PYRROLIDINYL PEPTIDE NUCLEIC ACID PROBE CAPABLE OF CROSSLINKING WITH DNA

Miss Penthip Muangkaew

A Dissertation Submitted in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy in Chemistry

Department of Chemistry

FACULTY OF SCIENCE

Chulalongkorn University

Academic Year 2020

พิร์โรลิดินิลเพปไทด์นิวคลีอิกแอซิดโพรบที่สามารถเชื่อมขวางกับดีเอ็นเอ

น.ส.เพ็ญทิพ เมืองแก้ว

Thesis Title	PYRROLIDINYL PEPTIDE NUCLEIC ACID PROBE CAPABLE OF CROSSLINKING WITH DNA
By	Miss Penthip Muangkaew
Field of Study	Chemistry
Thesis Advisor	Professor TIRAYUT VILAIVAN, Ph.D.

Accepted by the FACULTY OF SCIENCE, Chulalongkorn University in Partial Fulfillment of the Requirement for the Doctor of Philosophy

--- Dean of the FACULTY OF SCIENCE

(Professor POLKIT SANGVANICH, Ph.D.)

DISSERTATION COMMITTEE

--- Chairman

(Associate Professor VUDHICHAI PARASUK, Ph.D.)

--- Thesis Advisor

(Professor TIRAYUT VILAIVAN, Ph.D.)

--- Examiner

(Professor THAWATCHAI TUNTULANI, Ph.D.)

--- Examiner

(Professor NONGNUJ MUANGSIN, Ph.D.)

--- Examiner

(Associate Professor VORAVEE HOVEN, Ph.D.)

--- External Examiner

(Assistant Professor Chaturong Suparpprom, Ph.D.)

เพ็ญทิพ เมืองแก้ว : พิร์โรลิดินิลเพปไทด์นิวคลีอิกแอซิดโพรบที่สามารถเชื่อมขวางกับดีเอ็นเอ. (PYRROLIDINYL PEPTIDE NUCLEIC ACID PROBE CAPABLE OF CROSSLINKING WITH DNA) อ.ที่ปรึกษาหลัก : ศ. ดร.ธีรยุทธ วิไลวัลย์

การเชื่อมขวางระหว่างสายดีเอ็นเอเป็นปรากฏการณ์ที่ดีเอ็นเอสองสาย ซึ่งมักจะเป็นส่วนของสายที่เป็นเกลียวคู่เดียวกัน จับยึดกันโดยผ่านพันธะโควาเลนต์ กระบวนการนี้ขัดขวางการแยกสายดีเอ็นเอออกจากกันซึ่งเป็นสิ่งจำเป็นสำหรับหน้าที่ทางชีวภาพของมัน ดังนั้นจึงทำให้เกิดความเสียหายต่อเซลล์อย่างร้ายแรง เพื่อแก้ปัญหาความไม่จำเพาะเจาะจงของตัวที่ทำให้เกิดการเชื่อมขวางทั่วไป จึงได้มีการนำเสนอโพรบที่สามารถเชื่อมขวางได้ โดยมีหมู่ที่ว่องไวติดอยู่บนโพรบที่สามารถจดจำส่วนของดีเอ็นเอเป้าหมายได้อย่างจำเพาะเจาะจง ความสามารถในการจับยึดที่แข็งแรงและความเสถียรทางชีวภาพของพิร์โรลิดินิลเพปไทด์นิวคลีอิกแอซิด (เอซีพีซีพีเอ็นเอ, acpcPNA) เป็นแรงบันดาลใจให้พัฒนาโพรบชนิดใหม่ที่สามารถเกิดการเชื่อมขวางได้เมื่อถูกกระตุ้นตามความต้องการ โดยอาศัยหมู่ฟิวแรนเป็นหมู่ว่องไวที่ซึ่งสามารถถูกกระตุ้นได้โดยกระบวนการออกซิเดชัน ในงานวิจัยได้มีการพัฒนาวิธีการสังเคราะห์ โพรบดังกล่าวด้วยวิธีการดัดแปรโครงสร้างภายหลังการสังเคราะห์ที่ง่ายและมีประสิทธิภาพ และได้ศึกษาปฏิกิริยาการเชื่อมขวางของมันกับดีเอ็นเอโดยอาศัยเทคนิคการศึกษาการสูญเสียสภาพธรรมชาติด้วยความร้อน พอลิอะคริลาไมด์เจลอิเล็กโตรโฟรีซิสแบบถูกทำให้เสียสภาพ รีเวิร์สเฟสโครมาโทกราฟี และมัลดิ-ทอฟ แมสสเปกโทรเมทรี งานวิจัยถูกแบ่งเป็นสองส่วน ในส่วนแรกเป็นการศึกษาผลของตำแหน่งของหมู่ฟิวแรนบนสายของเอซีพีซีพีเอ็นเอต่อประสิทธิภาพการเกิดการเชื่อมขวาง โพรบถูกออกแบบให้หมู่ฟิวแรนถูกติดอยู่บนโพรบอย่างอิสระจากนิวคลีโอเบส ดังนั้นมันจึงไม่รบกวนกระบวนการเข้าคู่กันของเบส ผลการศึกษาพบว่าเอซีพีซีพีเอ็นเอที่ถูกดัดแปรด้วยหมู่ฟิวแรนที่ปลายสายสามารถเกิดการเชื่อมขวางกับเบสไซโตซีน > อะดีนีน >> กวานีน ในขณะที่เอซีพีซีพีเอ็นเอที่ถูกดัดแปรด้วยหมู่ฟิวแรนบริเวณข้างในสายไม่สามารถเกิดได้ โดยมีการพบเป็นครั้งแรกว่าเบสกวานีนสามารถเกิดการเชื่อมขวางกับฟิวแรน การเพิ่มเบสไซโตซีนที่ไม่ได้เข้าคู่แทรกลงไปบนลำดับเบสของดีเอ็นเอสามารถเพิ่มประสิทธิภาพการเชื่อมขวางของโพรบที่ดัดแปรด้วยหมู่ฟิวแรนบริเวณข้างในสายได้ ข้อสรุปที่ได้คือปัจจัยสำคัญที่ส่งผลต่อความสำเร็จของการเชื่อมขวางคือการมีนิวคลีโอเบสที่ไม่เข้าคู่อยู่ใกล้หมู่ฟิวแรนที่ถูกกระตุ้น นอกเหนือไปจากการที่โพรบจะต้องจับยึดกับเป้าหมายได้อย่างมีเสถียรภาพ งานในส่วนที่สองเป็นการศึกษาเพื่อเปรียบเทียบประสิทธิภาพของการเชื่อมขวางของเอซีพีซีพีเอ็นเอและเออีจีพีเอ็นเอที่ถูกดัดแปรด้วยหมู่ฟิวแรนที่หลากหลายชนิดบนปลายสาย ผลการทดลองพบว่าชนิดของหมู่ฟิวแรนและชนิดของพีเอ็นเอโพรบที่แตกต่างกันส่งผลต่อประสิทธิภาพการเชื่อมขวางและความเลือกจำเพาะที่แตกต่างกัน โดยส่วนใหญ่เออีจีพีเอ็นเอโพรบแสดงประสิทธิภาพการเชื่อมขวางที่สูงกว่าเอซีพีซีพีเอ็นเออยู่พอสมควร แต่เอซีพีซีพีเอ็นเอแสดงความจำเพาะเจาะจงต่อเบสไซโตซีนที่สูงกว่า นอกจากนั้น ยังได้มีการศึกษาเอซีพีซีพีเอ็นเอโพรบที่ถูกดัดแปรด้วยหมู่ฟิวแรนข้างในสายกับกลุ่มดีเอ็นเอเป้าหมายที่หลากหลาย โดยพบว่าประสิทธิภาพของปฏิกิริยาการเชื่อมขวางสูงสุดจะเกิดเมื่อมีเบสไซโตซีนที่ไม่เข้าคู่อยู่ตำแหน่งตรงข้ามหรืออยู่ติดกันกับตำแหน่งตรงข้ามของหมู่ฟิวแรน โดยข้อสรุปหลักจากการศึกษานี้คือทั้งความสามารถในการจับยึดของโพรบกับเป้าหมายเกิดเป็นสายคู่ที่เสถียร และการมีดีเอ็นเอนิวคลีโอเบสที่มีหมู่อะมิโนอยู่นอกวง (อะดีนีน ไซโตซีน กวานีน) ที่ไม่ได้เข้าคู่เป็นปัจจัยสำคัญต่อการเกิดปฏิกิริยาการเชื่อมขวาง ซึ่งให้ข้อมูลเชิงลึกถึงความต้องการที่จำเป็นต่อการเกิดการเชื่อมขวางที่มีประสิทธิภาพ

สาขาวิชา   เคมี   ลายมือชื่อนิสิต ....

ปีการศึกษา   2563   ลายมือชื่อ อ.ที่ปรึกษาหลัก ....

5772833323 : MAJOR CHEMISTRY

KEYWORD: cross-link, acpcPNA, furan

Penthip Muangkaew : PYRROLIDINYL PEPTIDE NUCLEIC ACID PROBE CAPABLE OF CROSSLINKING WITH DNA. Advisor: Prof. TIRAYUT VILAIVAN, Ph.D.

DNA interstrand cross-linking (ICL) is a phenomenon in which two strands of the DNA (most often as part of the same DNA duplex) bind together via a covalent bond. This process prevents the DNA strand dissociation, which is essential for its biological functions, and thus is extremely damaging to cells. To overcome the problem of non-specificity associated with general DNA cross-linking agents, a cross-linkable probe in which the reactive group is placed on the probe that can specifically recognize part of the DNA target. The strong binding affinity and biological stability of the pyrrolidinyl peptide nucleic acid (acpcPNA) inspired us to develop a new activate-on-demand cross-linkable acpcPNA probe with furan as the reactive moiety, which can be activated by an oxidative process. A simple post-synthetic modification protocol was developed to synthesize such probes and their cross-linking with DNA were studied by thermal denaturation, denaturing PAGE, reverse phase HPLC, and MALDI-TOF MS. The work was divided into two sections. In section I, the effect of the furan position on the acpcPNA strand towards the cross-linking efficiency was investigated. The probe was designed so that the furan moiety was attached to the probe independent from the nucleobase so that it does not interfere with the normal base-pairing process. The results revealed that terminally furan-modified acpcPNA could undergo cross-linking with C > A >> G while the internally furan-modified acpcPNA could not. The G-base was found to participate in such cross-linking with furan for the first time. The C-inserted DNA sequences could enhance the cross-linking efficiency of the internally furan-modified probe. It was concluded that the essential factor of the successful cross-linking reaction is the availability of the nucleobase nearby the furan activated moiety in addition to the ability of the probe to form a stable hybrid with the target. In section II, the cross-linking efficiency of the acpcPNA and aegPNA modified with various furan building blocks at the terminal were compared. The results showed that various furan building blocks and various PNA types of probe offered different cross-linking efficiency and selectivity. In most cases, aegPNA probes showed somewhat higher cross-linking yields over acpcPNA probes, but the latter showed a higher selectivity toward C-base. In addition, a study of internally cross-linking with acpcPNA probes was performed with various DNA targets, and the cross-linking reaction was most efficient when there was a free C-base present at or adjacent to the opposite position of the furan building block. Overall, the major conclusion obtained from the study is that both the ability to form a stable duplex and the availability of the unpaired DNA nucleobase with an exocyclic amino group (A, C, G) were essential for the cross-linking reaction, which provides further insights into the requirements for an efficient ICL formation.

Field of Study:	Chemistry	Student's Signature ....
Academic Year:	2020	Advisor's Signature ....

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my research supervisor, Professor Dr. Tirayut Vilaivan, for giving me the opportunity to do research and providing invaluable guidance of the research as well as thesis writing. He has taught me the methodology to carry out the research and to present the research works as clearly as possible. It was a great privilege and honor to work and study under his guidance. I am extremely grateful for what he has offered me. He is not only a teacher but also a family for me. My sincere thanks are also extended to his wife, Chotima Vlaivan, for helping me in every-single thing that I asked for. She is very gentle, kind and thoughtful. I am extremely grateful to be a member of TV-lab where people are kind, helpful and support each other.

Importantly, I would like to express my gratitude to Professor Dr. Annemieke Madder for giving me an opportunity to do research and taking care of me during my internship in Ghent University. She is very smart and kind. My thanks are also extended to her lab's member especially Dr. Alex Manicardi and Mr. Enrico Cadoni without them I cannot finish my research perfectly. My life in Ghent was fulfilled with them.

I am extremely grateful to my parents, my grandfather, and my grandmother, for their love, caring, and sacrifices for educating me. I am very much thankful to my sister, for understanding, and listening in every moment in my life. My Special thanks goes to my partner, Alexandre Marouen, for keeping me company, listening, and understanding. My life during thesis writing has passed thanks to him.

last but not least, I would link to thank the Science Achievement Scholarship of Thailand, and the Overseas Research Experience Scholarship for Graduate Students from Graduate School and Faculty of Science, Chulalongkorn University for the financial supports during my Ph.D life.

Finally, my thanks go to all the people who have supported me to complete the research work directly or indirectly.
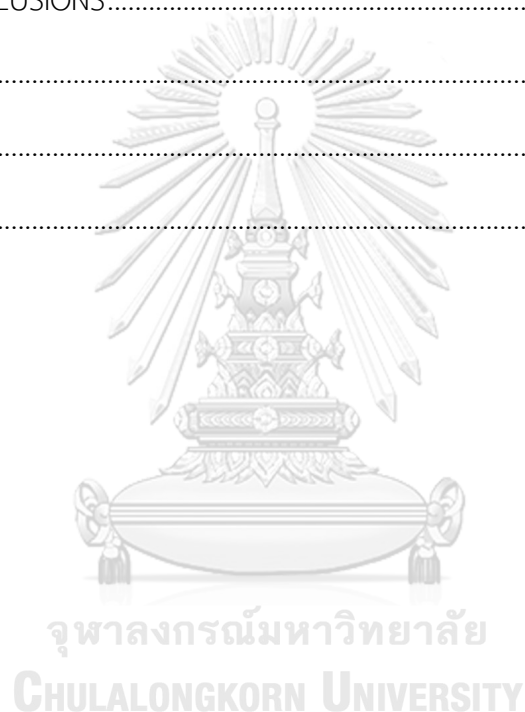

Penthip Muangkaew

TABLE OF CONTENTS

LIST OF FIGURES

จุฬาลงกรณ์มหาวิทยาลัย

CHULALONGKORN UNIVERSITY

LIST OF TABLES

LIST OF ABBREVIATIONS AND SYMBOLS

µL	microliter
µmol	micromole
A	adenine
A^{Bz}	N^6-benzoyladenine
Ac	acetyl
Ac_2O	acetic anhydride
APC	3-aminopyrrolidine-4-carboxylic acid
ACPC	(1*S*,2*S*)-2-amino-1-cyclopentanecarboxylic acid
Boc	*tert*-butoxycarbonyl
Bz	benzoyl
C	cytosine
cal.	calculated
C^{Bz}	N^4-benzoylcytosine
CCA	α-cyano-4-hydroxy cinnamic acid
DBU	1,8-diazabicyclo[5.4.0]undec-7-ene
DCM	dichloromethane
DIPEA	diisopropylethylamine
DMF	*N,N'*-dimethylformamide
DNA	deoxyribonucleic acid
Dpm	diphenylmethyl
ds	double strand
equiv	equivalent (s)
Fmoc	9*H*-fluoren-9-ylmethoxycarbonyl
g	gram
G	guanine
HATU	*O*-(7-azabenzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluorophosphate
HOAt	1-hydroxy-7-azabenzotriazole

HPLC	high performance liquid chromatography
Lys	lysine
M	molar
MALDI-TOF	matrix-assisted laser desorption/ionization-time of flight
MeCN	acetonitrile
MeOH	methanol
mg	milligram
min	minute
mL	milliliter
mM	millimolar
mmol	millimole
MS	mass spectrometry
m/z	mass to charge ratio
nm	nanometer
nmol	nanomole
°C	degree Celsius
Pfp	pentafluorophenyl
PNA	peptide nucleic acid or polyamide nucleic acid
RNA	ribonucleic acid
SPPS	Solid phase peptide synthesis
ss	single strand
T	thymine
TFA	trifluoroacetic acid
THF	tetrahydrofuran
T_m	melting temperature
UV	ultraviolet

CHAPTER I

INTRODUCTION

1.1 DNA Interstrand cross-links (ICLs)

Cross-linking agents are a type of compounds that carries two functional groups that are reactive towards two different parts of a DNA molecule. The formation of DNA cross-link can be divided into three main types. First, the link between a cross-linking agent and two DNA nucleobases from the same DNA strand is called "intrastrand cross-link". An example of the intrastrand cross-link process occurring naturally is the thymine dimerization,[1] taking place via a 2+2 cycloaddition reaction between two adjacent thymine bases of the DNA strand when they are exposed to UV light. Thymine dimer then creates a kink on the strand which can cause skin cancer, erythema or sunburn. Second, the cross-link process occurring with different DNA strands is called "interstrand cross-links (ICLs)". ICL is one of the processes that is toxic to cells. The ICL prevents the strands separation thus blocking the DNA replication, transcription and/or translation processes. The cells that cannot replicate themselves or produce proteins cannot grow and finally die. Thus, the formation of ICL can stop the growth of unwanted cells, such as cancer or tumor cells. For these reasons, ICL is an important phenomenon that may have pharmaceutical and medical applications. The study of DNA interstrand cross-links (ICLs) began after the sulfur mustard bombs in the second World War (1943). It was noticed that the chemical could exert cytotoxic action on various tissues and could attack the white blood cells.[2] In 1946, the first report involving the use of nitrogen mustard in chemotherapy was published.[3] Some drugs in this class such as cyclophosphamide and melphalan are still being used nowadays.[4]

The third type of cross-link can be found between DNA and proteins called "DNA protein cross-link (DPC)". The first publication about the DNA-protein cross-link took place when the *E. coli* was treated with acridine orange in the presence of visible light.[5, 6] It was found that the acridine orange dye – being able to absorb lights in the visible region – could induce the DNA-protein cross-link formation in a

higher yield than UV irradiation, thus could kill the cells more efficiently. The formation of DNA-protein cross-link can also be induced by chemical reagents such as formaldehyde,[7] by metal ions such as nickel,[8] or by IR radiation.[9]
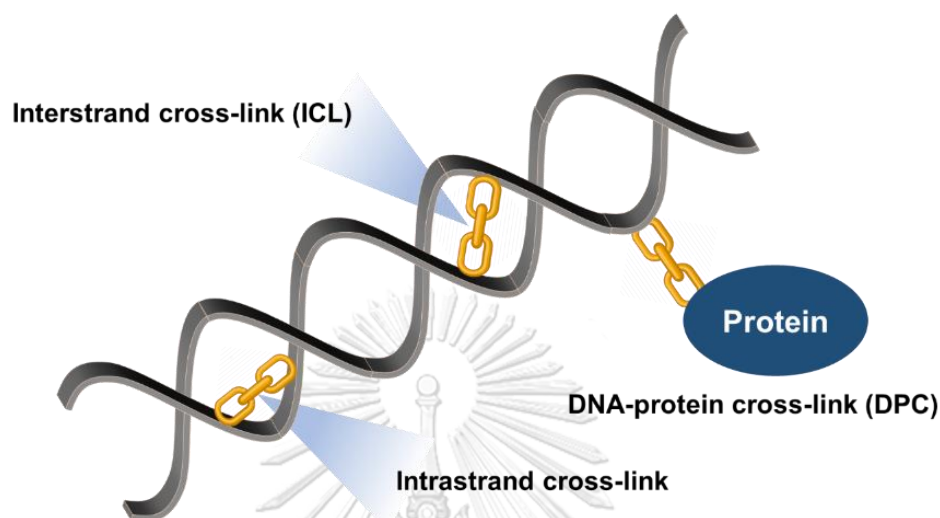
Figure 1.1 Illustration of different types of DNA cross-linking

1.2 General method of ICL

There have been much interest in developing cross-linking agents (artificial cross-linkers) to perform "on-demand" DNA cross-linking. Over the past several decades, various cross-linking agents have been developed and utilized in a vast area of biological applications.[10] They have been used as anti-cancer drugs,[11-14] for DNA/ RNA detection,[15] for DNA damage and repair studies[16] and in the creation of DNA nanostructures.[17-20] Most of these cross-linking agents are bifunctional, meaning that they contain two reactive functional groups that can be the same or different. Examples of common bifunctional cross-linkers are nitrogen mustard, platinum complexes, psoralen and mitomycin C.
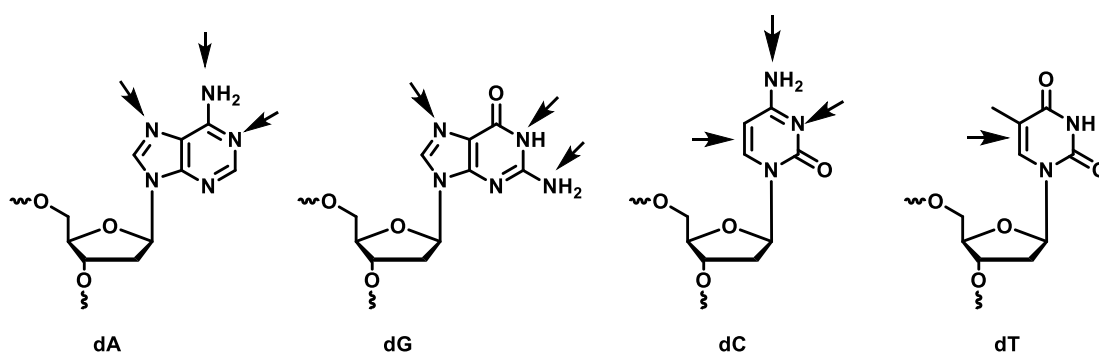
Figure 1.2 Potential reactive sites for the cross-link formation of natural nucleobases (dA, dG, dC and dT).

