatic incorporation. In contrast, KOD DNA polymerase demonstrates high processivity for these modified nucleotides, possibly because of features like large channel volume and B-form of DNA near the active site. Taken together, we present an efficient probing platform that seamlessly correlates biochemical data, fluorescence, NMR, and X-ray crystallography techniques to provide a detailed understanding of the enzymatic process. This approach has immense potential in developing nucleotide analogs compatible with polymerase for various biotechnological applications.

To extend the utility of such a biophysical platform, we studied the incorporation of these analogs by the terminal deoxynucleotidyl transferase (TdT) enzyme, a therapeutic target and biomarker for leukemia diseases. Modified nucleotide analogs demonstrate a distributive incorporation mechanism and exhibit certain stops after double modifications. To investigate the reason behind this incorporation behavior, we have studied the incorporation of a responsive benzofuran-modified dideoxy nucleotide. The analog exhibits preferential incorporation over natural analog primarily due to the extra stability from the stacking interaction with the 3' primer end. To this end, the probe's versatility enabled the creation of a detection platform for monitoring TdT activity and a screening platform for identifying novel TdT inhibitors.

The advantage of our nucleotides is that they serve as good substrates and also as probes to evaluate the processivity of enzymes. The method demonstrated in this thesis has the potential for further development into a high-throughput screening assay. Recently, Tor's group utilized emissive adenosine analogs for high-throughput screening and to discover novel adenosine deaminase (ADA) inhibitors, which are particularly important for treating certain types of leukemia.¹ Additionally, the Veelma group employed expanded adenosine analogs to develop a high-throughput assay, leading to the identification of a non-nucleoside TdT inhibitor.² Taken together, the techniques demonstrated in this thesis not only facilitate monitoring enzymatic incorporation mechanisms in real-time and at the atomic level but also pave the way for studying viral polymerases and engineered DNA and RNA polymerases. While Chaput's and Holliger's labs^{3,4} have explored engineered polymerases incorporating sugar-modified nucleotides, our platform offers a valuable tool for studying the incorporation of base-modified nucleotide analogs.

Over the past few years, various isomorphous emissive nucleobases have been developed. These substitutes have demonstrated the ability to effectively supplant their natural equivalents across various scenarios, offering a faithful optical window into biochemical transformations. This kind of probe system is employed to monitor the reactivity of ADAR2, an RNA editing enzyme.¹ Despite its efficacy in various fields like biotechnology, biophysics, chemical biology, and drug discovery platforms, its application in biological imaging or cell biology remains largely unexplored due to significant technical hurdles. Challenges such as short emission wavelengths, low emissive, and poor cellular uptake are commonly encountered with small heterocycles. However, recent progress in multiphoton spectroscopy and further structural modifications have shown potential for overcoming these barriers.^{5,6} Novel

alternative modifications and advancements in multiphoton excitation could address these perceived limitations and broaden the possibilities for small fluorophores.

References

1. Sinkeldam, R. W.; McCoy, L. S.; Shin, D.; Tor, Y. Enzymatic interconversion of isomorphous fluorescent nucleosides: adenosine deaminase transforms an adenosine analogue into an inosine analogue. *Angew. Chem., Int. Ed.* **2013**, *52*, 14026–14030.
2. Pals, M. J.; Velema, W. A. A quenched size-Expanded nucleotide reports activity of the leukemia biomarker terminal deoxynucleotidyl transferase (TdT)
3. Cozens, C.; Pinheiro, V. B.; Vaisman, A.; Woodgate, R.; Holliger, P. A short adaptive path from DNA to RNA polymerases. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 8067–8072.
4. Dunn, M. R.; Otto, C.; Fenton, K. E.; Chaput, J. C. Improving polymerase activity with unnatural substrates by sampling mutations in homologous protein architectures. *ACS Chem. Biol.* **2016**, *11*, 1210–1219.
5. Hadidi, K.; Steinbuch, K. B.; Dozier, L. E.; Patrick, G. N.; Tor, Y. Inherently Emissive Puromycin Analogues for Live Cell Labelling. *Angew. Chem., Int. Ed.* **2023**, *62*, e202216784.
6. Nobis, D.; Fisher, R. S.; Simmermacher, M.; Hopkins, P. A.; Tor, Y.; Jones, A. C.; Magennis, S. W. Single-molecule detection of a fluorescent nucleobase analogue via multiphoton excitation. *J. Phys. Chem. Lett.* **2019**, *10*, 5008–5012.

Copyrights and Permission



Sign in/Register ?



Most Trusted. Most Cited. Most Read.

Microenvironment-Sensitive Fluorescent Nucleotide Probes from Benzofuran, Benzothiophene, and Selenophene as Substrates for DNA Polymerases

Author: Pulak Ghosh, Heike M. Kropp, Karin Betz, et al
Publication: Journal of the American Chemical Society
Publisher: American Chemical Society
Date: Jun 1, 2022

Copyright © 2022, American Chemical Society

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms and Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your RightsLink request. No additional uses are granted (such as derivative works or other editions). For any uses, please submit a new request.

If credit is given to another source for the material you requested from RightsLink, permission must be obtained from that source.

BACK

CLOSE WINDOW