Nitrogen mustard and its analogs are powerful cross-linking agents that were originally developed as chemical warfare agents, but also find some uses for the treatment of various cancer diseases.[21] The active functionality is the *N,N*-bis-(2-chloroethyl)amine, which preferentially reacts with the nucleophilic N^7 position of guanine or adenine in the DNA, thus forming the ICL product and blocking the synthesis of new DNA.[10] There are several nitrogen mustard derivatives which have been or are being used for cancer therapy including chlorambucil,[22] melphalan,[23] bendamustine,[24] and cyclophosphamide, as shown in **Figure 1.3**.[25] Chlorambucil is the chemotherapy of choice for the treatment of chronic lymphocytic leukemia[22] and melphalan is being used for multiple myeloma.[23]
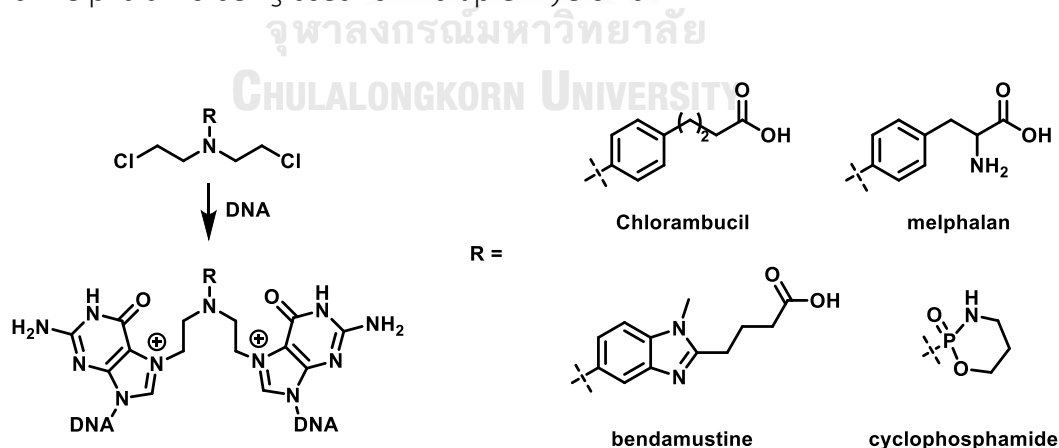
Figure 1.3 Structural and cross-linking mechanism of nitrogen mustard equipped with its derivatives used for cancer therapy

The platinum complex cisplatin is a commercially available drug approved by the FDA for the chemotherapy of advanced ovarian and bladder cancer since 1978. Cisplatin preferentially reacts with the nucleophilic N^7 of guanine to form both interstrand and intrastrand cross-linking products. Cisplatin derivatives, such as carboplatin, and oxaliplatin, constitute the second generation of anti-cancer drugs in this class in order to show a greater and more wide-spread activity against cancer with lower toxicity. They are used for various types of cancer: bladder, cervical, small cell lung cancer, lymphomas, and germ cell tumors.[26]
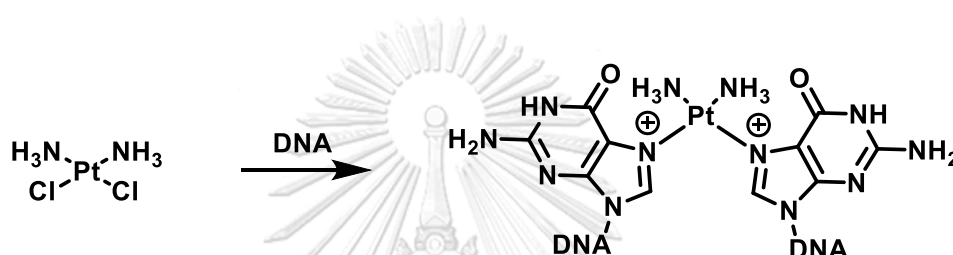
Figure 1.4 Structural and cross-linking mechanism of cisplatin

Psoralen[27] is a class of natural products known as furocoumarins. Upon activation by UV light, it preferentially cross-links with the thymine base in DNA via a [2+2] cycloaddition reaction to form a monoadduct between the pyrone or the furan ring. Then the second [2+2] cycloaddition reaction occurs with another thymine from the same or another DNA strand. Although it's being used for the treatment of various diseases including cutaneous T-cell lymphoma,[28] the non-selectivity of the cross-linking reaction leads to high risks of skin cancer associated with psoralen.[29]
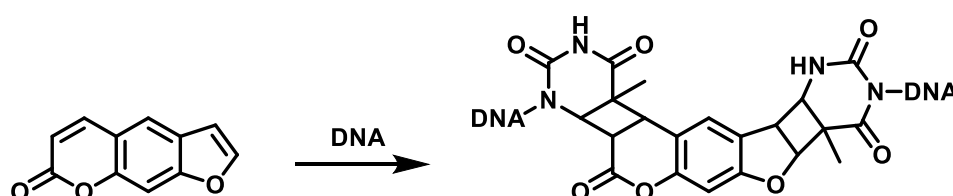
Figure 1.5 Structural and cross-linking mechanism of psoralen

Mitomycin C is another naturally occurring compound produced by *Streptomyces caespitosus*. It has an antibiotic and antitumor activity that has been widely used in chemotherapy against gastrointestinal tumors and anal cancer.[30] Mitomycin C preferentially forms cross-links with guanine via the exocyclic amino group. Although Mitomycin C is known as an interstrand cross-linking agent, an intrastrand cross-linked monoadduct had also been observed.[31]
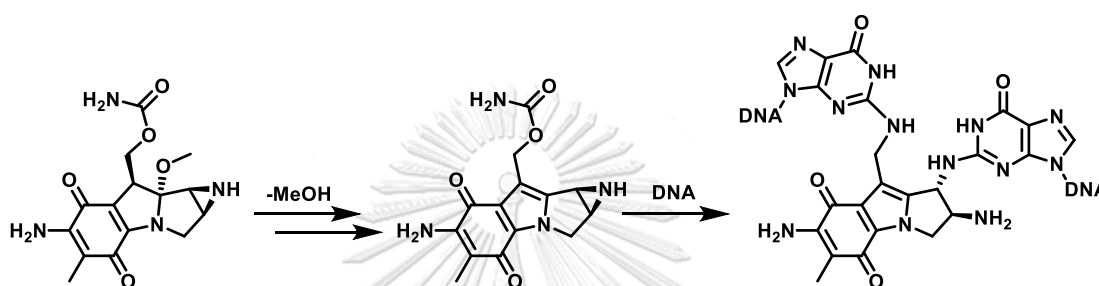
Figure 1.6 Structural and cross-linking mechanism of mitomycin C

Such DNA cross-linking agents have been developed and widely used in chemotherapy, but a major weakness remains: the poor selectivity between cancer and normal cells. The non-specific nature of the "off-target" cross-link formation can result in a high toxicity leading to the regrowth of drug-resistant tumor cells. To enhance the "on-target" cross-linking efficiency, many DNA cross-linking agents that can target specific DNA sequences have been developed by conjugating one or more reactive groups with a short DNA strand as a recognition element. These cross-linkable DNA probe can bind with the specific DNA sequence of interest and form the cross-links between the probe and the bound DNA target in a sequence-specific fashion. A desirable cross-linkable probe should be activable on demand, i.e. the reactive groups should be exposed only after the binding to the target to avoid self-destruction of the probe, non-specific cross-linking, and side reactions with other nucleophilic species in the biological systems.

1.3 Reactive oligonucleotides

1.3.1 Oligonucleotides modified with phenylsulfide/phenylsulfoxide nucleotides

In 1999, Nagatsugi et al. reported a phenylsulfide-masked 2-amino-6-vinylpurine modified oligonucleotide probe as a cytidine selective cross-linker.[32] The phenylsulfide protecting group could be converted to its sulfoxide form by an oxidation reaction. After the hybridization of the probe with the complementary DNA, the sulfoxide moiety was eliminated to form the corresponding 2-amino-6-vinylpurine probe that could subsequently react with the opposite cytidine. Moreover, the same research group reported a new 6-vinylpurine modified probe that could cross-link faster than the previous probe (2-amino-6-vinylpurine). The researchers explained that the absence of the electron donor group (amino group) increased the reactivity of the vinyl group.[33]

2-amino-6-vinylpurine **6-vinylpurine**
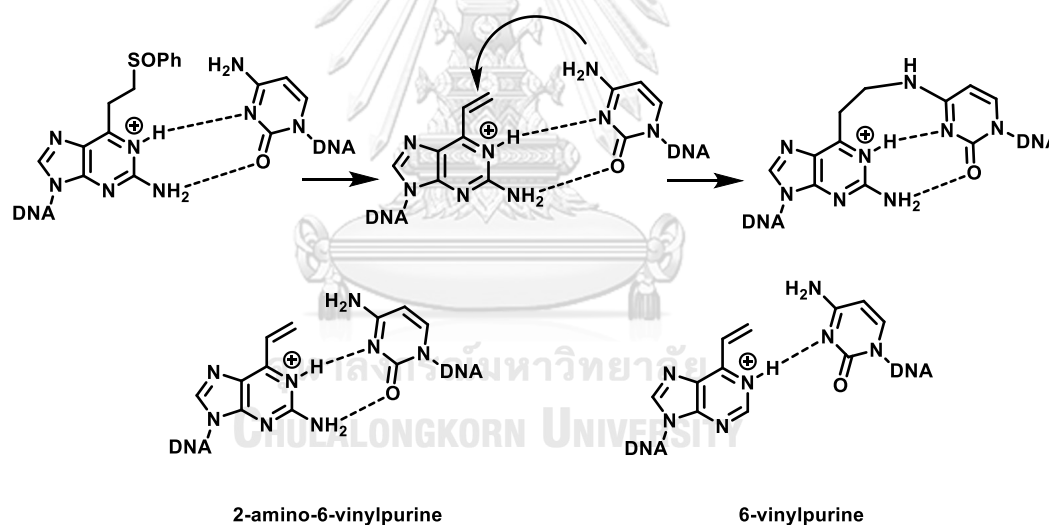
Figure 1.7 Cross-linking mechanism of 2-mino-6-vinylpurine and 6-vinylpurine toward cytidine DNA target

1.3.2 Oligonucleotides modified with phenylselenide

The phenylselenide group can be activated by a mild oxidation reagent ($NaIO_4$ or $^1O_2$) to create a reactive selenoxide group that can undergo a [2,3]-sigmatropic rearrangement to form a reactive methide intermediate which then forms cross-links with dA or dG on the opposite DNA strand.[34] In 2008, Peng and Greenberg synthesized a bifunctional oligonucleotide probe containing a phenylselenide-modified thymidine that could form cross-linking products with specific DNA targets under mild conditions.[35] After inducing the cross-linking between the biotin-labeled probe and its DNA target, the hybridization was fluorometrically detected by utilizing conjugates of avidin and horseradish peroxidase (avidin-HRP) in the presence of Amplex Red as a substrate.[15] The removal of the unbounded avidin-HRP from the biotinylated target was facilitated by the cross-linking reaction resulting in the reduction of the background signal. At the very least, 250 fmol of DNA could be detected without using PCR. Moreover, the discrimination ratio between complementary (dA) and single-base mismatched (dG) sequences was in the order of 200:1. Furthermore, the same research group applied a strategy to modify the phenylselenide moiety on the uridine of an RNA strand to act as a cross-linking probe targeting an adenine of the opposite RNA strand under mild conditions (H_2O_2, $NaIO_4$).[36] It was found that the cross-linking reactivity depended on the conformation of the selenide part because the glycosidic bond could rotate, and the cross-linking reaction preferred the C2′-endo conformation where the selenide group could fit in the duplex. This conformation helped the electrophile to preferentially cross-link with the opposite adenosine and cytidine because their N1 and N3 were more nucleophilic than the N1 and N3 of G and U.
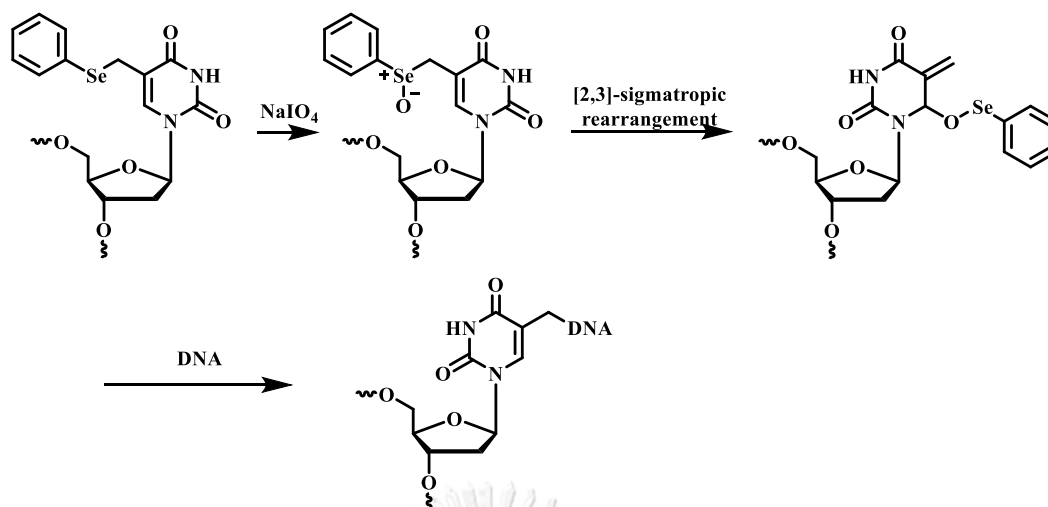
Figure 1.8 Cross-linking mechanism of phenylselenide-modified oligonucleotide towards DNA targets

In a related example, a benzyl phenylselenide modifier could be activated by the formation of a quinone methide (QM) when it was exposed to periodate oxidation.[37] QM is a reactive intermediate that can react quickly with nucleophiles. In this work, the benzylphenylselenide modifier on DNA was oxidized by $NaIO_4$. The cross-linking reaction occurred when the sequences of the DNA target and the probe were perfectly complementary. The selectivity of the cross-linking reaction with cytosine was 10 times higher than with purine bases.
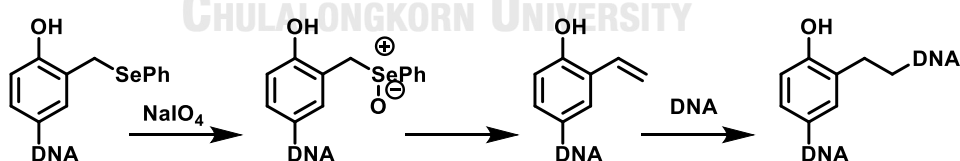
Figure 1.9 An example of quinone methide generation and the formation of the cross-linking adduct

1.3.3 Oligonucleotides modified with derivatives reacting through a [2 + 2] cycloaddition

The psoralen moiety can cross-link with the DNA strand via a [2+2] cycloaddition reaction but the selectivity is usually poor. Kobertz and Essigmann accomplished the synthesis of psoralen modified thymine and incorporated it into an oligonucleotide as a selective cross-linking probe. Psoralen is known to be a photo-reversible cross-linker.[38] The probe was irradiated at 366 nm to generate the cross-linked product and at 254 nm to reverse it. The cross-link product selectively formed with the thymine base of the DNA strand. Furthermore, Murakami et al. reported the synthesis of a 2′-O-psoralen-conjugated adenosine and incorporated it in the middle of an oligodeoxyribonucleotide. After UV irradiation, the psoralen-modified RNA probe preferred to cross-link with the uridine on the complementary RNA strand.[39]

To form a cyclobutane product, a [2+2] cycloaddition reaction must be initiated by exposing two alkene groups to light at a specific wavelength corresponding to the absorption of the unsaturated system. This reaction readily occurs with pyrimidine bases of DNA when exposed to UV light. Several cross-linkable DNA probes based on this principle have been reported. Haque and co-workers reported a cross-linkable probe carrying a coumarin-modified thymine analog in 2014,[17] which acted as a photo-reversible cross-linker which could be controlled by light switching (350 nm/254 nm) as shown in **Figure 1.10A**. The cross-linking reaction occurred upon irradiation with 350 nm light for 50 min, while the cross-linking was reversed by 254 nm irradiation during 6 min. The probe preferred to cross-link with thymine over cytosine and adenine.
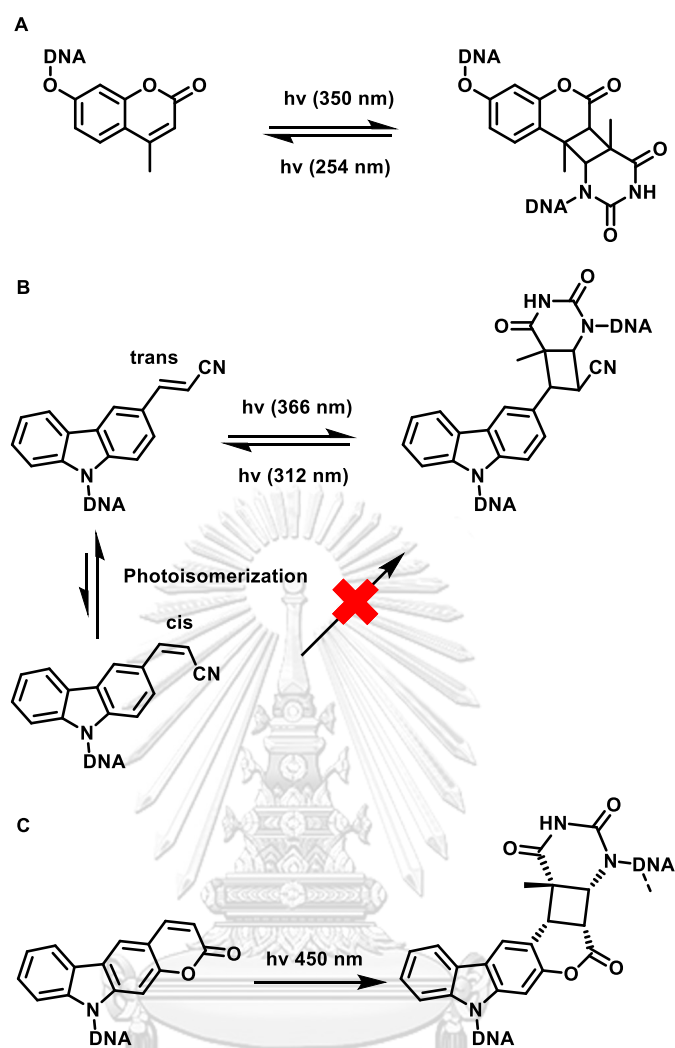
Figure 1.10 A [2 + 2] cycloaddition reaction of coumarin-modified DNA (A), 3-cyanovinyl-carbazole-modified DNA (B), and pyranocarbazole (C)

Relating to this concept, in 2008 Sakamoto and co-workers reported an oligonucleotide carrying 3-cyanovinylcarbazole modified nucleosides that could cross-link with its complementary DNA or RNA after UV irradiation.[40, 41] The fast and efficient cross-linking reaction of the 3-cyanovinylcarbazole-modified probe with a pyrimidine base on the opposite strand was achieved within a few seconds upon irradiation at 366 nm. Next, the cross-link adduct was broken when it was exposed to 312 nm light during 60s, as shown in **Figure 1.10B**. The probe preferred to cross-link with the pyrimidine base (T and C) similar to other previously discussed systems. The

scope of this strategy was expanded to RNA cross-linking.[42] The 3-cyanovinylcarbazole-modified oligonucleotide also acts as a cross-linking probe targeting the complementary RNA strand. After 366 nm irradiation, the probe selectively cross-linked with the fully complementary RNA strand in the presence of three other mismatched RNA sequences, providing a highly RNA selectivity. As expected, the probe only form cross-links with the uridine and cytidine bases. Recently, Fujimoto et al[43] reported a new pyranocarbazole photo-inducible DNA cross-linker as shown in **Figure 1.10C**. The cross-linker was placed in the middle of the DNA oligonucleotide strand targeting an opposite pyrimidine base of the complementary DNA. When activated by a 450 nm light, the probe cross-linked with pyrimidine base via a [2+2] cycloaddition reaction. The probe preferred to cross-link with the adjacent thymine base (90% yield) over the cytosine base (20%) within 5 min. The extended π-conjugate system of the pyranocarbazole allowed using visible light to induce the cross-linking. This is considered advantageous since it reduces the phototoxicity to the cells.

Another example of cross-linking via a [2+2] cycloaddition involving *p*-stilbazole analogs as an artificial base was described by Asanuma's group.[18] The cross-linking reaction of two *p*-stilbazoles between two complementary DNA strands occurred when the two π-π stacked chromophore were activated by a light of 340 nm (**Figure 1.11A**). The cross-linking reaction was completed within 5 min resulting in the formation of an artificial base pair in the DNA duplex. The same research group subsequently reported a new photo-switchable cross-linkable styrylpyrene as an artificial base.[44] The cross-linking reaction between the two styrylpyrene molecules inserted in the middle of two oligonucleotide strands via a [2+2] cycloaddition reaction was triggered by 455 nm light irradiation. The reverse reaction was quantitatively induced by irradiation with 340 nm light (**Figure 1.11B**).

Moreover, the styrylpyrene functionality was also modified on the adenine base at the C8 position. The styrylpyrene modified adenine was incorporated into the serinol nucleic acid (SNA) to act as a photo-reversible cross-linker targeting the complementary RNA strand.[45] Two styrylpyrene modified adenines were placed

adjacently in the middle of the SNA strand. The probe exhibited significant fluorescence in the free state and when bound to its complementary RNA. When triggered by 455 nm light irradiation, the intrastrand cycloaddition reaction between the two adjacent styrylpyrene units was induced. The change in the DNA conformation resulted in loss of fluorescent properties and dissociation of the duplex between the DNA probe and the RNA target. The reverse reaction could be achieved by 340 nm light irradiation. Thus, the reversible light-induced cross-linking reaction can be used to control the hybridization state of the probe and the RNA target strand employing light as the external stimulus. In 2020, the same research group extended the same strategy to create a dual cross-linking photo-switchable probe that was modified by 8-naphthylvinyl adenine and 8-pyrenylvinyl adenine residues.[46] These two residues were placed on the SNA strand that formed a duplex with the target RNA strand. The hetero [2+2] cycloaddition was triggered by 405-465 nm irradiation, and by less than 340 nm for the reversal.
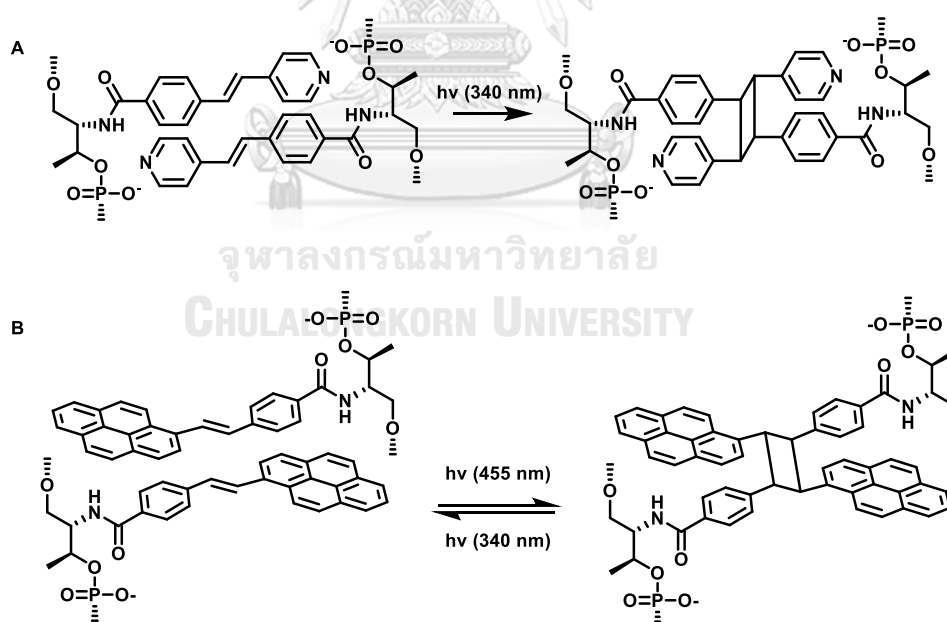
Figure 1.11 Cross-linking mechanism of *p*-stilbazole modified oligonucleotides as an artificial base (A) and styrylpyrene modified oligonucleotides via [2+2] cycloaddition reaction (B)

1.3.4 Oligonucleotides modified with furan

Furan is a relatively stable small aromatic compound found in nature as a decomposition product of organic matters such as thermal degradation of smog and cigarette smoke. It is known as a liver toxicant of rats and mice and is also genotoxic.[47] Its toxicity to humans has been recognized,[48, 49] notably as a carcinogen in 2B group by the International Agency for Research on Cancer (IARC).[50] However, kinetin or furfuryl-modified adenine is a compound generated from the degradation of natural DNA under oxidative stress, that is beneficial for antiaging or antiplatelet aggregation.[51] Furthermore, it is known that furan can generate a highly reactive aldehyde species upon being oxidized by the cytochrome P450 family of enzyme (**Figure 1.12**).[52] The initially formed aldehyde species can react with many nucleophiles, such as amino side chains of proteins or exocyclic amino groups of nucleobases.[53, 54] Similar oxidation of furan leading to the same aldehyde species can be achieved in vitro via *N*-bromosuccinimide (NBS) (**Figure 1.12**) or by using singlet oxygen (**Figure 1.12**). Inspired by the known reactivity of furan, Madder and colleagues have explored the possibility of using furan as a masked reactive cross-linker.[55] Thus, several furan-based cross-linkable oligonucleotide probes have been reported.
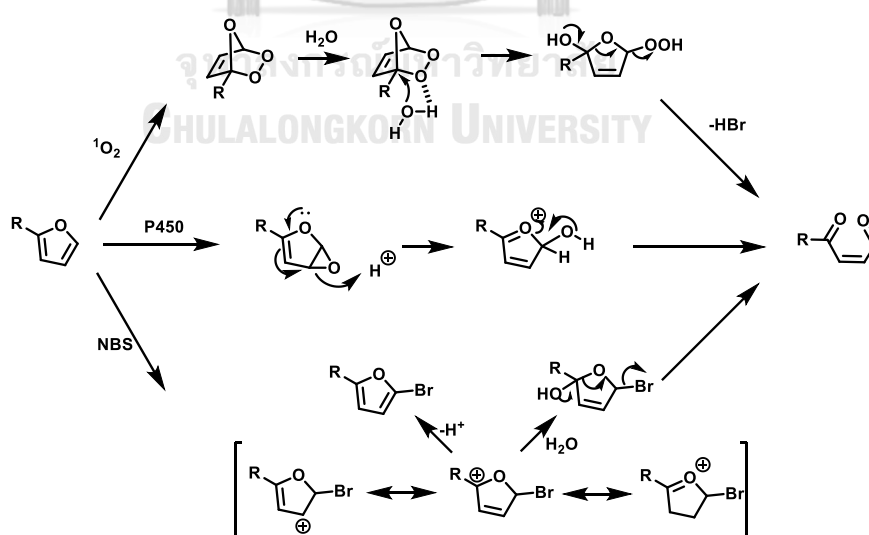
Figure 1.12 Overview of furan oxidation mechanism via $^1O_2$, NBS, and Cytochrome P450.

In 2005, Halila and co-workers published the first generation furan-modified oligonucleotide probes capable of producing the cross-linking reaction with unmodified nucleic acid targets upon being oxidized by NBS.[55] The furan building block **2** (**Figure 1.13**) was introduced into the oligonucleotide strand, and the variations of the nucleic acid targets such as a flanking and opposing nucleobase, allowed the group to study the selectivity of the furan-modified probe. NBS was used as an activator and it was found that 4 equiv. of NBS were necessary to activate the probe. The furan-modified probe selectively cross-linked with adenine and cytosine via their exocyclic amino groups. The molecular modeling of the B-form of the DNA-DNA duplex revealed that the furan moiety was located in the major groove of the duplex. Molecular modeling suggested that adenine (opposite strand) and its neighboring cytosine (5′- terminus) instead of the opposite cytosine itself were the most favorable ligation sites.

Carrette and coworkers from the same research group studied the new furan modified building block **1** (**Figure 1.13**) by modifying the furan moiety on the natural T-base and inserting it into the middle of the DNA strand, thus creating a cross-linkable probe targeting the lysine residue of the protein.[56] The results revealed that the duplex stability of the furan-modified building block with its complementary strand containing adenine as the opposite base was similar to that of the unmodified strand due to the favorable base stacking formation. The study found that the reactive aldehyde species which was generated by NBS oxidation of the furan moiety could react with a lysine residue in the general control protein (GCN4). Adding the $NaBH_3CN$ enhanced the formation of the DNA-protein cross-linked product by irreversibly converting the Schiff base into the corresponding amine. In the absence of the nucleophile, the NBS-oxidized probe rapidly degraded.

A study of the influence of the linker was carried out by changing the amide linker (based on the furan building block **2** from the Halila's work) to an ureide linker in the furan building block **3**.[57] It was found that the oligonucleotide modified with the furan building block **3** formed a more stable duplex with its complementary DNA (A at the opposite) than the furan building block **2**. They explained that the extra

nitrogen atom may catalyze the dehydration reaction, and thus activating the keto functionality of the aldehyde species. The building blocks **2** and **3** could form cross-linked products with C more preferentially than with A because the A base formed the base pairing with the U-modified building blocks, and thus was unavailable to react with the furan.

In addition, the furan moiety was incorporated into adenosine as a pendant group (**4** in **Figure 1.13**) by Jawalekar and coworkers in 2011.[58] In this work, *N*-iodosuccinimide (NIS) was used as an alternative activator to NBS. The reactivity of the NIS activation was better than NBS in terms of stoichiometric and fast ICL formation. The building block **4** was placed between two guanine bases to form a strong base dyads that forced the activated furan moiety to interact with the mismatched A or C opposing bases. In contrast to other systems, it was found that the furan building block **4** preferentially cross-linked with A over C. This result could be explained by the stability of the duplex that helped the mismatched A to cross-link with the A-modified building block **4**.
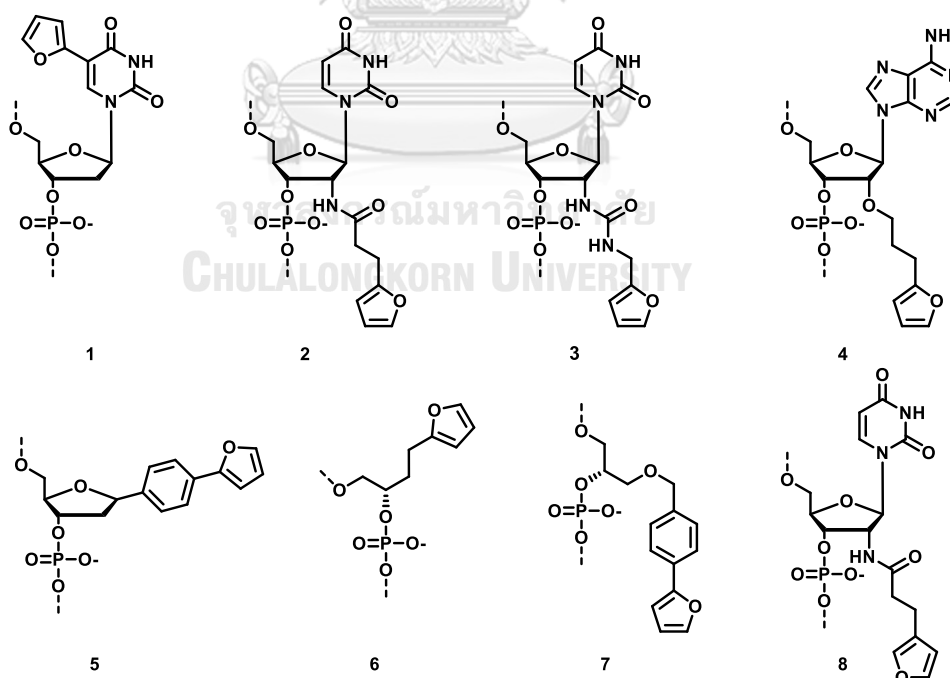
Figure 1.13 Furan modified nucleosides and nucleoside analogues published by Madder's group.

The second generation furan-modified probes was subsequently developed by the same research group in 2009.[59] These furan-modified building block does not carry a nucleobase that can form a proper base pairing as in the cases of furan-modified uridine or adenine described previously. The influence of the duplex stability towards the cross-linking reaction was studied by introducing simple acyclic furan building blocks **6** and **7** (**Figure 1.13)** into the middle of the oligonucleotide strand. In these cases, the furan was incorporated as a base surrogate that could not form a base pair with the opposite nucleobase in the DNA strand. When the furan building block **6** was introduced into the oligonucleotide strand, it was found that only the base located opposite to **6** could be involved in the cross-linking reaction. Its strong selectivity with the complementary A or C has been demonstrated. The G-base did not form a cross-link with this probe due to its bulky nature which could not reach the furan moiety in the major groove according to a molecular modeling data. As expected, the building block **7** bearing an extra aromatic ring enhance the stacking interaction with the complementary nucleobase, and thus displayed a higher duplex stability and cross-linking efficiency than the building block **5**. Moreover, the building block **7** revealed a fast and high yielding selective cross-linking with adenine and cytosine (up to 73% and 49% for C and A respectively) upon NBS activation. Importantly, the less stable duplexes deriving from oligonucleotide containing the building block **6** showed lower cross-linking efficiency when compared to the more stable duplexes of oligonucleotide containing the building block **7**.[60] Later, in 2014, the same research group developed cross-linkable furan-modified RNA probes (building block **2** and **7**) [61] to target its complementary RNA. The results showed that the furan-modified nucleoside selectively cross-linked with cytidine in a 42% yield by employing *N*-bromosuccinimide (NBS) as an activator.

Since the use of NBS to activate the furan moiety cannot be performed in living cells, red light was employed as an alternative way for the oxidation of furan to generate the aldehyde species. In 2012, Beeck and Madder reported a cross-linking study of yet another furan-modified oligodeoxynucleotides with DNA. In this work, the same furan building block **7** (**Figure 1.13**) was employed as in the previous study,

but the furan-modified probe was activated by red light (620–750 nm) in the presence of methylene blue as a photosensitizer.[62] This biocompatible oxidation method enabled the cross-linking reaction with a nearby A or C base in the opposite DNA strand in 13%-57% yields.

Furan-modified oligonucleotide was also used for the cross-link of double-stranded DNA through a triplex formation, as reported by Gyssels and coworkers.[63] The furan building block **7** was introduced at the internal position and the 3'-end of the DNA strand to investigate the cross-linking efficiency towards double stranded DNA upon NBS activation as shown in **Figure 1.13**. It was observed that the internally-modified probe could form the cross-linked product with single stranded DNA but not with double stranded DNA targets. The results were explained in terms of the unavailability of the exocyclic amino groups in the nucleobases that participated in the Watson-Crick base pairing with its DNA complementary strand. The cross-linking would require breaking the base pairs, and thus an energetically unfavorable process. The hypothesis was supported by the cross-link formation after heating to destroy the base pairing. On the other hand, the cross-linking reaction took placed smoothly in the case of the 3'-furan modified probe that has free-bases (3' terminus) to react with.

	TFO1:	5'-CT CTT CC**7** CTC TTT C-3'		TFO3:	5'-CT CTT CCT CTC TTT **7**-3'
ON1:	5'-GCG CGA GAA GG**A** GAG AAA GCC GG-3'		ON1:	5'-GCG CGA GAA GGA GAG AAA GCC **G**G-3'	
ON2:	5'-CGC GCT CTT CC**T** CTC TTT CGG CC-3'		ON2:	5'-CGC GCT CTT CCT CTC TTT CGG **C**C-3'	
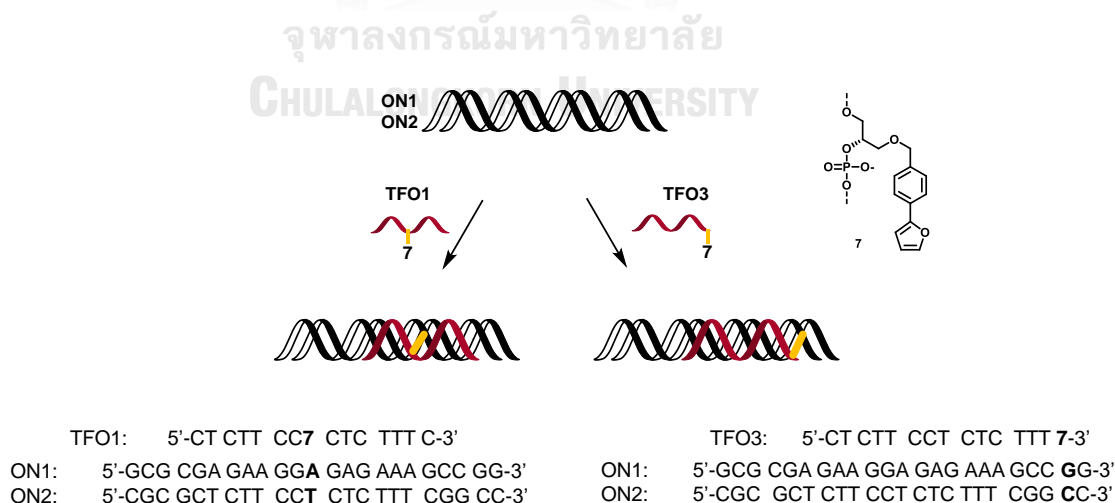
Figure 1.14 Illustration of cross-linking reaction of furan modified oligonucleotide towards double strand DNA duplex studied by Gyssels and coworkers[63]

In terms of applications, the furan-modified DNA probes have been used to improve the sensitivity of DNA detection via the cross-link formation. For example, a biotinylated furan-modified DNA probe was immobilized on the neutravidin (NA)-coated on a gold chip.[64] Following the DNA target hybridization, the probe was oxidized with singlet oxygen generated by red light (658 nm) irradiation in the presence of methylene blue as a sensitizer to induce the cross-linking. Then the FITC-modified DNA target was detected by SPR after treatment with an HRP-labeled anti-FITC antibody. The cross-link enhanced the duplex stability and allowed extensive washing with Na_2CO_3 which had a strong anion effect and would normally destroy the non-cross-linked duplex.[65] Thus, this method provided a high immobilization efficiency and enhanced the stability of the short oligonucleotide duplex to increase the efficiency of the DNA detection.

1.3.5 Cross-linking between two complementary modified oligonucleotides through Click chemistry

The copper-catalyzed 1,3-dipolar cycloaddition reaction of a terminal alkyne with an azide was popularized by Sharpless since 1988 under the name of "Click chemistry".[66] This reaction is high yielding, fast, and tolerant to a large variety of functional groups. Shortly after its introduction, the Click chemistry has been expanded into the biological fields, including the DNA cross-linking. In 2008, Brown's group synthesized modified-uracil analogs with different linkers containing alkyne and azido groups and incorporated each of them separately into two complementary DNA strands as shown in **Figure 1.15**.[67] In the presence of a Cu(I) catalyst, the cross-linking reaction was completed within 5 min - 2h and cleanly afforded the ICL product in >90% yield. Moreover, the cross-linking reaction via Click chemistry was applied to DNA ligation. The bases of the two oligonucleotide strands were modified with alkyne moiety which could react with bifunctional bis-azide cross-linkers.[68-70] Although the Click chemistry is very useful in the biological field, the cytotoxicity of copper ion remains an important obstacle for in vivo applications.
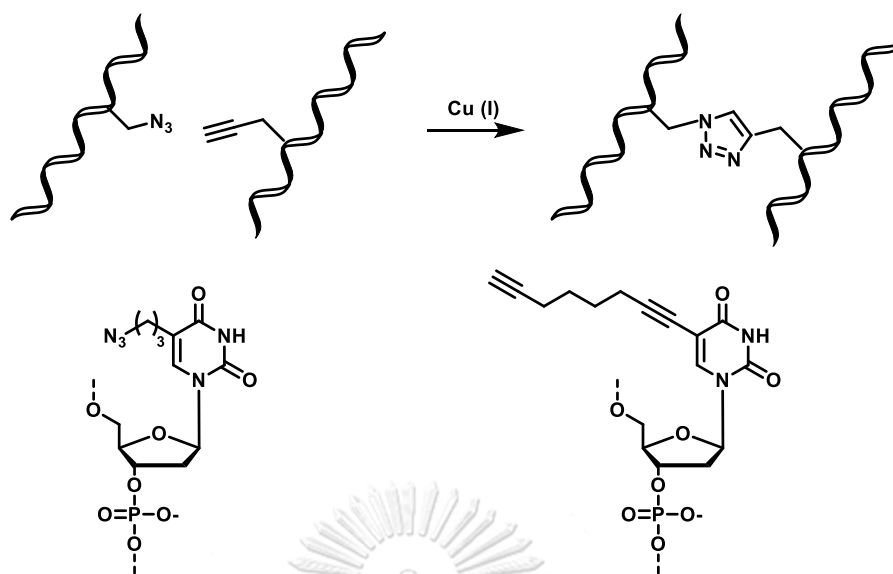
Figure 1.15 Schematic illustration of oligonucleotide cross-linking reaction via Click-chemistry

1.4 Biological roles of ICL and its repair

The change in the chemical structure of the DNA caused DNA damage that can affect its function in the cells. The damage can be divided into two main types: endogenous damage and exogenous damage. The endogenous damage is induced by reactive oxygen species (ROS) while exogenous damage is induced by hydrolysis, radiation (UV, X-ray, gamma) and viruses.[71] Some common DNA-damaged lesions are shown in **Figure 1.16** including damaged base, mismatch, abasic site (apurine/apyrimidimic or AP site), single strand break (SSB) or double strand break (DSB), and ICL adduct. In nature, cells can avoid these damages through various DNA repair processes which encompass base excision repair (BER), nucleotide excision repair (NER), and homologous recombination (HR).
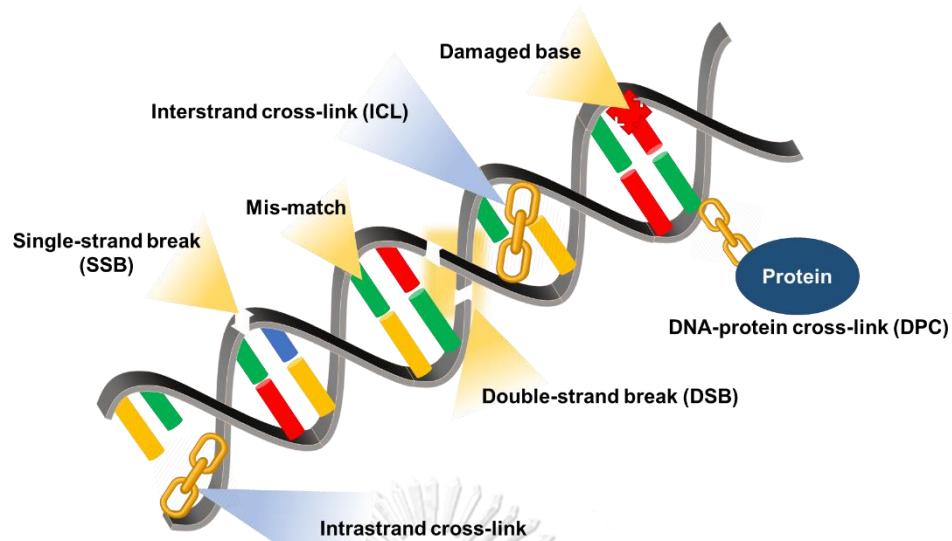
Figure 1.16 Examples of DNA damage consisting of mismatched base, damaged base, single-strand break (SSB), double-strand break (DSB) and cross-linking adduct

However, the DNA repair process involving ICL is more complicated than other DNA damaged lesions because the ICL process simultaneously damages both strands of the DNA and thus stalling the replication process. Accordingly, this process is highly damaging to the cells and can have many potential biological consequences. In general, the ICL repair process consists of four steps involving the recognition of the ICL part, the incision of the lesion, the removal of the ICL part and patching the strand. The first proposed mechanism is the recombination-independent mechanism (**Figure 1.17A**).[72-74] According to this mechanism, a nucleotide excision repair (NER) cut out one strand of the duplex at the phosphodiester bonds on either side of the lesion, generating a gap in the DNA duplex. Then the gap was repaired by translesion DNA polymerases (TLS) and the process was repeated with another strand. However, the proposal of this mechanism was based on bacteria studies, which may not apply to the case of mammalian cells. Furthermore, some indirect evidence suggested that in metazoans, the ICL repair occurred in the S-phase of the cell cycle,[75-77] while the NER pathway involved in this mechanism had a role in the G_0/G_1 phase. Therefore, this mechanism was declared as a repairing process occurring in the G_0/G_1 phase in the cell cycle.[78]

The second proposed mechanism is the recombination-dependent mechanism which is typical for the S-phase of cell cycle. This mechanism involved the Fanconi anemia, a cancer predisposition syndrome.[2] Because the Fanconi anemia gene products are found to resistance to ICL-inducing agent, thus it is used as a model disease to understand the ICL repair.[79] The FA repair mechanism involved three main steps, as shown in **Figure 1.17B**. The ICL repair started when the cell recognizes the replication fork from the ICL adduct in the S-phase of the cell cycle that stalling the replication. The damaged recognition was noticed by the signal from the combination of the FA protein with other proteins. Then, an endonuclease cut the damaged base adjacent to the stalled replication fork, which is called "unhooking". The translation synthesis polymerase, NER and homologous recombination (HR) factors were then able to complete the repair process.

Although the FA mechanism is generally accepted,[2, 79, 80] some studies provided an additional mechanism that helped cell survival might also be operating in the ICL repair. The mechanism involved a base excision repair (BER) pathway. The studies found that a XRCC1 protein, the base excision protein, participated in the removal of the damaged base that was bound to the cisplatin monoadduct.[81] The BER pathway was divided into four main steps as shown in **Figure 1.17C**.[82] First, DNA glycosylase recognized and released the damaged base by breaking the deoxyribose-phosphate backbone and flipped the damaged base out of the duplex, leaving the AP site. Second, the AP endonuclease cleaved the phosphodiester bond adjacent to the AP site creating a 3'-hydroxyl priming group and a 5'-phosphate. Third, the DNA polymerase beta (pol β) synthesized the missing nucleotide. Then, the phosphodiester bond was cleaved to generate the substrate for DNA ligase I. Finally, the DNA ligase sealed the nick to complete the repair process.[83]

Although some repair mechanisms for ICL exist, the process is still considered highly damaging to the cell. Hence, the failure of the cell to repair itself from the ICL can be useful for chemotherapeutic treatment.
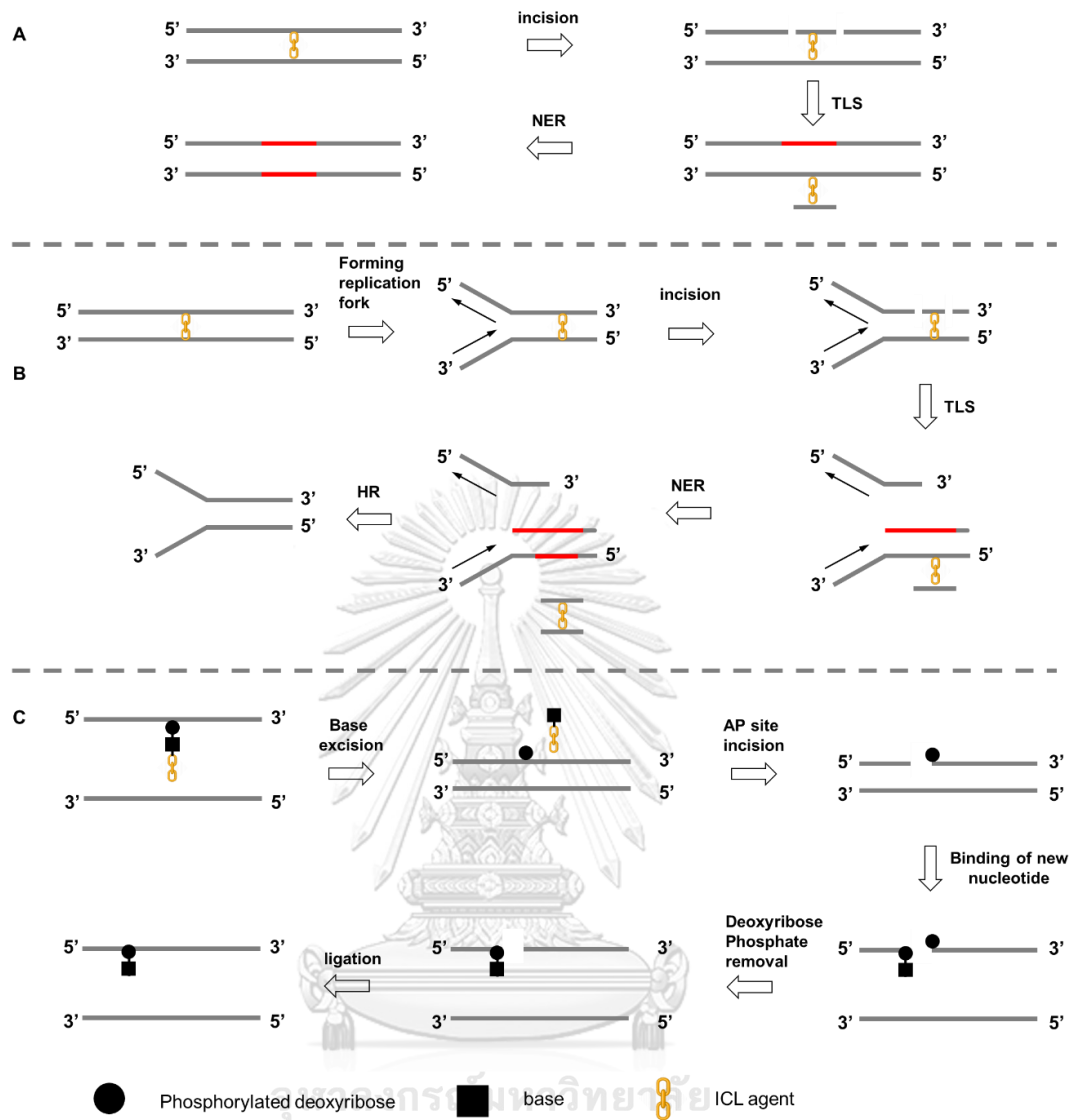
Figure 1.17 The ICL repair mechanism including nucleotide excision repair (NER) (A), Fanconi aneamia (FA) pathway (B), and base excision repair (BER) pathway (C).[79]

1.5 ICL involving PNA

Although DNA/RNA-modified oligonucleotides are typically employed as the recognition probes for the ICL process, their poor cellular stability, and sub-optimal hybridization efficiency and specificity remain an issue. Various strategies to enhance the stability of the DNA duplex and increase the efficiency of the ICL process have been proposed,[18, 20, 44-46, 57, 64, 84] one of which is the use of alternative oligonucleotide probe such as peptide nucleic acid (PNA). PNA is a synthetic nucleic acid mimic[85] that can bind to DNA/RNA with exceptionally strong affinity and high sequence specificity according to the Watson–Crick base-pairing rules. PNA contains an entirely neutral charged peptide backbone instead of the anionic phosphate backbone in DNA. Therefore, it can strongly bind with DNA without the unfavorable electrostatic repulsion.[86] The unnatural backbone of the PNA also contributes to the greater chemical and physiological stabilities than natural DNA or RNA, which are reflected in the PNA's resistance towards nuclease and protease.[87-89] The first PNA system was reported by Nielsen et al. in 1991. It consists of an *N*-(2-aminoethyl) glycine backbone (hence being named aegPNA) instead of the deoxyribose phosphate backbone of natural DNA.[85, 90] This PNA system showed strong binding affinity with DNA in both parallel and antiparallel directions. Over the past three decades, numerous new PNA systems have been developed to enhance the stability of the PNA-DNA duplex as well as to improve its directional preference (parallel and antiparallel) and selectivity of the binding toward DNA/RNA targets.[91] Examples of such PNA modifications include the pyrrolidine oxy-PNA,[92] pyrrolidinyl PNA,[93] aminoethylprolyl peptide nucleic acids (aepPNA),[94] cyclohexylPNA,[95] γPNA,[96] etc.

In 2005, Vilaivan and co-workers[93, 97, 98] presented a new pyrrolidinyl peptide nucleic acid (acpcPNA)[99] consisting of a nucleobases-modified proline (monomer) and a 2-amino-1-cyclopentanecarboxylic acid (ACPC, spacer) as shown in **Figure 1.18**. This conformationally constrained structure of acpcPNA allows it to bind with DNA only in the antiparallel direction similar to the natural DNA-DNA duplex. Moreover, the rigidity of acpcPNA structure leads to a higher thermal stability with its complementary target with improved specificity, thus allowing more accurate distinction between complementary and mismatched DNA targets than aegPNA.
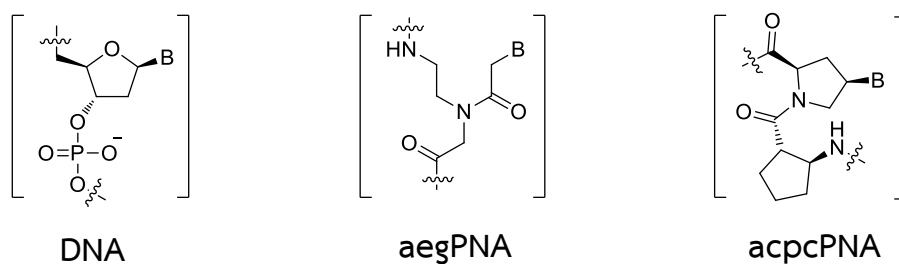
Figure 1.18 Structural comparison of DNA, aegPNA and acpcPNA

Few examples of cross-linkable PNA probes have been reported in the literature. In 2016, Manicardi and co-workers synthesized furan-modified aegPNA probes and demonstrated their ability to cross-link with DNA using NBS as an activator. The furan moiety was placed in the middle of the PNA strand as a base surrogate (PNA1 and PNA4 in **Figure 1.19**), and as a modifier on the base T (PNA7 in **Figure 1.19**).[100] After hybridization with the DNA target and subsequent activation by NBS oxidation, the furan-modified PNA cross-linked preferentially with adenine and cytosine, as determined by denaturing PAGE, melting temperature (T_m) and HPLC analyses. The study showed that the PNA4, with a long linker directly attached to the PNA backbone, displayed a stronger reactivity towards cytosine, while PNA1 favored cross-linking with adenine. The longer spacer of PNA4 provided a higher duplex stability than PNA1 due to forming of π-π stacking of aromatic ring with nucleobase resulting forcing the furan moiety into the duplex and enhancing the cross-linking efficiency. With the PNA7 that can form a stable base pair with the base A on the opposite DNA strand, no ICL products were observed for both A and C, confirming the results from the studies of cross-linkable DNA/RNA probes that the nucleobases involved in the Watson–Crick base pairing cannot participate in the cross-link formation. In addition, the furan ring attached to the 5-position of the uracil base in the PNA7 was placed in the major groove of the duplex, which may have been too far away from the exocyclic amino group as the nucleophile in the complementary DNA strand.

In 2017, the same research group[101] developed new γ-L-lysine and γ-L-arginine furan-modified PNA monomers and incorporated them in the aegPNA

sequence in order to improve the efficiency of the cross-linking as shown in **Figure1.19**. A new method for incorporating the acid-sensitive furan moiety into the PNA oligomer via a transient protection with maleimide was introduced. It was found that the introducing of only one residue of γ-L-lysine or γ-L-arginine furan-modified PNA monomer into the PNA strand did not affect the duplex stability when compared to unmodified sequence. The result showed that the γ-L-lysine modification showed a higher cross-linking efficiency than the γ-L-arginine modification because of the combination of an appropriate orientation of the helix and a favorable electrostatic interaction between the positively charged lysine and the negatively charged DNA backbone. Although, the L-arginine modification had a stronger positively charge than the lysine, the cross-linking efficiency was less effective probably due to a greater steric effect.
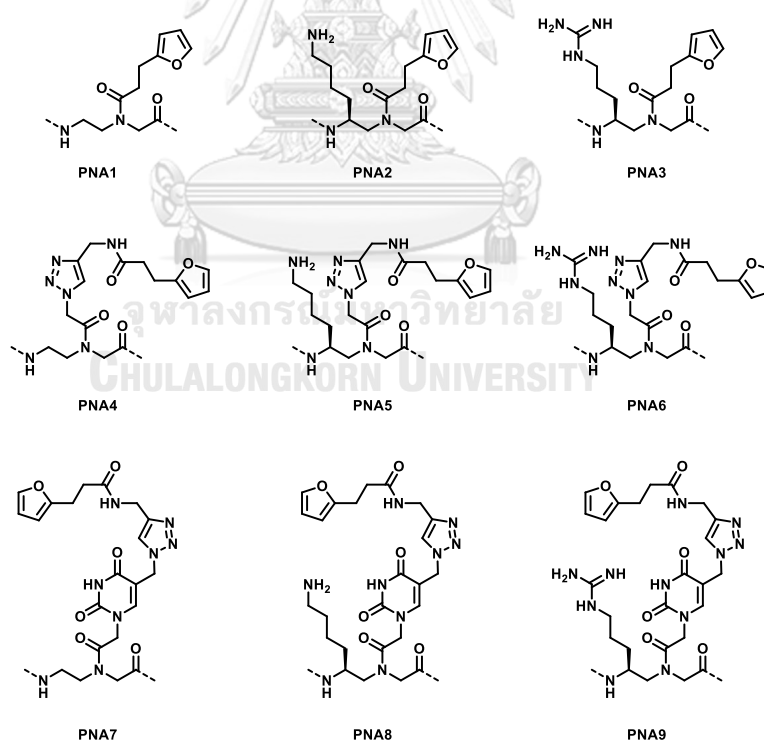
Figure 1.19 Monomers of furan modified PNA probes[100, 101]

1.6 Rationale and objective of this work

Following the successes of the use of furan-modified DNA, RNA, aegPNA and γPNA as cross-linkable probes with DNA and RNA, this research aims to develop furan-modified acpcPNA probes as reactive functional PNA probes that can be activated on-demand by treatment with *N*-bromosuccinimide (NBS). Since the efficiency and selectivity of the ICL process are very important for most applications, ideally, the ICL should occur selectively when the probe forms a specific base-pairing hybrid with the DNA target. In this respect, we propose that the high affinity and specificity of acpcPNA should improve the performance of such probes over the previously reported ones. From previous reports, most of the furan modifications were placed within the PNA strand and a base mismatch or an abasic site was required in order to allow the cross-linking to occur. This in turn decreased the stability of the duplex and thus created an issue. It was proposed that the high affinity of acpcPNA would also solve this problem.

The study is divided into two parts: in the first part, the furan modified acpcPNA is synthesized to investigate the cross-linking efficiency towards the complementary DNA target. Two kinds of PNA probes are included in the study – acpcPNAs with the furan modification at the terminal [PNA(T)] and internal [PNA(I)] positions of the PNA strand. The probes are designed so that all base pairs are fully formed. The cross-linking efficiency is then evaluated by various techniques including T_m, gel electrophoresis, HPLC and MALDI-TOF MS. The parameters to be investigated include: the position of the furan moiety on the PNA strand both internal [PNA(I)] and terminal modifications [PNA(T)]; the reactivity of the nucleobase (A, T, C, G) towards furan cross-linking reaction; the influence of the distance between the nucleobase and the furan moiety; the selectivity of the furan probe towards mismatched sequences.

In the second part, the comparative studies of the cross-linking reaction between the furan modified acpcPNA and the furan modified regular PNA (aegPNA) towards the complementary DNA target are investigated and compared. As in the first part, two kinds of furan modified PNAs within internal and terminal positions are

studied. Three furan building blocks (f, F, T) are modified at the end of the PNA strand, and two building blocks (F, T) are studied further with the internal probes. The PNAs sequences, studied in this part, are chosen from previous works of Madder's group. The cross-linking efficiency is assessed by gel electrophoresis, HPLC and MALDI-TOF MS. The parameters to be investigated are similar to the ones of the first part but the conditions are different. In order to compare the reactivity side by side, the same cross-linking protocol and the detection condition will be adopted.

We expect that this present research will develop new cross-linkable pyrrolidinyl PNA probes that display high affinity and specificity for DNA targets as well as to provide some further insights into the ICL formation involving PNA probes.
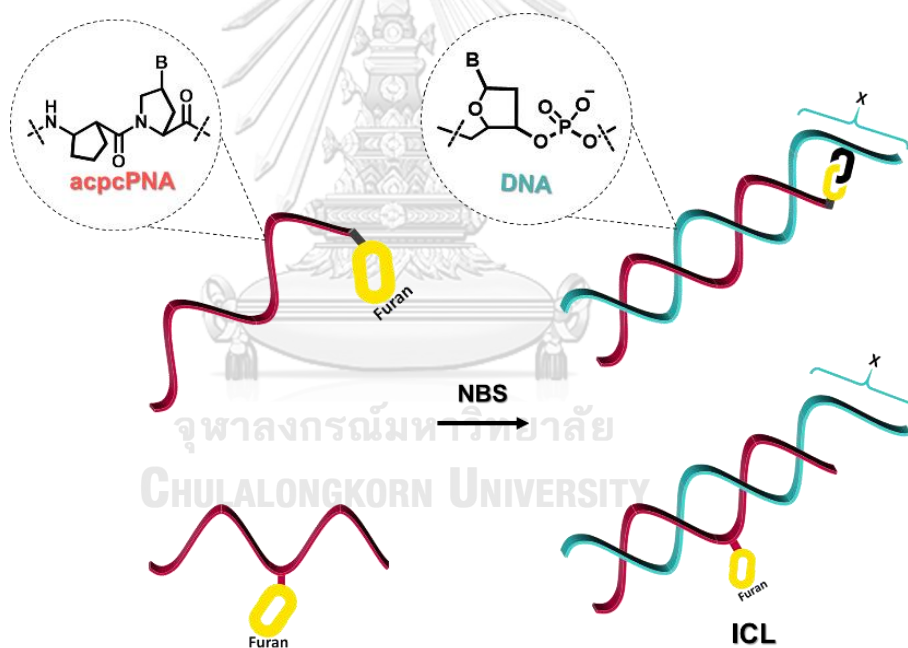
Figure 1.20 Illustration of the cross-linking reaction of PNA(I) vs PNA(T) towards DNA target

CHAPTER II

EXPERIMENTAL

2.1 Section I (Experiments performed at Chulalongkorn University)

2.1.1 General

All of the chemical reagents involved in the experiments were purchased from standard suppliers (Fluka, Acros, TCI, and Thermo Fisher) and used without further purification. Fmoc-Lys(Mtt)-OH was perchased from Novabiochem®. Fmoc-AEEA-OH was purchased from TCI. Anhydrous N,N-dimethylformamide (<50 ppm H_2O) used for the solid phase peptide synthesis was purchased from RCI Labscan (Thailand) and dried over activated 4 Å molecular sieves before use. DNA oligonucleotides were purchased from Pacific Science (Thailand) and BioDesign (Thailand). Milli-Q water was obtained from a Milli-Q® Reference water purification system (type 1) equipped with a Millipak® 40 filter unit 0.22 μm, Millipore (USA). T_m experiments were performed on a CARY 100 Bio UV-Visible spectrophotometer equipped with a thermal melt system. The melting curves were recorded at 260 nm at a heating rate of 1 ºC/min. CD spectra were recorded on a Jasco J-715 CD spectropolarimeter. ^1H and ^{13}C NMR spectra were recorded in the deuterated solvents on Varian Mercury-400 plus operating at 400 MHz (^1H) and 100 MHz (^{13}C), and JEOL JNM-ECZ500R/S1 operating at 500 MHz (^1H) and 125 MHz (^{13}C).

2.1.2 Synthesis of acpcPNA

2.1.2.1 Monomer synthesis

Three Fmoc-protected, Pfp-activated pyrrolidinyl PNA monomers (Fmoc-A^{Bz}-OPfp, Fmoc-C^{Bz}-OPfp, Fmoc-G-OH, and (1S,2S)-2-amino-1-cyclopentanecarboxylic acid (ACPC) spacer, 3-aminopyrrolidine-4-carboxylic acid (APC) spacer and Fmoc-D-proline were synthesized by Dr. Boonsong Ditmangklo (Fmoc-ACPC-OPfp, Fmoc-APC-OPfp, Fmoc-D-proline-OPfp), Mr. Nuttapon Jirakittiwut (Fomc-A^{Bz}-OPfp, Fmoc-C^{Bz}-OPfp) and Mr. Chayan Charoenpakdee (Fmoc-G-OH). The Fmoc-T-OPfp monomer was

synthesized as described below. The protocol for the synthesis of all PNA monomers was adapted from previous publications.[102]
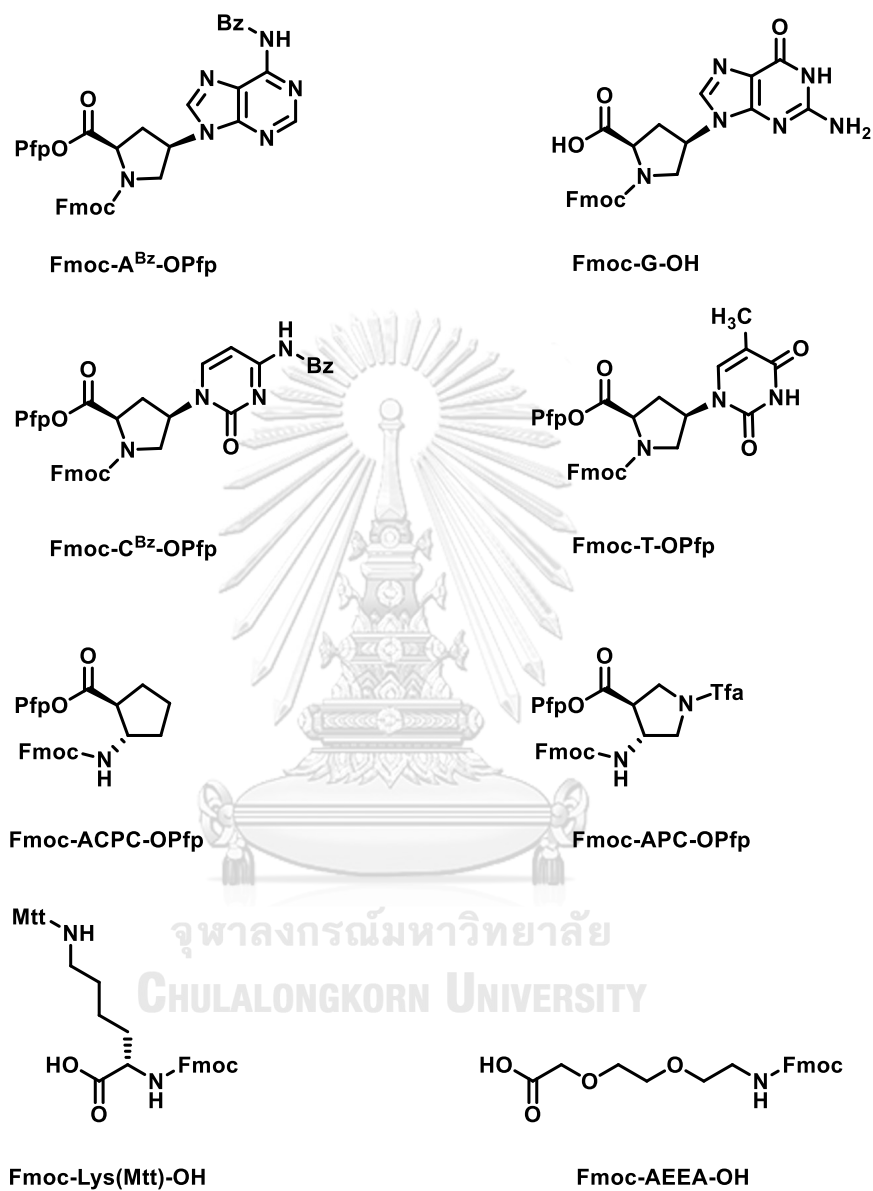
Figure 2.1 The structure of monomers and spacers used for the solid phase synthesis of acpcPNA

2.1.2.1.1 Fmoc-T-OPfp

cis-4-Hydroxy-D-proline[102]
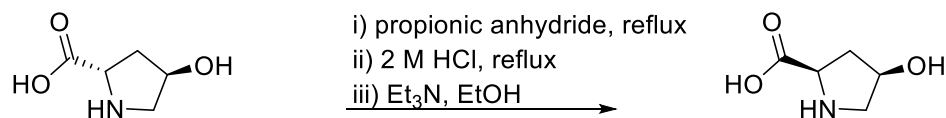
trans-4-Hydroxy-L-proline (10 g, 76 mmol) was suspended in propionic anhydride (50 mL) in a 250 mL round-bottomed flask and heated to reflux (150 °C) in an oil bath for 1 h. The reaction mixture was cooled down to room temperature, then 50 mL of water was added to the reaction mixture and the solvent was removed by a rotary evaporator. The procedure was repeated until all the excess propionic anhydride and propionic acid were removed. Next, 2 M HCl (50 mL) was added to the residue and the mixture was refluxed for 3h. After 3 h, the reaction mixture was decolorized by heating with activated charcoal for 10 min and then filtered to obtain a clear solution. The solvent was removed by a rotary evaporator to obtain a mixture of hydrochloride salts of *cis*-4-hydroxy-D-proline and *trans*-4-hydroxy-L-proline. For each 15.3 g of the mixture, a mixture of 15.3 mL of triethylamine and 413 mL of absolute ethanol (ratio 1: 27) was added. The solution was cooled at 20 °C for 15 min to allow precipitation of the produce. The needle-like crystals were filtered off and washed with ethanol (40–50 mL) to obtain the pure *cis*-4-hydroxy-D-proline (4.5 g, 45% yield). The absence of contamination by the *trans*-isomer (<1%) was confirmed by ^{1}H NMR analysis. ^{1}H NMR (400 MHz, D$_2$O) δ (ppm): 2.05 (1H, m), 2.30 (1H, ddd *J* = 14.6, 10.5, 4.4 Hz), 3.11 (1H, dd *J* = 12.5, 3.9 Hz), 3.23 (1H, m), 4.01 (1H, dd *J* = 10.5, 3.5 Hz), 4.38 (1H, m).

N-Boc-cis-4-hydroxy-D-proline[102]
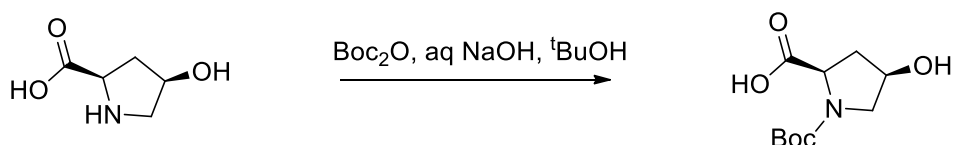
cis-4-Hydroxy-D-proline (4.2 g, 32 mmol) was dissolved in a solution of NaOH (1.92 g, 48 mmol, 1.5 equiv.) in water (25 mL). A solution of di-*tert*-butyl dicarbonate (Boc$_2$O) (8.38 g, 38 mmol, 1.2 equiv.) in *tert*-butanol (25 mL) was slowly added to the reaction mixture with continuous stirring. After stirring at room temperature overnight, the solvent was removed by a rotary evaporator. The residue was re-dissolved in 30 mL of water and the pH was adjusted to 2-3 by adding NaHSO$_4$. The solution was extracted by EtOAc (30 mL × 3) and the organic layer was concentrated by a rotary evaporator. The residue was triturated with 40 mL of EtOAc and left at 20 °C overnight. The white crystalline precipitate was collected by filtration and washed with cold EtOAc to give a white solid in 7.2 g (85% yield). ^{1}H NMR (400 MHz, DMSO-d_6) δ (ppm): 1.31 and 1.37 (9H, 2 × s), 1.81 (1H, m), 2.29 (1H, m), 3.08 (1H, m), 3.45 (1H, m), 4.07 (1H, m), 4.16 (1H, m).

Diphenyldiazomethane[102, 103]
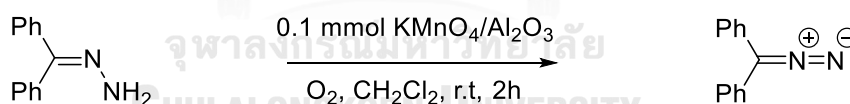
This procedure was developed based on a literature process[104] to replace the older process using HgO as the oxidant[105] which is both dangerous to the operator and generates excessive amounts of mercury waste. An alumina-supported potassium permanganate was prepared by crushing KMnO$_4$ (7.90 g, 50 mmol, 1.9 equiv.) and alumina (25 g) together in a mortar until homogeneous. Benzophenonehydrazone (5.13 g, 26 mmol) was dissolved in 125 mL of dichloromethane in an aluminum foil-wrapped round-bottomed flask (500 mL) to protect from light. Next, the prepared oxidant mixture (alumina-supported potassium permanganate) was slowly added into the reaction mixture and stir vigorously under

normal atmosphere (stopper loosely) at room temperature in the absence of light for 2 h. A purple solution was obtained, which was separated from the alumina support by filtration. The support was extensively washed with hexanes (15 mL × 3) and the combined extracts were evaporated to dryness on a rotary evaporator without heating and light. The pure diphenyldiazomethane was obtained as a dark-purple solid (4.8 g, 95 % yield) and was stored in a freezer at 20 °C in an aluminum foil-wrapped flask.

N-Boc-cis-4-hydroxy-D-prolinediphenylmethyl ester[102]
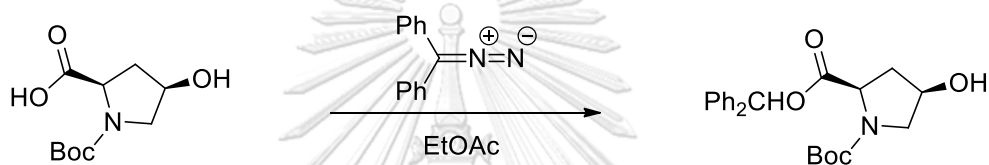
N-Boc-*cis*-4-Hydroxy-D-proline (6.5 g, 25 mmol) was dissolved in 50 mL of EtOAc in a foil-wrapped round bottom flask and cooled in an ice-bath. A solution of diphenyldiazomethane (4.9 g, 25 mmol, 1 equiv.) was added dropwise to give a purple solution that rapidly disappeared. More diphenyldiazomethane was added to get a permanent purple color and the stirring was continued overnight in the absence of light. The solvent was removed by a rotary evaporator to obtain a pink syrup. Next, 150 mL of hexanes was added and sonicated until a white solid formed. The solid was collected by filtration and washed with hexanes to obtain the product (10.5 g, 84% yield). ^{1}H NMR (400 MHz, CDCl$_3$) δ (ppm): 1.17 and 1.39 (9H, 2 × s), 1.64 (1H, m), 2.28 (1H, m), 3.56-3.71 (2H, m), 4.35 (1H, m), 4.35 and 4.47 (2H, 2 × s), 6.81 and 6.89 (1H, 2 × s), 7.20- 7.37 (10H, m).

N-Boc-trans-4-hydroxy-D-prolinediphenylmethyl ester[102]

i) HCOOH, Ph$_3$P, DIAD
ii) aq NH$_3$
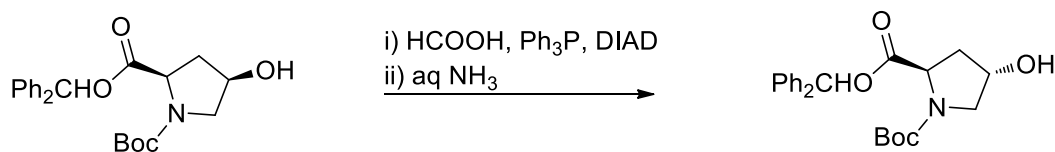
N-Boc-*cis*-4-hydroxy-D-prolinediphenylmethyl ester (3.0 g, 7.55 mmol) and triphenylphosphine (2.38 g, 9.06 mmol, 1.2 equiv.) were dissolved in 20 mL of dry THF. The flask was flushed with N$_2$ (g) and the reaction was cooled down to 0 °C. Next, formic acid (0.31 mL, 8.31 mmol, 1.2 equiv.) was added into the solution with continuous stirring. Diisopropyl azodicarboxylate (DIAD) (1.78 mL, 9.06 mmol, 1.2 equiv.) was then added dropwise over 20 min and the stirring was continued overnight. The solvent was removed, and the crude product was purified by column chromatography [hexanes:EtOAc (3:1)] to afford the formate ester as a colorless oil. The formate ester was hydrolyzed by stirring with a solution of a concentrated aqueous ammonia solution (2 mL) in 20 mL methanol for 3 h. The solvent was removed by a rotary evaporator and the residue was dried under vacuum to give the product as colorless oil (2.7 g, 92 %yield). ^{1}H NMR (400 MHz, CDCl$_3$) δ (ppm): 1.27 and 1.49 (9H, 2 × s), 1.98 (1H, m), 2.03 (1H, m), 2.35 (1H, m), 3.64 (2H, m), 4.35 (1H, m), 4.45 and 4.57 (2H, 2 × s), 6.91 and 6.99 (1H, 2 × s), 7.31-7.39 (10H, m).

N-Boc-cis-4-(N^3-benzoylthymin-1-yl)-D-prolinediphenylmethyl ester[102]

PPh$_3$, DIAD , THF
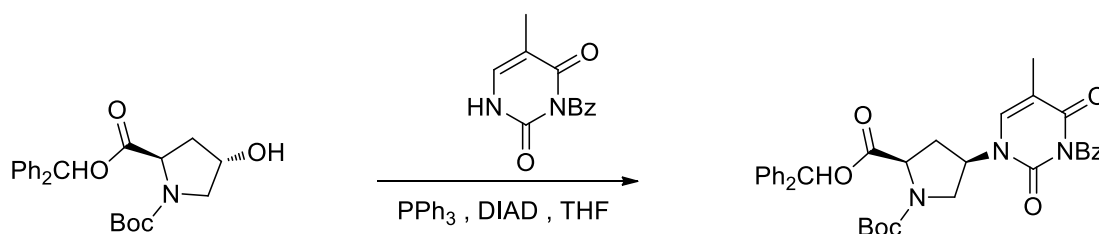
N-Boc-*trans*-4-hydroxy-D-prolinediphenylmethyl ester (2.7 g, 6.8 mmol), N^3-benzoylthymine (1.88 g, 8.2 mmol, 1.2 equiv.), and triphenylphosphine (2.68 g, 10 mmol, 1.5 equiv.) were dissolved together in 20 mL of dry THF. The flask was flushed with N$_2$ (g) and the reaction was cooled down to 0 °C. Diisopropyl azodicarboxylate

(DIAD) (2 mL, 10 mmol, 1.5 equiv.) was then added dropwise over 20 min and the stirring was continued overnight. The solvent was removed by a rotary evaporator to obtain a colorless residue, which was recrystallized by MeOH to obtain a white solid (3.7 g, 90% yield). ^{1}H NMR (400 MHz, CDCl$_3$) δ (ppm): 1.27 and 1.48 (9H, 2 × s), 1.64 (3H, s), 1.80 (1H, m), 2.84 (1H, m), 3.47–3.70 (1H, 2 × m), 3.98 (1H, m), 4.49–4.59 (1H, 2 × m), 5.22 (1H, m), 6.92 (1H, s), 7.16 and 7.33 (1H, 2 × s), 7.33–7.35 (10H, m), 7.47 (2H, t *J* = 7.9 Hz), 7.64 (1H, t *J* = 7.6 Hz), 7.78 (2H, d *J* = 7.4 Hz).

N-Fmoc-cis-4-(N[3]-benzoylthymin-1-yl)-D-proline[102]
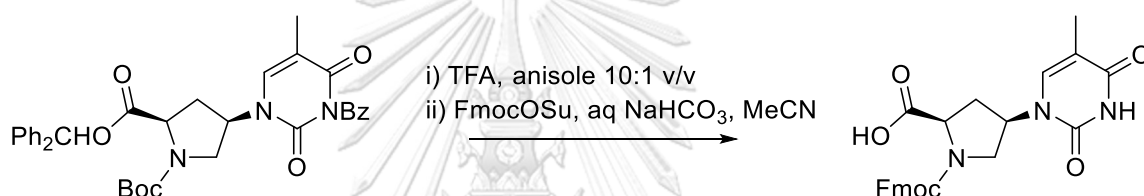
N-Boc-*cis*-4-(N[3]-benzoylthymin-1-yl)-D-prolinediphenylmethyl ester (2.0 g, 3.2 mmol) was dissolved in a mixture of trifluoroacetic acid and anisole (2:1 v/v) (6 mL) in a stoppered round bottomed flask. The reaction was stirred at room temperature overnight. The volatiles were removed by a gentle stream of nitrogen gas. Next, diethyl ether (2 × 10 mL) was added to the residue to obtain a white precipitate. The solid was filtered off and dried under vacuum. This was re-dissolved in 1:1 water/acetonitrile (10 mL). The pH of the solution was adjusted to 8-9 by adding solid NaHCO$_3$ (0.68 g, 8 mmol, 2.5 equiv.). Then, 9-fluorenylmethyl *N*-succinimidyl carbonate (Fmoc-OSu) (1.1 g, 3.2 mmol, 1 equiv.) was added portion wise to the reaction mixture with continuous stirring. After stirring at room temperature overnight, the residue was re-dissolved in 20 mL of water and extracted by diethyl ether (20 mL × 3). The pH was adjusted to 2-3 by adding concentrated HCl to obtain a white solid, which was collected by filtration and dried under vacuum (1.3 g, 87% yield). ^{1}H NMR (400 MHz, DMSO-*d$_6$*) δ (ppm): 1.78 (3H, s), 2.15–2.32 (1H, m), 2.60–2.71 (1H, 2 × m), 3.50 (1H, m), 3.87 (1H, m), 4.21–4.42 (4H, m), 4.97 (1H, m), 7.29 (2H, m), 7.32 (2H, m),

7.41-7.44 (2H, m), 7.58 (1H, s), 7.67-7.69 (2H, m), 7.84-7.96 (2H, m), 11.30 (1H, br s), 12.89 (1H, br s).

N-Fmoc-cis-4-(N³-benzoylthymin-1-yl)-D-proline pentafluorophenyl ester[102]
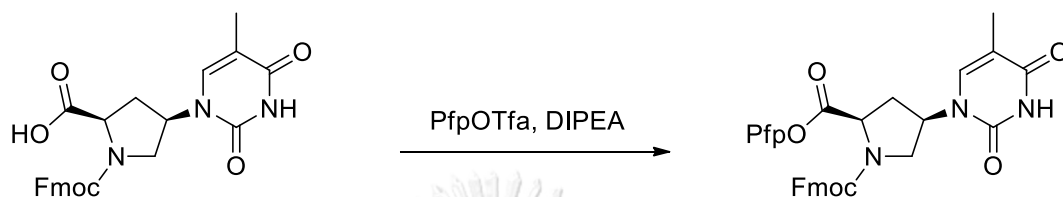
N-Fmoc-*cis*-4-(*N³*-benzoylthymin-1-yl)-D-proline (0.25 g, 0.54 mmol) was dissolved in 10 mL of dichloromethane. Next, diisopropylethylamine (DIPEA) (0.18 mL, 1.08 mmol, 2 equiv.) was added to the solution mixture. Pentafluorophenyl trifluoroacetate (PfpOTfa) (0.19 mL, 1.08 mmol, 2 equiv.) was then added in one portion with continuous stirring. The reaction was next stirred for 30 min at room temperature. The reaction was monitored by TLC (50% EtOAc: Hexane). If the starting material was still present at the base line, more PfpOTfa (0.095 mL, 1 equiv.) and DIPEA (0.09 mL, 1 equiv.) were added. The crude mixture was extracted with 10% HCl (10 mL × 2), water (10 mL × 1), and saturated aqueous NaHCO$_3$ (10 mL × 2). The dichloromethane layer was dried over anhydrous MgSO$_4$ and concentrated by a rotary evaporator. The residue was re-dissolved in 3 mL of dichloromethane, and this solution was added dropwise into a beaker of stirring hexanes (100 mL). The white solid formed was filtered and washed with hexanes and dried under vacuum (0.4 g, 89% yield). ^{1}H NMR (500 MHz, CDCl$_3$) δ (ppm): 1.92 (3H, s), 2.30 and 2.41 (1H, 2 × m), 2.96 (1H, m), 3.56 and 3.74 (1H, 2 × m), 3.87 and 4.03 (1H, 2 × m), 4.24 (1H, m), 4.53 (2H, m), 4.65 and 4.77 (1H, 2 × m), 5.17 and 5.27 (1H, 2 × m), 7.09 (1H, s), 7.30 (2H, m), 7.40 (2H, m), 7.55 (2H, d *J* = 7.5 Hz), 7.76 (2H, d *J* = 7.5 Hz), 8.24 (1H, s).

2.1.2.2 Solid phase peptide synthesis for acpcPNA
Procedure

The following stock solutions were prepared to be used in the solid phase synthesis of PNA:

Stock #1 (20% piperidine, 2% DBU in DMF): 20 µL of DBU and 200 µL of piperidine in 780 µL DMF

Stock #2 (7% DIPEA in DMF): 70 µL of DIPEA in 930 µL DMF

Stock #3 (5.5% or 0.4 M HOAt in DMF): 5.5 mg of HOAt in 100 µL DMF

The PNA oligomers, to be later modified by the furan part, were prepared by solid phase peptide synthesis (SPPS) on a Tentagel S-RAM resin at 1.5 µmol scale. The synthesis was performed by stepwise couplings of the PNA monomers (A, T, C, G) and ACPC/APC spacer as reported previously.[106] The Fmoc-protection of the amino group on the Rink amide resin was first removed by treatment with stock #1 (30 µL, 5 min) to generate the free *N*-terminal amino group. The resin was washed extensively with DMF. In the coupling step, the coupling of the desired Pfp-activated monomers or spacer (4 equiv. + stock #2 15 µL and stock #3 15 µL) (40 min for Fmoc-T-OPfp, 30 min for other monomers). In the case of Lys or G monomer, the free acids were used instead of the unstable Pfp esters. They were first activated by HATU prior to the coupling (4 equiv. Fmoc-G-OH or Fmoc-Lys(Mtt)-OH, 4 equiv. HATU (6 µmol, 2.2 mg) in 15 µL of stock #2). The coupling reaction was allowed to proceed for 30 minutes. Next, the resin was washed exhaustively with DMF. Finally, the resin was treated with stock #2 (30 µL) and 5 µL of Ac$_2$O for 5 min to cap the unreacted amino group from further reacting in the next steps. This deprotection-coupling-capping process was performed alternately for the PNA monomers and spacers and was repeated until the desired PNA sequence was obtained.

For the terminal furan modification, the fully protected PNA(T) precursor, Fmoc-GTAGATCACT-K(Mtt), was first synthesized on the solid support as described above. The PNA on the resin was next split into 0.5 µmol portions for further reactions. The *N*-terminal Fmoc group was removed and the PNA was subsequently

modified with 4 equiv. of the Fmoc-AEEA-OH [2-(2-(Fmoc-amino)ethoxy)ethoxy]acetic acid to expand the length of the furan part. The Fmoc-AEEA-OH monomer (6 µmol, 2.3 mg) was activated with 4 equiv. HATU (6 µmol, 2.2 mg) in 15 µL of stock #2, and were subsequently coupled with the last ACPC spacer. Next, the PNA(T) precursor was repeatedly treated with 5% TFA in DCM (5 × 500 µL × 1 min) until the yellow color in the cleavage solution turned to colorless to remove the Mtt protecting group from the lysine side chain. Then, the free amino group of the C-terminal lysine was reductively dimethylated by treatment with 30 equiv. of formaldehyde (38% w/w aqueous solution) in the presence of 60 equiv. of $NaBH_3CN$ and 3 µL of AcOH in 300 µL of methanol for 5 h.[107] This step was necessary to protect the lysine amino group from reacting with the furan-carboxylic acid part during the post-synthetic modification while still maintaining the positive charge necessary for the water solubility. The reaction was monitored by MALDI-TOF MS, which showed that the reaction was completed within 5 h as shown by the disappearance of the mass signal of the PNA and concomitant formation of a new peak with the mass increased by 29 units, which corresponded to the addition of two methyl groups. Next, the *N*-terminal Fmoc group was removed by treatment with stock #1 as usual followed by removal of the nucleobase protecting groups by heating the PNA resin with 1:1 ammonia-dioxane at 65 ºC (overnight). After the side chain deprotection step, the PNA was cleaved from the solid support by treatment with trifluoroacetic acid (TFA) for (3 x 500 µL x 50 min). The TFA was removed from the cleavage solution by purging with N_2 gas and the residue was centrifugally washed with diethyl ether to give the crude PNA for further post-synthetic modification with the furan moiety.

For the internal furan modification, the fully-protected PNA(I) precursor - Fmoc- GTAGA(apc^{Tfa})TCACT-K(Mtt) - was first synthesized on the solid support following the similar protocol as the PNA(T) precursor. The only difference was one of the ACPC spacer in the middle of the strand was replaced by 3-aminopyrrolidine-4-carboxylic acid (APC) spacer to provide a handle for the internal modification by furan. For the coupling step involving APC spacer, the Fmoc-APC-OPfp spacer (4 equiv.) was activated by stock #2 (15 µL) and stock #3 (15 µL) as usual, but the coupling time was increased to 1h. The deprotection and capping steps were

performed as usual. After finishing the synthesis, the *N*-terminal Fmoc group of the PNA was first removed by the stock solution #1 followed by acetylation to protect the *N*-terminal amino group for further reactions. After that, the Mtt protecting group of the lysine side-chain was removed prior to the dimethylation. Next, the nucleobase side chain deprotection and TFA cleavage were performed as described for the PNA(T). Note that the Tfa protecting group of the APC spacer was simultaneously removed under the same condition as the nucleobase side chain deprotection. MALDI-TOF MS was used to monitor the progress of the reactions.

2.1.2.3 Pfp-activation of 3-(2-furyl) propionic acid

For the Pfp-activation, the 3-(2-furyl) propionic acid (2 mg, 30 equiv. at 0.5 µmol scale reaction) was dissolved in 1 mL of DCM. Next, 39 µL (450 equiv.) of DIPEA was added and stirred for 1 min. This was followed by adding 38 µL (450 equiv.) of pentafluorophenyl trifluoroacetate and the stirring was continued for another 3 min. The color of the solution mixture changed from yellow to brown. Next, the crude reaction was diluted by more DCM and was extracted by 10% HCl, H_2O and sat. $NaHCO_3$. The DCM extract was dried over $MgSO_4$ and the solvent was removed by a gentle stream of N_2 gas to obtain a dark brown residue, which was directly used for the coupling with the PNA in the solution phase without characterization.

2.1.2.4. Post synthetic modification of acpcPNA by furan

The post-synthetic modification was performed by dissolving the crude PNA after the deprotection and cleavage from the solid support (0.5 µmol) in 30 µL of dry DMF. After which, 7 µL of DIPEA and Pfp-activated 3-(2-furyl) propionic acid (10 equiv.) were added. The reaction was completed in 5 min as confirmed by MALDI-TOF MS. Next, 800 µLs of diethyl ether were added to precipitate the modified PNA. After extensive ether washing and air drying, the crude PNA was purified by reverse phase HPLC on a C_{18} HPLC column (4.6 × 150 mm). The elution was performed using a water (A) and methanol (B) containing 0.1% trifluoroacetic acid gradient system. The system was first equilibrated at 90:10 A:B for 5 min. Then a linear gradient was

applied from 90:10 to 10:90 (A: B) over 60 min at a flow rate of 0.5 mL/min. The peaks were monitored at 260 nm, and the fractions showing the desired mass according to MALDI-TOF analyses were collected and freeze dried. Characterization was achieved by MALDI-TOF MS employing α-cyano-4-hydroxycinnamic acid (CCA) was used as the matrix. Characterization data: PNA(T) *m/z* observed 3812.3, calculated for M·H$^+$: 3811.9 Da (5.9% isolated yield); PNA(I) *m/z* observed: 3709.1 Da, calculated for M·H$^+$: 3710.8 Da (9.6% isolated yield).

2.1.3 Melting temperature experiments

The samples containing 1 µM of both PNA and DNA in 10 mM phosphate buffer (pH 7) containing 100 mM NaCl with a total volume of 1000 µL was placed in a quartz cuvette with a Teflon stopper. After that, 1 equiv. of NBS (1 µL of 1 mM NBS stock solution) was added and left for 15 min at 37 °C. Next, the second equiv. of NBS was added and the reaction was subsequently incubated for another 45 min at 37 °C (total equiv. of NBS = 2, total time = 1 h). The UV absorbance was recorded at 260 nm from 20 to 90 °C (or 95 °C in some cases) at the heating rate of 1.0 °C/min. The melting temperature graphs were performed by KaliedaGraph 4.0 (Synergy Software) and Microsoft Excel (Microsoft Corporation). The melting temperature (T_m) was calculated from the maximum value of the first derivative plot.

2.1.4 Cross-linking protocol

The mixture of furan modified PNA and DNA (20 µM of PNA and 20 µM of DNA in a total volume of 10 µL for denaturing PAGE, 5 µM of PNA and 7.5 µM of DNA in a total volume of 100 µL for HPLC, 1 µM of PNA and 1 µM of DNA in a total volume of 1 mL for melting temperature) in 10 mM phosphate buffer pH 7 containing 100 mM NaCl was denatured at 95 °C for 5 mins and then kept in an ice bath for 1 h before adding freshly prepared NBS solution. The first equiv. of NBS (1 µL of 0.2 mM NBS stock solution for denaturing PAGE, 1 µL of 0.5 mM NBS stock solution for HPLC, 1 µL of 1 mM NBS stock for melting temperature) was added and left for 15 min at 37 °C. Next, the second equiv. of NBS was added and the reaction was subsequently

incubated for another 5 or 45 min at 37 °C (total equiv. of NBS = 2, total time = 20 min to 1 h).

2.1.5 Denaturing PAGE

A mixture consisting of 20 µM of the furan modified-PNA and 20 µM of DNA in 10 mM phosphate buffer (pH 7) containing 100 mM NaCl in a total volume of 10 µL was denatured at 95 °C for 5 min and then annealed in ice-bath temperature over 1 h. The cross-linked samples were analyzed using a 17% polyacrylamide gel (29:1 acrylamide: *N,N'*-methylenebis(acrylamide) 40%) in TBE buffer containing 7 M urea. The gel was prepared with the following composition: 4.6 g of urea, 1 mL of 10× TBE buffer, 4.3 mL of 40% w/v 29:1 acrylamide: *N,N'*-methylenebis(acrylamide), 33 µL of 10% ammonium persulfate, and 10 µL of *N,N,N',N'*-tetramethylethylenediamine (TEMED). The temperature of the gel experiments was externally controlled (ice bath). The gels were pre-run at 300 V for 15 min before being run at 225 V for 1 h using a NANOPAC-300 power supply (Cleaver scientific, United Kingdom). Formamide solution (5 µL) was added into the samples (10 µL). Then 10 µL of the mixture solution was loaded into the gel. The gels were photographed under UV light (256 nm) by UV shadowing over a 20 × 20 cm TLC plates that was coated with silica gel GF254.

2.1.6 RP-HPLC of ICL products

RP-HPLC analysis of the ICL was performed on a Water Delta 600 system equipped with an ACE C_{18} HPLC column (4.6 × 50 mm) and a Water 996 photodiode array detector (monitored by UV absorbance at 260 nm). The sample consisted of 7.5 µM of the DNA and 5 µM of the PNA strand in 10 mM phosphate buffer (pH 7) containing 100 mM NaCl in a total volume of 100 µL. The cross-linking was induced by 2 equiv. NBS for 30 to 1 h before injection to the HPLC column. The elution was performed using a 0.1 M TEAA buffer (pH 7) (A) and MeCN (B) at a flow rate of 0.5 mL/min. The column was flushed with 95% A for 5 min, then to 80% B over 30 min at a flow rate of 0.5 mL/min.

2.1.7 MALDI-TOF MS

MALDI-TOF mass spectra were recorded on a Microflex MALDI-TOF mass spectrometer (Bruker Daltonics). For analysis of the PNA samples, α-cyano-4-hydroxycinnamic acid (CCA) was used as a matrix. In the case of cross-linked product, the sample was desalted by treatment with Dowex® 50W X8 beads prior detection. 3-Hydroxypiconilic acid: ammonium citrate (9: 1) was used as a matrix.

2.2 Section II (Experiments performed at Ghent University)

2.2.1 General

All reagent grade chemicals and solvents were purchased from Fluka, Sigma-Aldrich, Merck, TCI Europe, Fluorochem and used as received without further purification. Anhydrous *N,N*-dimethylformamide (<50ppm) used for the solid phase peptide synthesis dried over activated 4 Å molecular sieves. DNA oligonucleotides were purchased from IDT (Leuven, Belgium).

2.2.2 AcpcPNA synthesis

DNA: 5'-AGATCATGCCC-**X**-3'
PNA: H₂N-rrr-TCTAGTACGGG-**M**

X:
A1 = A   A5 = AAAAA
C1 = C   C5 = CCCCC
G1 = G   G5 = GGGGG
T1 = T   T5 = TTTTT

R: regPNA
P: acpcPNA
The PNA sequence was written from C to N direction.
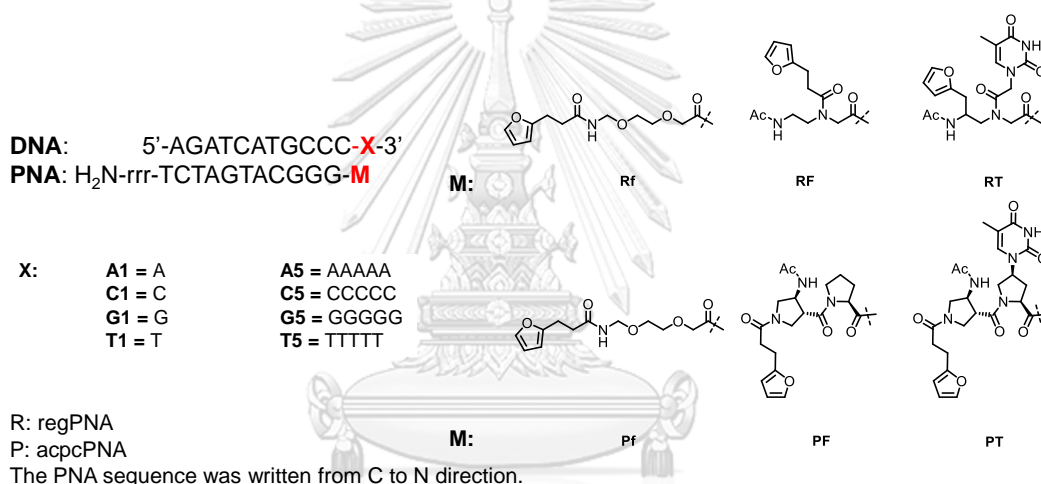
Figure 2.2 Furan building blocks and PNA/ DNA sequences in terminal modified PNA

DNA: 5'-GCAGATC-**X Y**-GCCCGGC-3'
PNA: H₂N-rrr-TCTAG-**M P**-CGGG-Ac

X: the facing nucleobase of **M** building block
Y: the previous DNA nucleobase
P: the nucleobase before **M** building block
M: furan modified acpcPNA building block
The PNA sequence was written from C to N direction.

Fmoc-D-proline-OPfp   **Fmoc-Arg(Pbf)-OH**
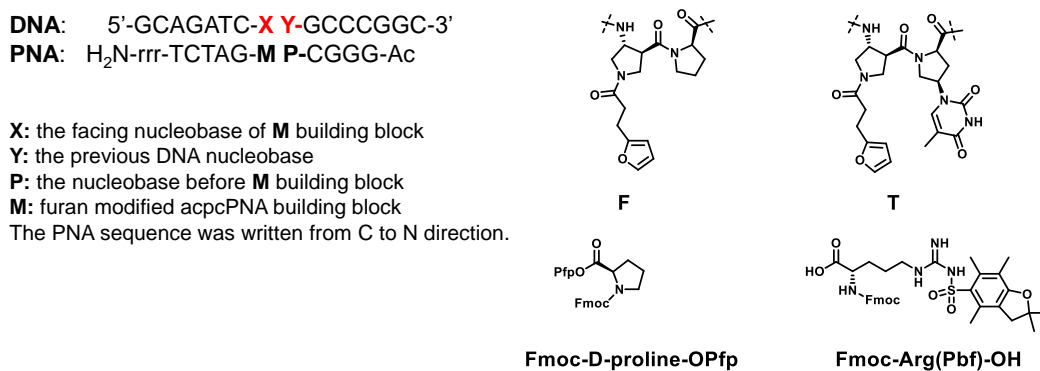
Figure 2.3 Furan building blocks and PNA/ DNA sequences in internal modified acpcPNA and structure of Fmoc-ᴅ-proline-OPfp and Fmoc-Arg(Pbf)-OH

The synthesis of acpcPNA was performed at Chulalongkorn University using the same monomers as described in section I (**2.1.2.1**), but with additional monomers including Fmoc-D-proline-OPfp (synthesized by Dr. Boonsong Ditmanklo) and Fmoc-Arg(Pbf)-OH (commercially available from TCI). The sequences of terminal furan modified acpcPNA and internal furan modified acpcPNA are shown in **Figure 2.2** and **Figure 2.3**, respectively. The PNAs were synthesized by SPPS on a Tentagel S-RAM resin at 1.5 µmol scale. The SPPS cycles (deprotection, coupling, and capping steps) were performed as described in section I (**2.1.2.2**). To avoid the lengthy and time-consuming lysine dimethylation step, arginine was employed as the solubility enhancer instead. Three arginine residues were first loaded onto the resin by using the coupling solution consisting of Fmoc-Arg(Pbf)-OH (4 equiv., 3.8 mg), HATU (4 equiv., 2.2 mg), in 15 µL of stock #2 for 40 min. The PNA monomers and spacers were then incorporated as described previously in section I (**2.1.2.2**).

Table 2.1 Activation conditions for the monomers for the acpcPNA synthesis in section II

monomer	coupling condition
Fmoc-AEEA-OH	4 equiv. Fmoc-AEEA-OH (2.3 mg), 4 equiv. HATU (2.2 mg) in 15 µL of stock #2 for 40 min
Fmoc-D-proline-OPfp	4 equiv. Fmoc-D-proline-OPfp in a solution of stock #2 (15 µL) and stock #3 (15 µL)
Fmoc-T-OPfp	4 equiv. Fmoc-T-OPfp in a solution of stock #2 (15 µL) and stock #3 (15 µL)

For terminally-modified PNAs Pf, PF, and PT, the resin-supported PNAs were modified prior to the coupling with the furan monomers as follows: After obtaining the PNA's precursor (Fmoc-GGGCATGATCT-rrr), the *N*-terminal Fmoc group was removed and the PNA was subsequently modified whilst still on the solid support with the following monomers: Fmoc-AEEA-OH for Pf, Fmoc-D-proline-OPfp for PF, and Fmoc-T-OPfp for PT sequences. The coupling conditions were presented in **Table 2.1**. Next, the *N*-terminal Fmoc group was removed again by treatment with stock #1

followed by the side chain deprotection step as described in section I (**2.1.2.2**). The PNA was subsequently cleaved from the solid support by treatment with TFA (2 x 500 μL x 2 h). The incorporation of the furan moiety by the post-synthetic method was performed at 0.5 nmol scale as described in section I (**2.1.2.4**). The reactions were monitored by MALDI-TOF MS which indicated that the reaction proceeded to completion within 2 h. Next, the crude PNA were purified by RP-HPLC as described in section **2.1.2.4**. The characterization data of the three terminal furan-modified acpcPNAs (Pf, PF, and PT) are reported in **Table 2.3**.

The internal furan modified acpcPNA sequences were synthesized similarly. For the precursor of the F-series (Ac-GGGCXFGATCT-rrr), an abasic site (in the form of D-proline) was inserted at the position marked with F using Fmoc-D-proline-OH instead of the regular PNA monomer. Fmoc-APC^{Tfa}-OPfp was subsequently loaded as the spacer. The acpcPNA monomer with various nucleobase (A, C, G, T) was next loaded at the position X, which led to four acpcPNA sequences, namely P-AF, P-CF, P-GF, and P-TF. The synthesis of the precursors of the T-series (Ac-GGGCXTGATCT-rrr), was the same as with F-monomer, but the acpcPNA T-monomer (Fmoc-T-OPfp) was incorporated in place of the Fmoc-D-proline-OH. After finishing the synthesis, the *N*-terminal Fmoc group of the PNA was removed by treatment with stock solution #1 followed by acetylation. After that, the nucleobase/APC deprotection and the TFA cleavage were performed under the same conditions as described in section I (**2.1.2.2**). The post-synthetic modification with the furan moiety was performed as described in section I (**2.1.2.4**). The reactions were monitored by MALDI-TOF MS which indicated that the reaction proceeded to completion within 2 h. Next, the crude PNA were purified by RP-HPLC as described in section **2.1.2.4**. The characterization data of the internal furan-modified acpcPNA are reported in **Table 2.3**.

2.2.3 Solid phase peptide synthesis for aegNA

The following stock solutions were prepared to be used in the solid phase synthesis of aegPNA:

Deprotection: 20% piperidine in DMF [Piperidine/DMF (1:4 v/v)]

Coupling: 5 equiv. monomer in 300 μL DMF, 5 equiv. HBTU in 300 μL DMF, 10 equiv. DIPEA

Capping: Ac$_2$O/DIPEA/DMF (5:6:89v/v)

An aegPNA synthesis was performed by using a standard Fmoc-based manual synthesis protocol on a Rink amide resin at a 20 μmol scale (200 mg) employing the four Fmoc/Bhoc protected aegPNA monomers. Three arginine residues were inserted into the C-terminus to improve the solubility of the PNA as explained in section **2.2.2** The SPPS steps of aegPNA consisted of Fmoc-deprotection, coupling, and capping similar to the acpcPNA. The synthesis was performed in a 10 mL polypropylene syringes vessel with filter equipped with Vacuum Manifold. Fmoc-deprotection was first carried out by a treatment of the DCM-swelled resin (20 μmol) with 20% piperidine in DMF (2 × 2 mL × 8 min). The reaction tube was shaken on a flask shaker (8 min). The resin was washed extensively and separately with DMF, DCM and extra dry DMF. After this step, the resin was tested positive by the Kaiser test (green). The arginine or aegPNA monomers were first activated by the following procedure: the monomer (5 equiv.) was dissolved in 300 μL of extra dry DMF; HBTU (5 equiv.) was dissolved in extra dry DMF in the same volume (300 μL); the two solutions were mixed and DIPEA (10 equiv.) was added for 2 min before coupling with the resin (shaking for 30 min). After the coupling step, the resin was washed with DMF (5 times) and the Kaiser test was performed again which should give a negative test (colorless). If a positive Kaiser test was obtained, the second coupling would be performed. In the capping step, the resin was shaken with the capping solution (1 × 2 mL × 2 min). Next, the resin was washed with 5% DIPEA in DMF (2 × 2 mL × 2 min) and then with DMF (5 times). All three steps were repeated until the desired PNA sequences were obtained.

Next, the aegPNA were modified with the furan moiety. In order to prevent the degradation of the furan moiety, the maleimide protected furan monomers (FurDA-OH and Fmoc-F*-OH)[100] (shown in **Figure 2.4**) synthesized by Dr. Alex Manicardi (Ghent University) were used for the solid phase synthesis of the furan-modified aegPNAs. For the solid phase synthesis of the T-sequences, the furan monomer (Fmoc-Fur-T-OH), synthesized by Dr. Alex Manicardi (Ghent University), was used without protection of furan moiety because of the problematic aspect during maleimide-modified furan synthesis.
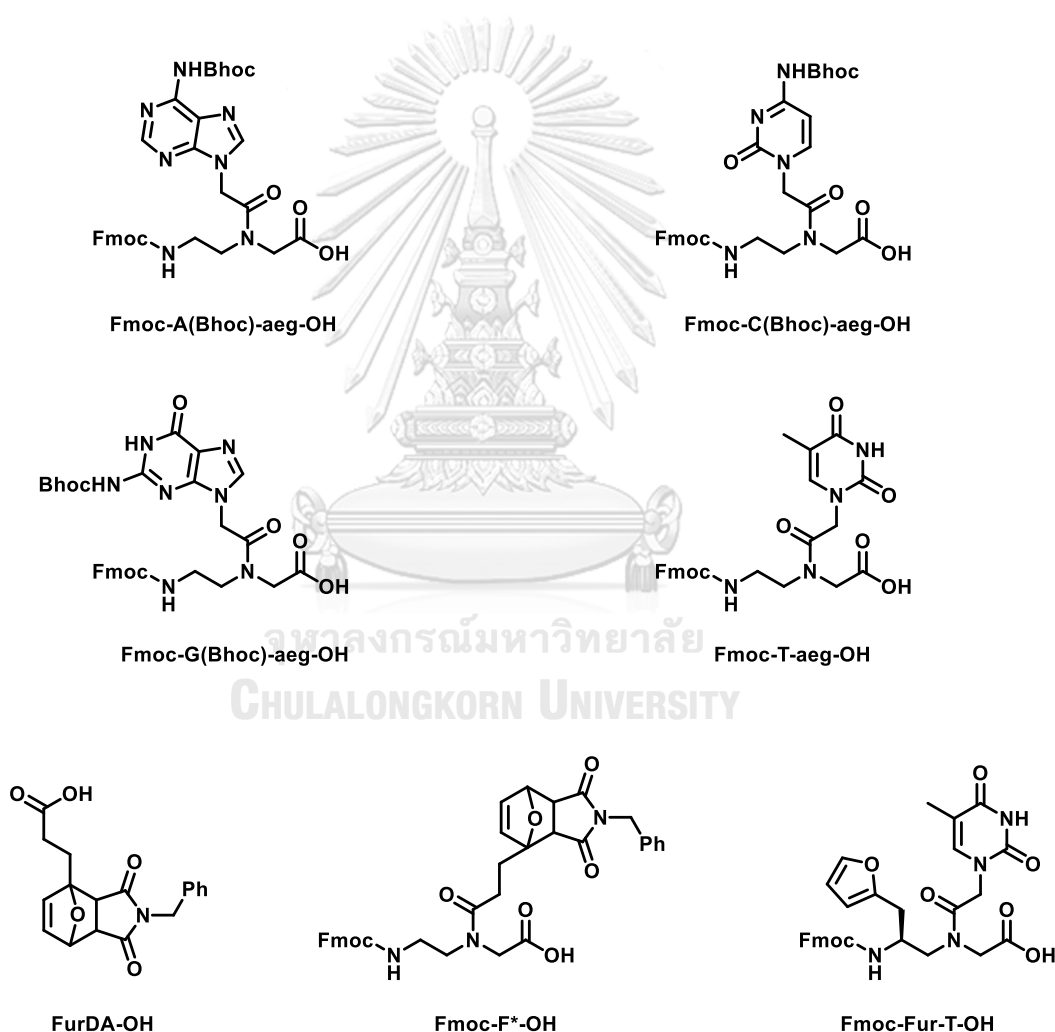
Figure 2.4 Structure of aegPNA monomers and the three furan monomers used in the synthesis of furan-modified aegPNA sequences

Table 2.2 Activation conditions for the monomers for the aegPNA synthesis in section II

monomer	coupling condition
Fmoc-AEEA-OH	5 equiv. Fmoc-AEEA-OH, 5 equiv. HBTU, 10 equiv. DIPEA
FurDA-OH	5 equiv. FurDA-OH, 5 equiv. HBTU, 10 equiv. DIPEA
T-monomer	5 equiv. Fmoc-Fur-T-OH, 5 equiv. HBTU, 10 equiv. DIPEA
F-monomer	5 equiv. Fmoc-F*-OH, 5 equiv. HBTU, 10 equiv. DIPEA

For the terminal furan-modified aegPNA, the aegPNA precursor (Fmoc-GGGCATGATCT-rrr) was synthesized according to the general protocol described above. The precursor was then coupled whilst still on the solid support with Fmoc-AEEA-OH followed by FurDA-OH for Rf, with F-monomer for RF, and with T-monomer for RT. The coupling conditions of each monomer are shown in **Table 2.2**. After obtaining the desired sequences, the PNAs were cleaved from the solid support by a treatment with a solution of TFA/m-cresol 9:1 for Rf and RF PNA sequences. The RT PNA sequence was cleaved in a TFA/m-cresol/thioanisole 8:1:1 solution to prevent the alkylation of the furan by the diphenylmethyl cation obtained as the by-product of the Bhoc- group cleavage. The crude PNAs were precipitated by adding diethyl ether (5 mL). After extensive ether washing and air drying, the crude PNA was split into four portions (5 μmol each). The retro-Diels-Alder reaction was performed on the crude Rf and RF PNAs to remove the maleimide protecting group by adding 1.5 mL of milli-Q water and the pH of the solution mixture was adjusted to 11.5 by adding 1 M NaOH. After that, the solution mixture was heated at 90 °C for 2 h. The reaction was monitored by MALDI-TOF MS. Once the reaction was completed, the crude deprotected PNA solutions were dried by lyophilization and purified by RP-HPLC. The purification of the crude aegPNA was performed on an Agilent 1100 Series equipped with a Luna C18(2) (5 μm, 100 Å, 250x4.6 mm). The elution was performed using 0.1% TFA in H_2O (A) and MeCN (B). The system was first equilibrated at 100% A for 5 min. Then a linear gradient was applied to 50% B over 30 min at a flow rate of

4 mL/min. The characterization data of the terminal furan-modified aegPNA are shown in **Table 2.3**.

Table 2.3 Characterization data of acpcPNA (P) and aegPNA (R) used in section II

name	sequence	m/z calcd.	m/z found	% yield
Rf	f-O-GGGCATGATCT-rrr-NH$_2$	3768.8	3765.3	11.3
RF	Ac-F*-GGGCATGATCT-rrr-NH$_2$	3766.7	3766.2	11.8
RT	Ac-T*-GGGCATGATCT-rrr-NH$_2$	3890.8	3886.9	5.7
Pf	f-O-GGGCATGATCT-rrr-NH$_2$	4497.8	4498.2	3.9
PF	Ac-F*-GGGCATGATCT-rrr-NH$_2$	4603.9	4604.0	6.7
PT	Ac-T*-GGGCATGATCT-rrr-NH$_2$	4728.0	4729.3	6.4
P-AF	Ac-GGGCA-FGATCT-rrr-NH$_2$	4271.5	4269.5	3.8
P-CF	Ac-GGGCC-FGATCT-rrr-NH$_2$	4262.5	4263.6	4.1
P-GF	Ac-GGGCG-FGATCT-rrr-NH$_2$	4247.5	4247.8	3.9
P-TF	Ac-GGGCT-FGATCT-rrr-NH$_2$	4287.5	4289.0	4.0
P-AT	Ac-GGGCA-TGATCT-rrr-NH$_2$	4395.6	4395.5	4.1
P-CT	Ac-GGGCC-TGATCT-rrr-NH$_2$	4386.6	4386.6	4.3
P-GT	Ac-GGGCG-TGATCT-rrr-NH$_2$	4371.6	4372.1	5.9
P-TT	Ac-GGGCT-TGATCT-rrr-NH$_2$	4411.6	4412.4	6.5

The internally furan-modified aegPNA were synthesized previously by Miss. Francesca Pennati (Università degli Studi di Milano) in 2019-2020.[50]

2.2.4 Cross-linking protocol

The mixture of furan modified PNA and DNA (5 µM of PNA and 5 µM of DNA for denaturing PAGE and HPLC) in 10 mM phosphate buffer (pH 7) containing 100 mM NaCl was incubated in an Eppendorf thermomixer comfort 5355 at 25 °C, 700 rpm for 30 min before adding freshly prepared NBS. The first equiv. of NBS (2 µL of 250 µM NBS stock solution) was added and shaken for 15 min at 25 °C in the thermomixer. The reaction time was recorded immediately after addition of the first equiv. NBS. The second, the third, and the fourth equiv. of NBS was added after 15, 30, and 45 min, respectively. After the final portion of NBS was added, the reaction was left for another 75 min (total equiv. of NBS = 4, total time = 2 h). The temperature of the cross-linking experiments was kept constant at 25 °C.

2.2.5 Denaturing PAGE

A mixture consisting of 5 µM of the furan modified-PNA and 5 µM of DNA in 10 mM phosphate buffer (pH 7) containing 100 mM NaCl in a total volume of 100 µL was prepared. The cross-linking was induced by 4 equiv. NBS for 2h. The cross-linked samples were analyzed using a 20% polyacrylamide gel (19:1 acrylamide: *N,N'*-methylenebis(acrylamide) 40%) in TBE buffer containing 7 M urea. The gel was prepared with the following composition: 8.4 g of urea, 2 mL of 10× TBE buffer, 10 mL of 40% w/v 19:1 acrylamide: *N,N'*-methylenebis(acrylamide), 236 µL of 10% ammonium persulfate, and 20 µL of *N,N,N',N'*-tetramethylethylenediamine (TEMED). The temperature of the gel experiments was controlled by a Julabo F12 refrigerating circulator (25 °C). The gels were pre-run at 230 V for 15 min before being run at 230 V for 1 h and 40 min using a CBS Scientific MGV-202 gel electrophoresis apparatus, connected to a Consort EV202 power supply. The gels were stained using SYBR Gold® and the gel images were recorded by an Autochemi gel imaging system (UVP). Formamide (16 µL) was added into each of the samples (4 µL) to obtain 20 µL of the mixture solution. Then, 10 µL of the mixture solution was loaded into the gel.

2.2.6 RP-HPLC of ICL products

RP-HPLC analysis of the ICL was performed on an Agilent 1200b System (Agilent Technologies, Machelen, Belgium) equipped with a Waters Xbridge 130 Å Oligonucleotide C18 column at a column temperature of 60 °C. The sample was prepared by mixing 5 µM of the DNA and 5 µM of the PNA strand in in 10 mM phosphate buffer (pH 7) containing 100 mM NaCl in a total volume of 150 µL. The cross-linking was induced by 4 equiv. NBS for 2 h before injection to the HPLC column. The elution was performed using a 0.1 M TEAA buffer (pH 7) with 5% MeCN (A) and MeCN (B) gradient system at a flow rate of 0.8 mL/min. The column was flushed with 100% A for 1 min, then to 30% B over 14 min and then increased to 100% B in 15 min. The column was washed with 100% B for 3 min before returning to 100% A and equilibrated for 5 minutes before starting the next analysis cycle.

2.2.7 MALDI-TOF MS

MALDI-TOF mass spectra were recorded on an Applied Biosystems - 4800 Plus MALDI TOF/TOF™ Analyzer. For analysis of the PNA samples and cross-linked product, 2,5-DHB (2,5-dihydroxy benzoic acid) was used as the matrix. Each sample was prepared by spotting 0.3 µL of analyte, allowing it to dry, and then adding 0.3 µL of matrix solution (100 mg/mL 2,5-DHB in 1:2 H_2O:MeCN containing 0.1% TFA) on top of the dried sample. Laser intensity was maintained in the range of 5500-7900K, depending on the sample.

CHAPTER III

RESULTS AND DISCUSSION

SECTION I

3.1 FURAN-MODIFIED acpcPNA

This dissertation focuses on the development of a cross-linkable acpcPNA probe based on furan as a reactive functional moiety. The approach used in this work involves the incorporation of the furan molecule into the PNA strand at both internal [PNA(I)] and *N*-terminal [PNA(T)] positions of the PNA strand to investigate and compare the cross-linking efficiency. In the cross-linking studies, *N*-bromosuccinimide (NBS) was used to activate the furan moiety to generate the reactive aldehyde species that can react with the nucleobases in the DNA strand to form cross-linked products. In the cross-linking experiments, the PNA and DNA strands were first hybridized to form a duplex before adding the NBS to oxidize the furan to generate the reactive aldehyde species that subsequently undergoes the cross-linking reactions with the DNA bases as shown in **Figure 3.1**. Previous reports suggest that the cross-linking reaction only takes place with the nucleobases bearing an exocyclic amino group (A, C, and possibly G) that are not involved in hydrogen bonding.[63, 100]
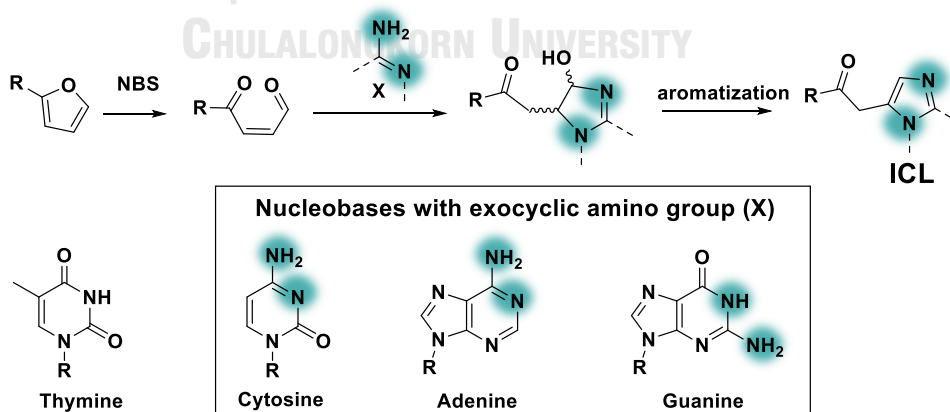
Figure 3.1 The cross-linking mechanism of oxidized furan with exocyclic amino bases upon NBS activation

In this study, the formation of the cross-linking products was investigated by four different techniques. These include thermal denaturation to determine the melting temperature (T_m), denaturing PAGE, MALDI-TOF MS and RP-HPLC. For the thermal denaturation technique, the temperatures at which 50% of the base pairs in the PNA:DNA duplex which are broken were determined by UV-vis spectrophotometry at 260 nm, both before and after NBS activation. The higher melting temperature indicates the stronger duplex stability. It is expected that if the cross-linking occurs, the melting temperature of the cross-linked PNA-DNA duplex should be higher than the non-cross-linked duplex due to the intramolecular nature of the base pairing.[59] The second technique is the denaturing PAGE, which separates DNA fragments according to their sizes. DNA fragments are negatively charged, so they move towards the positive electrode at the bottom of the gel. Smaller and/or more negatively charged fragments move through the gel faster than the large/less negatively charged ones. In the denaturing PAGE, the gel contains urea as a denaturing agent that can destroy the PNA-DNA duplexes as well as any DNA secondary structures which might complicate the analysis of the results. Thus, only single stranded DNAs should be observed. However, in the cross-linked product, the PNA and the DNA strands are joined together by a covalent bond. Thus, unlike the non-cross-linked PNA-DNA duplex, they would move together even in the denaturing PAGE at a slower rate than the single stranded DNA due to the presence of the PNA component which is uncharged. The next technique is MALDI-TOF MS, which involves using a laser light to ionize the DNA embedded in a light-absorbing matrix. The ionized DNA is detected according to their size and charge. However, previous works[100] reported that the cross-linking products are difficult to detect by mass spectrometry, so that the matrix and the sample preparation must be optimized. The last technique is RP-HPLC that separates the analytes according to their retention on a reverse-phase column. Because the size and polarity of the cross-linking product is different from the single stranded DNA or PNA alone, these species show different retention time in the HPLC chromatogram.

3.1.1 Furan-modified acpcPNA synthesis

The PNA(I) and PNA(T) were synthesized by solid phase peptide synthesis employing the Fmoc protecting strategy (Fmoc SPPS). A furan moiety was designed to place in the middle and the end of the PNA strand. Furan is known to be an acid-sensitive molecule. The modification of furan on the solid support was deemed difficult because in the Fmoc SPPS, the PNA must be cleaved from the solid support under acidic conditions (90 %TFA). Few examples reported the method to overcome this obstacle either to protect the furan by a temporary protection with maleimide[101] or by a post-synthetic modification via click-chemistry.[100] However, both methods required the synthesis of the specific monomers (maleimide protected furan or azide-modified monomers) that involves many preparation steps. Thus, we propose an alternative post-synthetic method that is more efficient and versatile by incorporating the furan part after the cleavage step via an amide bond formation as shown in **Figure 3.2**. However, the amino side chain of lysine, which is normally incorporated into the PNA sequence to improve its water solubility, needs to be protected to avoid reactions at this site. This was achieved by permanent blocking of the amino group by dimethylation prior to the post-synthetic modification with the furan moiety.[107] The dimethylated lysine should still serve as a solubility enhancer and can be left as such.

For the synthesis of the terminally modified PNA(T), the PNA precursor was synthesized by the manual Fmoc-SPPS. The success of the synthesis was confirmed by a full deprotection and cleavage of a small amount of the PNA from the resin followed by MALDI-TOF MS analysis which showed a mass peak at 3513 Da corresponded to the expected PNA sequence (**Figure 3.3**). Next, the dimethylation reaction was performed by reductive alkylation with excess formaldehyde in the presence of $NaBH_3CN$. The successful incorporation of the two methyl groups was confirmed by an increase of the mass by +29 Da. After that, the exocyclic amino protecting groups of the nucleobase on the PNA strand was removed by heating with a solution of ammonia:1,4-dioxane (1:1). This also simultaneously removed the N-terminal Fmoc group. A hydrophilic linker (Fmoc-AEEA or "O-linker") was next

incorporated as the last residue. The Fmoc-protected group was removed prior to the cleavage step, and the successful incorporation of the AEEA linker was confirmed by the mass increase to 3688 Da (+145 Da). Finally, the PNA was cleaved from the solid support and post-synthetically modified with the pentafluorophenyl-activated 3-(2-furyl) propionic acid to form an amide bond with the free *N*-terminal amino group of the O-linker on the PNA in the solution phase. The post-synthetic reaction took place smoothly after 10 min to provide the desired PNA as shown by the mass increase to 3812.3 (+124 Da).
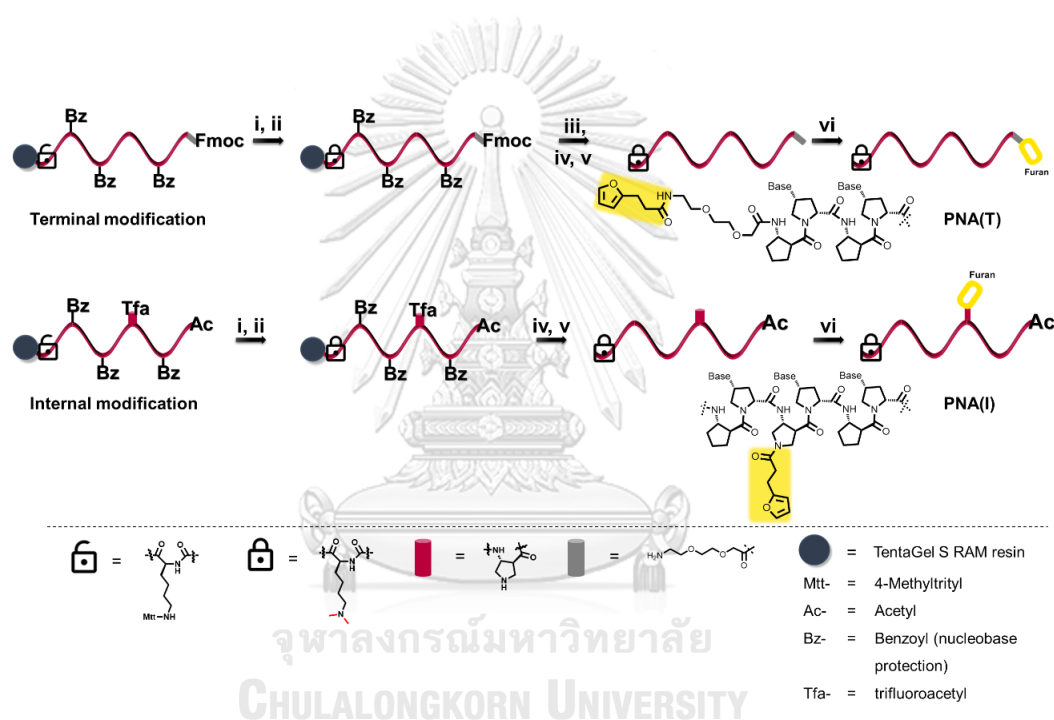
Figure 3.2 A scheme for the synthesis of furan-modified acpcPNA by the post-synthetic methodology ; conditions: **i)** 5% TFA in DCM; **ii)** 30 equiv. formaldehyde, NaBH$_3$CN, AcOH, MeOH, 5 h; **iii)** 20% piperidine in DMF containing 2% DBU; **iv)** NH$_3$: 1,4-dioxane (1:1), 65 °C, overnight; **v)** 95% TFA, 3 hr; **vi)** 10 equiv. pentafluorophenyl 3-(2-furyl)propionate, 10 equiv. DIPEA, DMF, 10 min.
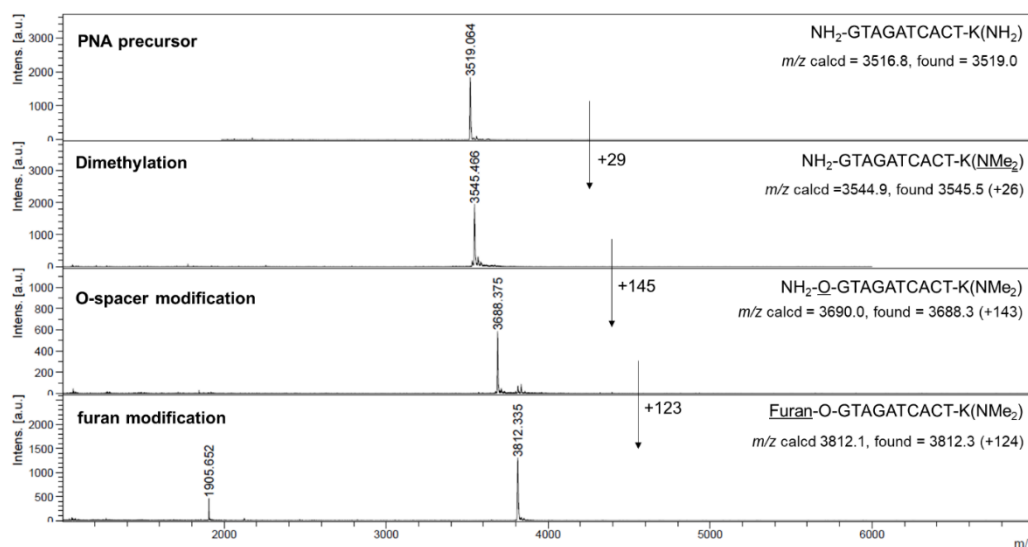
Figure 3.3 MALDI-TOF mass spectra demonstrating the successful synthesis of PNA(T). From top to bottom: crude of PNA(T) precursor, dimethylation, O-spacer modification and furan modification

For the synthesis of the internally-furan-modified acpcPNA, PNA(I), the precursor was synthesized similar to the PNA(T) but one ACPC spacer in the middle of the PNA strand was replaced by a Tfa-protected APC spacer to provide a handle for subsequent furan modification. The synthesis and modification steps thus followed the same protocol as described for PNA(T), but without the incorporation of the O-linker and the N-terminal amino group was capped by acetylation. **Figure 3.4** shows the MALDI-TOF MS trace of the PNA(I) synthesis. The crude PNA(I) precursor is presented on the top (3555 Da). Then, dimethylation reaction was performed to obtain a new product showing a mass of 3584 Da (+29 Da). After the nucleobase side chain as well as APC spacer deprotection by a hot ammonia:1,4-dioxane solution, the PNA was cleaved from the solid support by TFA treatment. Finally, the crude PNA was reacted with Pfp-activated 3-(2-furyl) propionic acid in the solution phase leading to the PNA(I) which showed the expected mass at 3708 Da (+123 Da). The crude PNA(I) and PNA(T) were purified to >90% purity by RP-HPLC and characterized by MALDI-TOF MS.
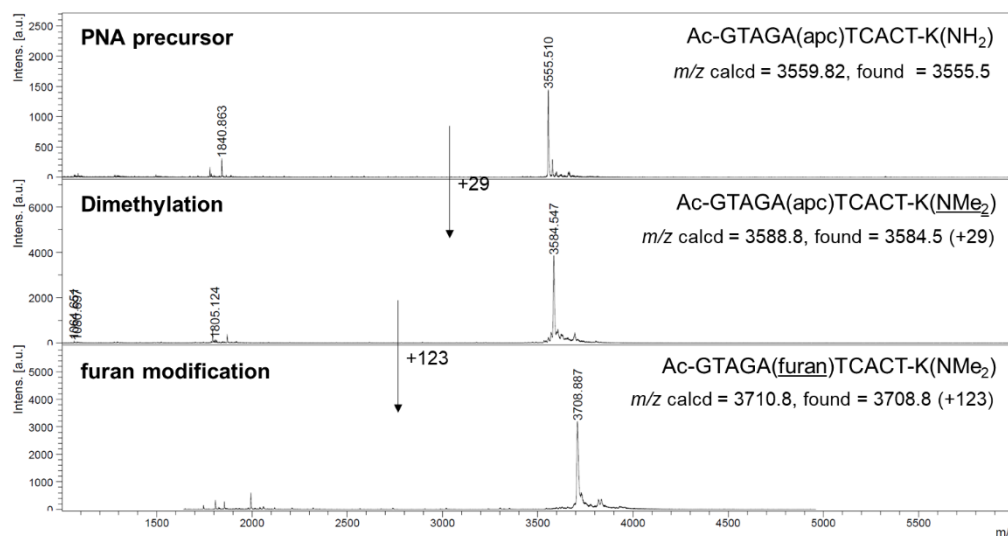
Figure 3.4 MALDI-TOF mass spectra demonstrating the successful synthesis of PNA(I). From top to bottom: crude of PNA(I) precursor, dimethylation and furan modification

Next, the PNA(T) and PNA(I) were investigated for their cross-linking efficiencies towards complementary DNA targets (DNA-com) involving five-base at the overhang sequences (DNA-C5 – DNA-T5), mismatched sequences (DNA-dm1 – DNA-dm2, DNA-sm1 – DNA-sm6) and C-inserted sequences (DNA-L1 – DNA-L8). All the target DNA sequences used in the study are presented in **Table 3.1**.

Table 3.1 PNA and DNA sequences used in this study

name	sequence[a] (5'-3')
DNA-com	AGTGATCTAC
DNA-C5	AGTGATCTACCCCCC
DNA-A5	AGTGATCTACAAAAA
DNA-G5	AGTGATCTACGGGGG
DNA-T5	AGTGATCTACTTTTT
DNA-dm1	AGTGATCT<u>TG</u>CCCCCC
DNA-dm2	AGTG<u>T</u>T<u>G</u>TACCCCCC
DNA-sm1	AGTGA<u>A</u>CTACCCCCC
DNA-sm2	AGTGA<u>C</u>CTACCCCCC
DNA-sm3	AGTGA<u>G</u>CTACCCCCC
DNA-sm4	AGTGATCTA<u>A</u>CCCCC
DNA-sm5	AGTGATCTA<u>G</u>CCCCC
DNA-sm6	AGTGATCTA<u>T</u>CCCCC
DNA-L1	A*C*GTGATCTAC
DNA-L2	AG*C*TGATCTAC
DNA-L3	AGT*C*GATCTAC
DNA-L4	AGTG*C*ATCTAC
DNA-L5	AGTGA*C*TCTAC
DNA-L6	AGTGATC*C*TAC
DNA-L7	AGTGATCT*C*AC
DNA-L8	AGTGATCTAC*C*
DNA-sm-L6	AGTGA<u>A</u>CCTAC

[a]The mismatched bases in the DNA targets are underlined. The base insertion position is italicized.

3.1.2 Internal cross-linking and terminal cross-linking reactions
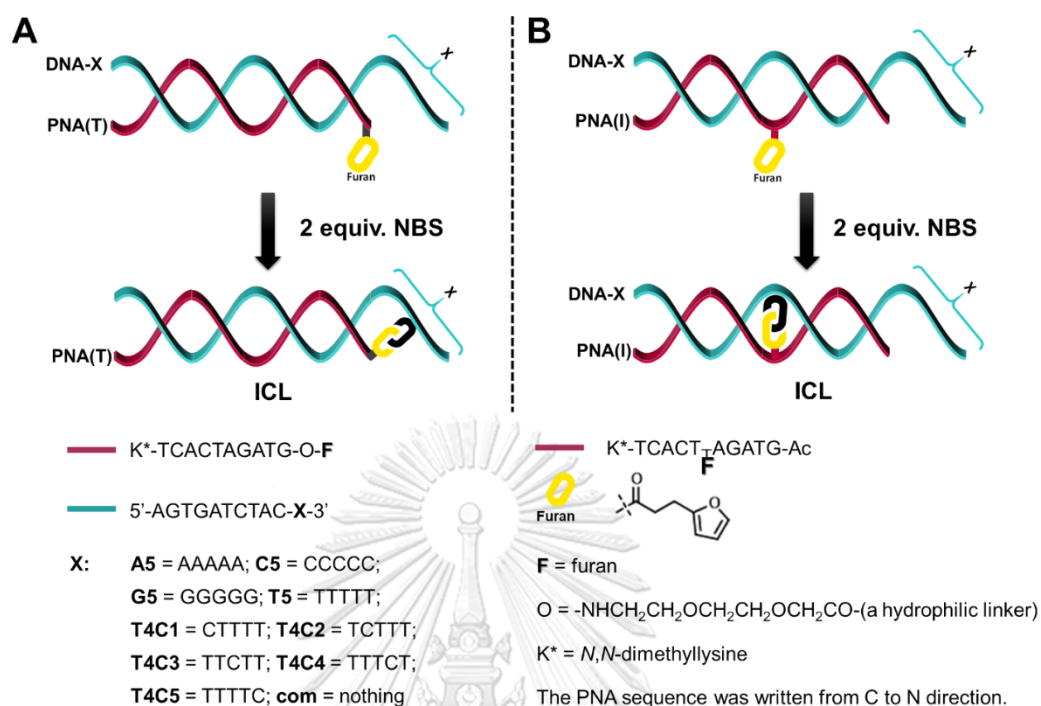
Figure 3.5 A schematic illustration of the cross-linking reaction between PNA(I) and PNA(T) towards complementary DNA targets

This study first investigated the effect of the position of the furan on the PNA strand (internal and terminal positions). The effect of the furan position towards the cross-linking efficiency was initially determined by melting temperature experiments. This experiment is a commonly used technique to determine the stability of DNA duplexes. The absorbance at 260 nm of the sample was monitored as a function of temperature during a slow heating. The hydrogen bonds that hold the base pairs are broken when the temperature was sufficiently high, resulting in a small but noticeable increase of the absorbance (hyperchromism). The plot between the absorbance (typically at 260 nm) as a function of temperature appears as a sigmoidal curve called "melting curve". Melting temperature (T_m), which represents the temperature at which the duplex is half-dissociated, is the temperature at the steepest part of the melting curve. In practice, the T_m can be obtained from the maxima of the first derivative plot of the melting curve. The melting curves of the cross-linking products was previously shown to increase in cross-linked products

which can be explained by the formation of a more stable intramolecular duplex as compared to the intermolecular duplex in the non-cross-linking cases.[59]

3.1.2.1 Furan-PNA probe activation

a) NBS activation

Furan activation was previously achieved by cytochrome P450, singlet oxygen or *N*-bromosuccinimide (NBS) oxidations. NBS was chosen as the oxidizing agent in this study which mainly involved in vitro reactions based on its simplicity and efficiency. However, it should be noted that this condition cannot be applied for in vivo or intracellular studies, and side reactions generating brominated side products had been observed in some cases.[108] The preliminary experiments focuses on the studies of activation of furan-modified acpcPNA by MALDI-TOF MS. The PNA(T) (5 µM) in 10 mM sodium phosphate buffer containing 100 mM NaCl was activated by treatment with NBS (0.5 mM) and the reaction mixture was aliquot sampled to detect the mass. After 15 min following the NBS addition, the original PNA signal at m/z = 3809 disappeared and a new signal at m/z = 3826 (+17 Da) was observed. The mass increase corresponded to addition of an oxygen atom to the furan to obtain aldehyde product as shown in **Figure 3.6**. No brominated byproduct was observed, and the activation was completed in 15 min. Thus, the NBS could efficiently activated the furan probe and only 2 equiv. of NBS was sufficient to activate the furan probe. This was considerably smaller than the 4 equiv. employed in the previous reports.[55] Attempted to use only 1 equiv. of NBS resulted in an incomplete activation of the PNA(T) probe as demonstrated by MALDI-TOF MS.
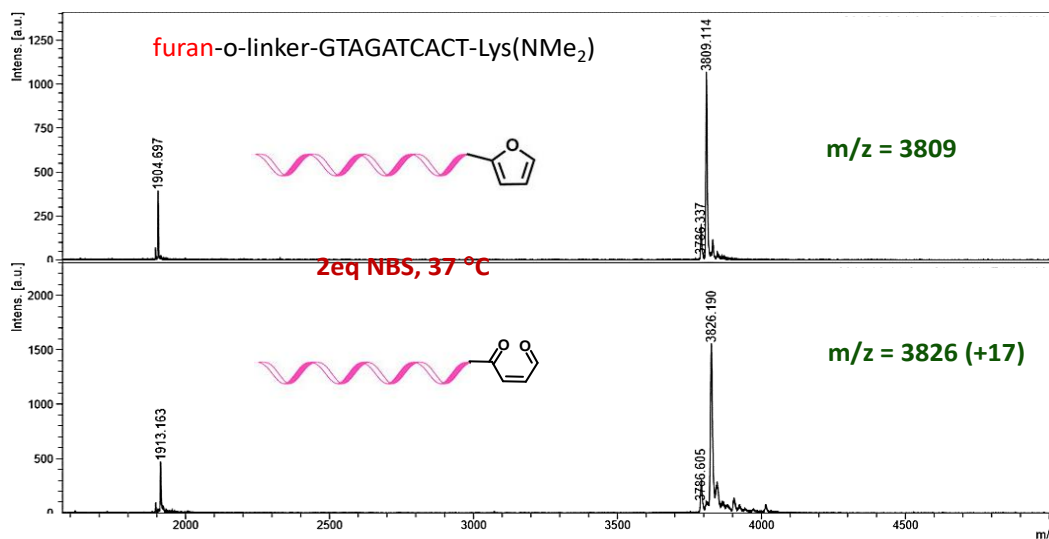
Figure 3.6 MALDI-TOF MS traces showing the activation of PNA(T) by NBS (2 equiv.)

b) Singlet oxygen oxidation

In addition, the activation of the furan acpcPNA probes by oxidation with singlet oxygen was also studied. The reaction was performed with PNA(T) (10 µM) in the presence of methylene blue (5 µM). After 10 min of the irradiation of the red light (LED source, 610-700 nm), the original PNA mass disappeared and the desired oxidation product (M+16) was formed only in very small amounts (**Figure 3.7**). The mass of the major product is larger than the original PNA by 32 Da (M+16+16). Additional side products containing multiple +16 mass unit difference were also observed. It was proposed that the M+32 might be the cyclized form of the aldehyde, and the multiple +16 peaks are due to overoxidation products as shown in **Figure 3.8**.
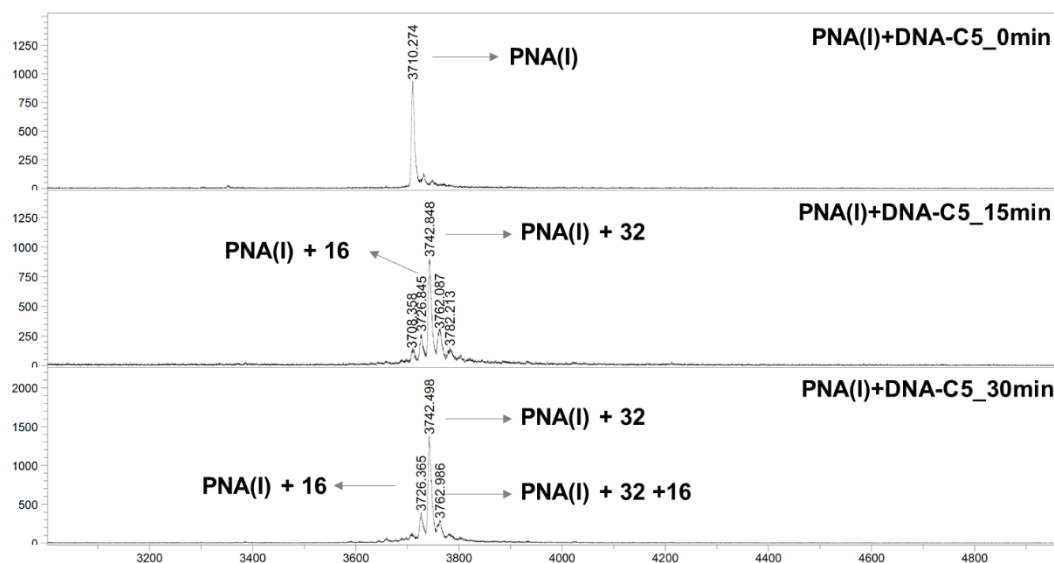
Figure 3.7 MALDI-TOF MS traces showing the activation of PNA(I) by red light/methylene blue
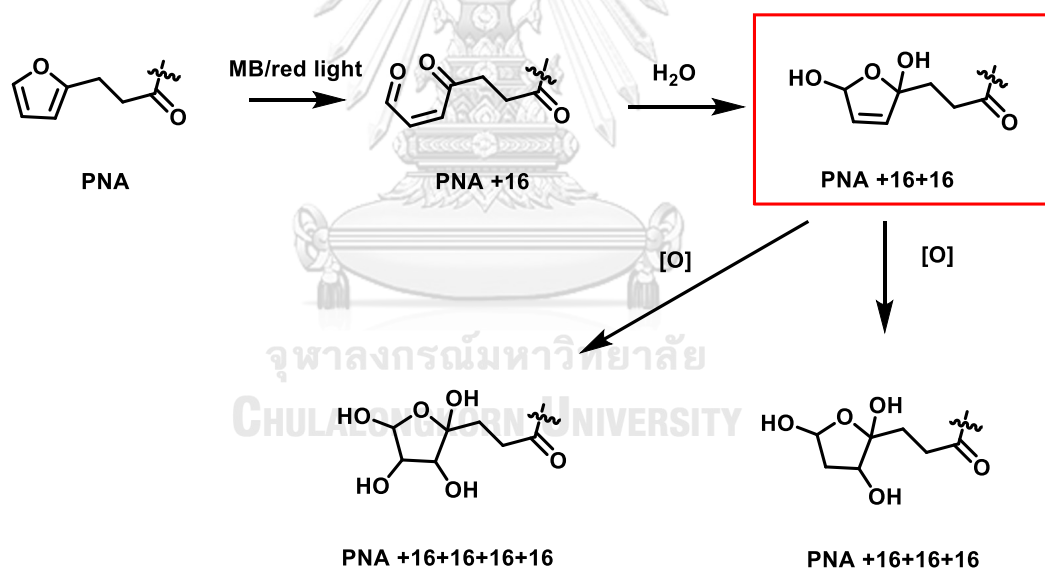
Figure 3.8 A proposed mechanism for the formation of multiple oxidation products in the activation of PNA(T) by red light/methylene blue

3.1.2.2 Cross-linking study

The cross-linking between PNA and DNA was preliminarily evaluated by performing the denaturing PAGE and melting temperature studies of the cross-linking reaction between PNA(I) or PNA(T) and their complementary DNA (DNA-com: 5'-AGTGATCTAC-3'). For the melting temperature studies, no change in the melting temperature was observed for both PNAs (**Figures 3.9A** and **3.9C**). It should be noted that the melting curves were flattened in both cases, indicating decomposition of either the PNA or DNA targets, or both. It was previously observed that NBS can cause peptide degradation by backbone cleavage.[109] In addition, the PAGE result revealed that no new band in the gel (lane 1 in **Figure 3.9B** and lane 3 in **Figure 3.9D**). All the results suggest that no cross-linked product was formed, even though the last base of the DNA strand carrying the cytosine which is known to readily form cross-links with furan probe.[61, 100, 101, 110] Moreover, when DNA-T5 and DNA-T4C5 were employed as the target, it was clearly shown that the last C-base of the complementary DNA did not involve in the cross-linking reaction. As shown in **Figure 3.9B**, It was found that the cross-linked product of the DNA-T4C5 (to the last C-base of the overhang region) with PNA(T) was observed in lane 3, while no cross-linked product was formed in the case of DNA-T5 (lane 2). The failure of both PNA(I) and PNA(T) to form cross-linked products with its exactly complementary DNA target is that all the DNA bases were involved in the Watson-Crick base pairing with the PNA strand, hence they were unavailable to form the cross-link with the activated furan as also noticed by previous reports.[61, 63] Thus, we designed new DNA target sequences by adding five-cytosine at the 3'-end of the DNA-com strand to provide DNA-C5 (5'-AGTGATCTACCCCCC-3') to be used in further studies.
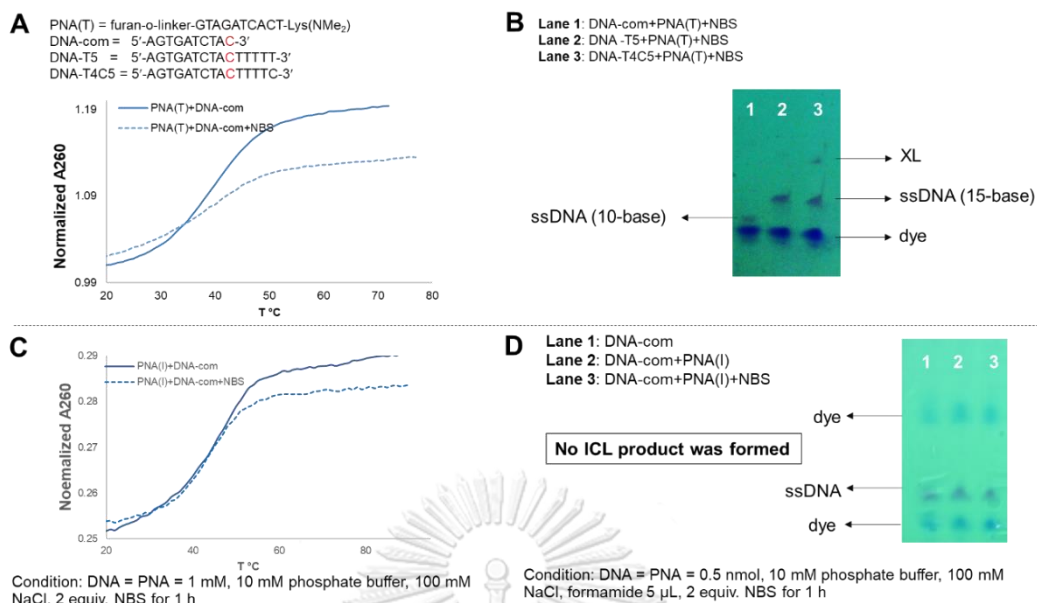
Figure 3.9 (A) Melting curves of cross-linking reaction of DNA-com with PNA(T) and with PNA(I) (C)at before (solid line) and after (dash line) the NBS addition (2 equiv., 37 °C, 1 h). Conditions: [DNA] = [PNA] = 1 μM, 100 mM NaCl, 10 mM sodium phosphate buffer, pH 7.0, heating rate 1.0 °C/min. (B) denaturing PAGE results of PNA(T) with DNA-com, DNA-T5, and DNA-T4C5 (in lane 1, 2, and 3, respectively) in presence of NBS. (D) denaturing PAGE results of PNA(I) with DNA-com in presence (lane 2) and absence (lane 3) of NBS. Conditions: [DNA] = [PNA] = 20 nmol in the total volume = 10 μL, 2 equiv. NBS (1 equiv. every 15 min), for 30 min. The experiment was performed in 17% polyacrylamide gel containing 7 M urea at 225 V for 1 h.

In the next studies, the PNA(I) and PNA(T) were cross-linked with a long complementary DNA target DNA(C5) which had five consecutive C-bases overhang at the 3'-end. The multiple C tract was added to ensure that there will always be a C base available for the cross-linking with the furan modification at the *N*-terminus (equivalent to the 5'-end) of the PNA(T) strand. The T_m studies results are shown in **Figure 3.10**. Prior to the NBS activation, the sigmoidal curves of the duplexes between DNA-C5 and PNA(I) as well as PNA(T) were observed as expected, which suggested the formation of a stable PNA-DNA duplex in both cases. The melting temperature of the internally modified PNA(I)-DNA-C5 duplex was observed at 44 °C, which was considerably lower than the unmodified PNA-DNA duplex (62 °C). The

decrease of melting the melting temperature suggested that the incorporation of the furan moiety in the middle of the PNA strand resulted in a significant steric hindrance which destabilized the duplex. In the case of the terminally modified PNA(T)-DNA-C5 duplex, the melting temperature (61 °C) was similar to that of unmodified PNA-DNA duplex (62 °C), indicating that the furan placed at the end of the strand has no effects on the duplex stability.
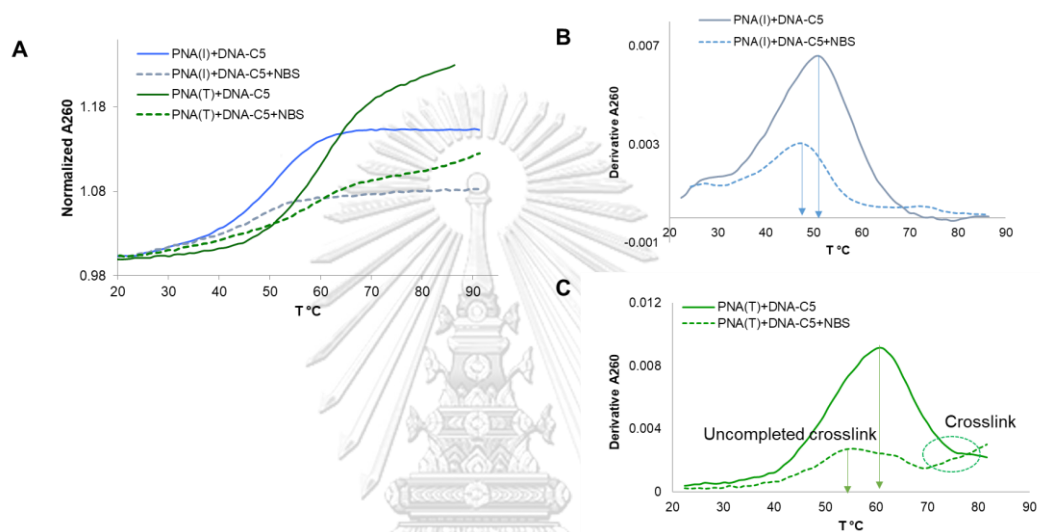
Figure 3.10 (A) Melting curves of cross-linking reaction of PNA(I) and PNA(T) with DNA-C5 before (solid line) and after (dash line) the NBS addition (2 equiv., 37 °C, 1 h). Conditions: [DNA] = [PNA] = 1 µM, 100 mM NaCl, 10 mM sodium phosphate buffer, pH 7.0, heating rate 1.0 °C/min. (B) derivative plot of cross-linking result of PNA(I) at before (blue solid line) and after (blue dash line) NBS addition (C) derivative plot of cross-linking result of PNA(T) at before (green solid line) and after (green dash line) NBS addition

After the addition of 2 equiv. of NBS to the PNA-DNA duplexes in each case, the melting curves were measured again under identical conditions. The results in **Figure 3.10** revealed that in the case of PNA(I), the melting curves in the presence and absence of NBS were flattened, with a slightly decreased T_m. The absence of T_m increase indicated that no cross-linking occurred in this internally modified system. On the contrary, a significant change in the melting profile was observed in the case of PNA(T) following the NBS activation. An additional sigmoidal curve was observed as shown by the significant rise of the absorbance after the first melting instead of the steady curve as observed before adding the NBS. The results suggest that after the NBS activation, the duplex stability increased which indicated that the cross-linking reaction between the PNA(T) and DNA-C5 had occurred. The second melting was not complete even at almost boiling water temperature. Hence it was estimated that the T_m of the cross-linked duplex was higher than 95 °C. However, the cross-linking reaction appeared to be incomplete as another melting curve with a T_m of around 55 °C (**Figure 3.10C**) which was slightly lower than the curve before NBS activated condition (61 °C) was still present. It should be noted that a drop in the hyperchromicity, resulting in flattening of the T_m curves, was observed in both cases following the NBS addition similar to the experiments with the fully complementary DNA target DNA-com (**Figures 3.9A** and **3.9C**) above.

Based on the results from the melting experiments discussed above, it was concluded that the cross-linking reaction could occur only in the case of terminally modified PNA(T) but not the internally labeled PNA(I). The results were further confirmed by denaturing polyacrylamide gel experiment (PAGE), RP-HPLC and MALDI-TOF MS as will be discussed next.

Figure 3.11 Denaturing PAGE results of terminally [PNA(T)] and internally furan-modified PNA [PNA(I)] with DNA-C5 in presence (+) and absence (-) of NBS. Conditions: [DNA] = [PNA] = 20 nmol in the total volume = 10 µL, 2 equiv. NBS (1 equiv. every 15 min), for 30 min. The experiment was performed in 17% polyacrylamide gel containing 7 M urea at 225 V for 1 h.

The cross-linking results were further confirmed by denaturing PAGE. The cross-linking reaction were performed in 20 µM of each PNA and DNA strands employing 2 equiv. NBS as an activator. As shown in **Figure 3.11**, lane 1 shows a fast moving band of the single stranded DNA used as a reference. Lane 2 and 3 are the internally furan-modified acpcPNA [PNA(I)] with complementary DNA (DNA-C5) before and after the NBS activation, respectively. Lane 4 and 5 are the terminally furan-modified acpcPNA [PNA(T)] with DNA-C5 before and after the NBS activation, respectively. The results revealed that in the absence of NBS, only the fast-moving band of single stranded DNA was observed (lanes 1, 2, and 4). Since the gel was run under denaturing conditions, the PNA-DNA duplex should completely dissociate and move separately in the gel unless a cross-link is formed. The PNA band could not move into the gel and thus was not detected due to the uncharged nature of the PNA backbone and the positively charged lysine modification. Gratifyingly, a slow-moving band corresponding to the cross-link product was observed in lane 5 that consisted of the mixture of PNA(T) and DNA-C5 after the NBS treatment. In addition to this cross-linked product band, a very small amount of single stranded DNA band

remained visible at the bottom of the gel. No cross-linked product band was observed in the case of similar experiments between PNA(I) and DNA-C5 as shown in lane 3. These results support the conclusion from the T_m experiments above that the cross-linking reaction was possible only in the case of PNA(T).

The different cross-linking behaviors between PNA(T) and PNA(I) to the same DNA target (DNA-C5) could be explained as follows. The furan at the end of the strand of PNA(T) is flexible enough to react with the unpaired cytosine bases in the overhang region that is readily available for the cross-linking reaction. In contrast, the furan in the middle of the strand PNA(I) are unlikely to react with the nearby A and C nucleobases because the exocyclic amino groups of these bases would get involved in the Watson and Click base pairing would therefore be unavailable for the cross-linking reaction.

The formation of the cross-linked product between PNA(T) and DNA-C5 following NBS treatment was further confirmed by MALDI-TOF MS. The cross-linking product was first purified by RP-HPLC and concentrated to 5 μM before submitting to the MALDI-TOF MS measurement (performed at the Institute of Molecular Biosciences, Mahidol University). The sample was desalted by treatment with Dowex® 50WX8 beads before being detected. 3-hydroxypiconilic acid: ammonium citrate (9:1) was used as a matrix for the cross-linked products. The mass spectrum was shown in **Figure A24**, which revealed the expected mass of the cross-linked product: PNA(T)+DNA-C5 ($C_{322}H_{405}N_{112}O_{124}P_{14}$) observed mass: 8308.08 Da (Mass + K); calculated mass 8269.87 Da. Accordingly, all three techniques confirmed the formation of a cross-linked product between PNA(T) and DNA-C5 upon activation by NBS.

3.1.3 Selectivity of PNA(T) towards different DNA nucleobases

3.1.3.1 Denaturing PAGE with unlabeled DNA targets
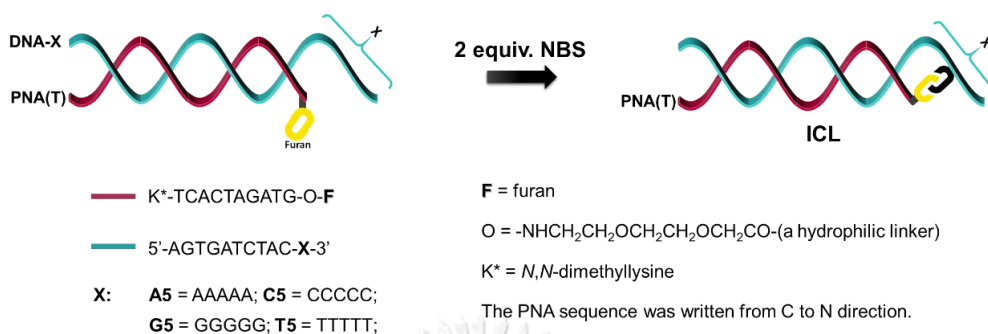
Figure 3.12 Representation of the cross-linking reaction of PNA(T) towards complementary DNA with the different overhanging bases

This part of the research focuses on the cross-linking reactivity of the PNA(T) towards different DNA nucleobases. After showing that the cross-linking reaction could occur only with the PNA(T) probe, the selectivity of the PNA(T) was further investigated by changing the 5-base at the overhang part from C in the DNA-C5 to other bases in DNA-A5, DNA-T5, and DNA-G5. The product(s) from the cross-linking reactions induced by NBS were analyzed by denaturing PAGE as shown in **Figure 3.13**. Lane 1 is single stranded DNA-A5. Lane 2 and 3 are the duplex of PNA(T) and DNA-A5 before and after the NBS addition, respectively. Similarly, lane 4-5, 6-7, and 8-9 represent the duplexes of PNA(T) and DNA-C5, DNA-G5 and DNA-T5 before and after activation, respectively. Previous works reported that furan-modified DNA and aegPNA probes can only cross-link with the nucleobase adenine (A) and cytosine (C) via their exocyclic amino groups.[61, 100, 110] In this work, the cross-linked products were observed in lanes 3, 5 and 7. These cross-linking bands corresponded to the cross-linking products of PNA(T) with the adenine, cytosine and guanine bases, respectively, in the overhang region of the DNA strand. No cross-linked product was observed in the case of thymine (lane 9). These results suggested that all the nucleobases carrying an exocyclic amino group including A, C and G could form the cross-linked product with the furan PNA(T) probe. This is the first observation that guanine (G) could participate in the cross-linking process with furan moiety. From the gel

experiment in lane 6, the DNA-G5 itself could form aggregated structures as shown in the slowest moving and smearing band in lane 6. CD-spectrometry of the DNA-G5 confirmed the presence of the secondary structure, most likely a parallel G-quadruplex, which showed distinctive features when compared to other DNAs (see supporting information **Figure A25**).
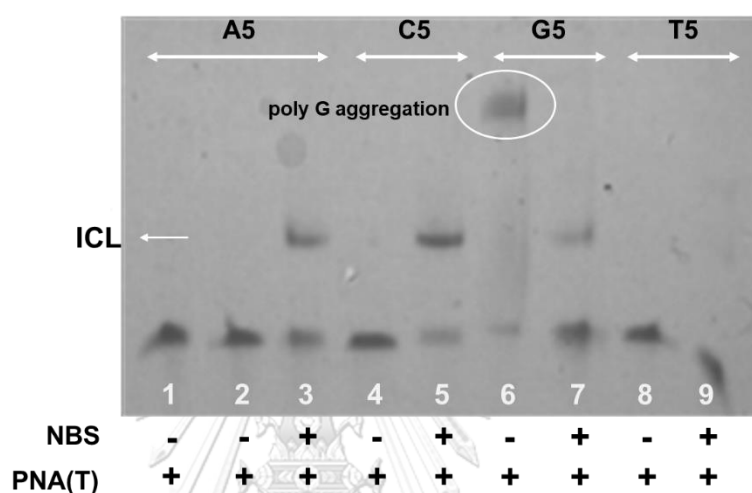
Figure 3.13 Denaturing PAGE results of terminally furan-modified PNA [PNA(T)] with different DNA bases changing the 5-base overhang in the DNA-C5 to other bases to give DNA-A5, DNA-T5, and DNA-G5. Conditions: [DNA] = [PNA] = 20 nmol in the total volume = 10 μL, 2 equiv. NBS (1 equiv. every 15 min), for 30 min. The experiment was performed in 17% polyacrylamide gel containing 7 M urea at 225 V for 1 h.

3.1.3.2 Denaturing PAGE with FAM-labeled DNA targets

To achieve a better visualization in the PAGE, we modified the targeted DNA strands (DNA-A5, DNA-C5, DNA-G5, DNA-T5), with fluorescein (FAM) at the 5'-end to give DNA-FA5, DNA-FC5, DNA-FG5, and DNA-FT5. It was expected to observe fluorescent-cross-linking bands under black light. The same cross-linking conditions and gel electrophoresis as for the non-fluorescent targets were performed in this experiment. It was found that when the first equivalent of NBS was added to a solution of PNA:DNA duplex, the color of the solutions already changed from green to pink. Moreover, the FAM-labeled DNA lost its fluorescence. When checked with

the literature, it was found that the fluorescein was brominated by NBS to obtain the pink species which is only weakly fluorescent as shown in **Figure 3.14**.[111] Thus, the 2 equiv. of NBS added to activate the furan moiety would not be sufficient because the fluorescein competed with the furan moiety to react with the NBS. Nevertheless, the cross-linking reaction still occurred as shown by the presence of a slow-moving and weakly fluorescent band in lane 5 of **Figure 3.15** and in lanes 3, 5, 7 in **Figure 3.16**. While the use of FAM-labeled DNA as the target DNA for the NBS-activated furan cross-linking probe did not offer advantages over the UV shadowing experiments, the results did confirm that the cross-linking was possible only with PNA(T) and DNA targets with overhanging A, C and G.
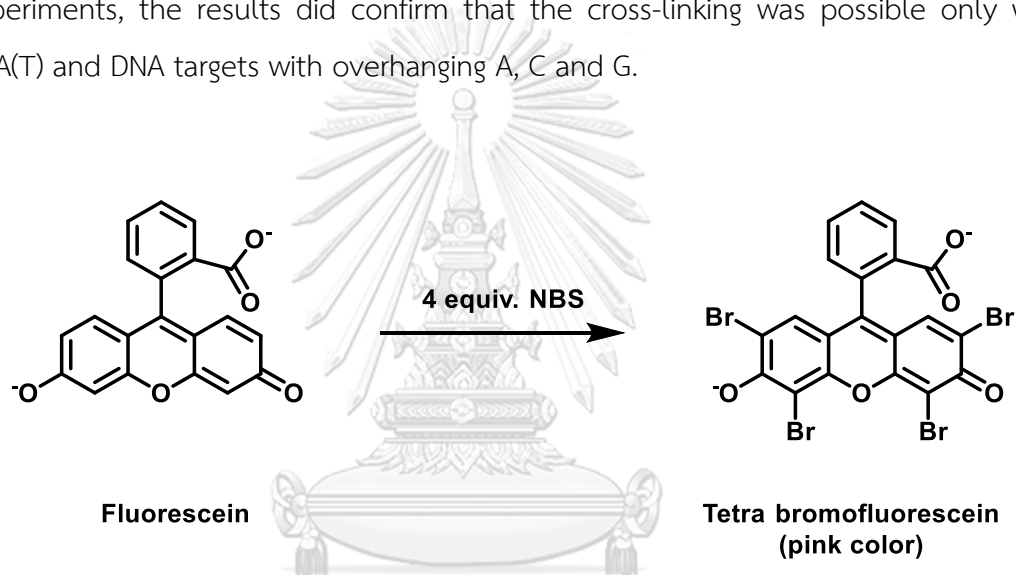
Figure 3.14 Bromination of fluorescein (FAM) by 4 equiv. NBS providing a tetra bromofluorescein (pink color)

Under UV light   Under black light

PNA(T)   : **furan-O**-GTAGATCACT-K(NMe₂)
PNA(I)   : Ac-GTAGA(**furan**)TCACT-K(NMe₂)
DNA-FC5: 5'-FAM-AGTGATCTACCCCCC-3'

Lane 1 = DNA-FC5
Lane 2 = DNA-FC5+PNA(I)
Lane 3 = DNA-FC5+PNA(I)+NBS
Lane 4 = DNA-FC5+PNA(T)
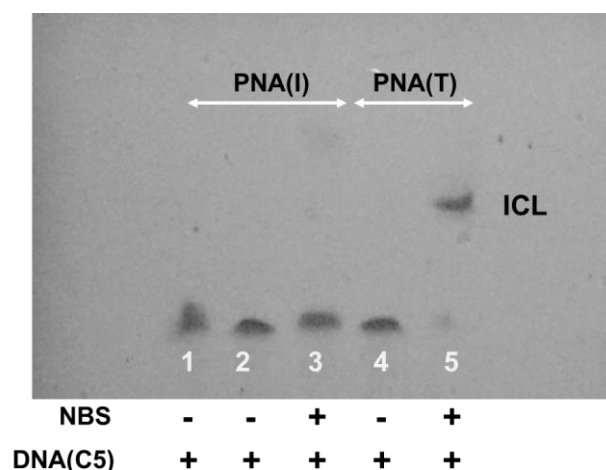Lane 5 = DNA-FC5+PNA(T)+NBS
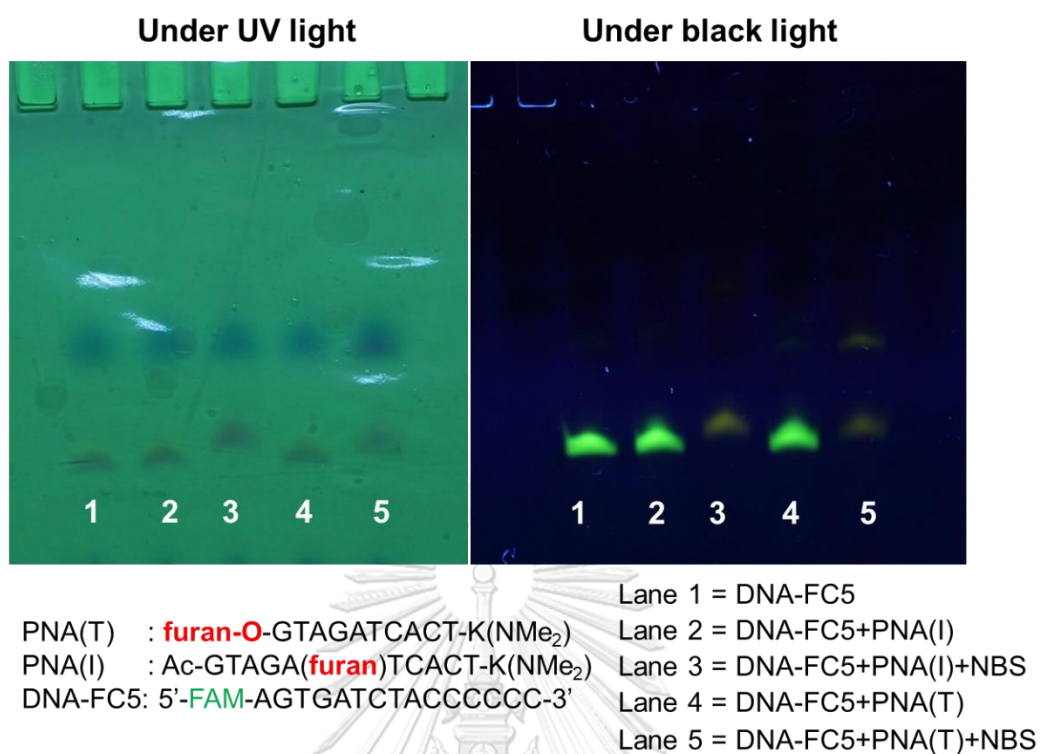
Figure 3.15 Denaturing PAGE result of terminally [PNA(T)] and internally furan-modified PNA [PNA(I)] with FAM-labelling DNA-FC5 in the presence (+) and absence (-) of the NBS. Conditions: [DNA] = [PNA] = 20 nmol in the total volume = 10 µL, 2 equiv. NBS (1 equiv. every 15 min), for 30 min. The experiment was performed in 17% polyacrylamide gel containing 7 M urea at 225 V for 1 h.

Under UV light	Under black light

DNA-FA5 : 5'-FAM-AGTGATCTACAAAAA-3'
DNA-FC5 : 5'-FAM-AGTGATCTACCCCCC-3'
DNA-FG5 : 5'-FAM-AGTGATCTACGGGGG-3'
DNA-FT5 : 5'-FAM-AGTGATCTACTTTTT-3'
PNA(T) : **furan-O**-GTAGATCACT-K(NMe₂)

Lane 1 = DNA-FA5
Lane 2 = DNA-FA5+PNA(T)
Lane 3 = DNA-FA5+PNA(T)+NBS
Lane 4 = DNA-FC5+PNA(T)
Lane 5 = DNA-FC5+PNA(T)+NBS
Lane 6 = DNA-FG5+PNA(T)
Lane 7 = DNA-FG5+PNA(T)+NBS
Lane 8 = DNA-FT5+PNA(T)
Lane 9 = DNA-FT5+PNA(T)+NBS
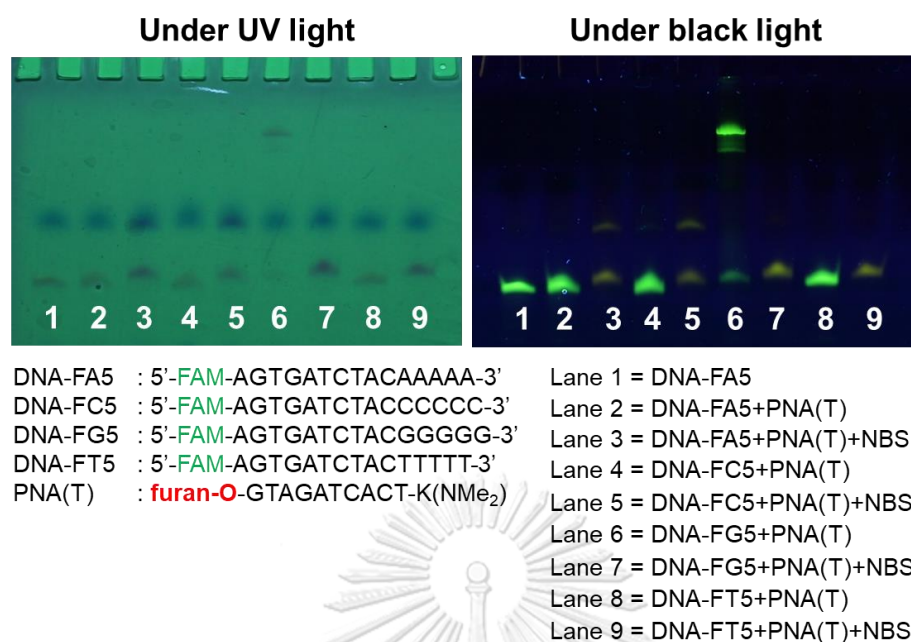
Figure 3.16 Denaturing PAGE result of terminally [PNA(T)] with different exocyclic nucleobase (A, C, G, T) on FAM-labelling DNA (DNA-FA5, DNA-FC5, DNA-FG5, DNA-FT5) in presence (+) and absence (-) of NBSConditions: [DNA] = [PNA] = 20 nmol in the total volume = 10 μL, 2 equiv. NBS (1 equiv. every 15 min), for 30 min. The experiment was performed in 17% polyacrylamide gel containing 7 M urea at 225 V for 1 h.

3.1.3.3 Melting temperature

The melting temperature experiments of PNA(T) towards DNA targets with different nucleobase overhangs (DNA-A5, DNA-C5, DNA-G5, DNA-T5) were further studied. The melting curves of the duplexes formed between PNA(T) with the DNA targets before the NBS activation is shown in **Figure 3.17**. The results showed that the T_m values in every case were very similar ($\sim$61 °C), indicating that the different nucleobases in the overhang region did not affect the duplex stability. After adding 2 equiv. of NBS, the melting curves became quite different from the original curves in most cases as shown in **Figure 3.17**. The exception was observed in the case of DNA-T5, whereby the melting curve and T_m were similar to those of the duplex before the NBS activation. This indicates that the DNA-T5 could not form the cross-linked product, which is in good agreement with the denaturing PAGE experiments. In the

corresponding cases of DNA-A5 (blue line), DNA-C5 (red line) and DNA-G5 (green line), the melting curves showed two transitions with a continuing rise of the melting curves after the first melting. The results indicate that the cross-linked products were formed in with the DNA carrying adenine, cytosine and guanine overhangs. The second transition was most pronounced in the case of DNA-C5. In this, and all other cases, the melting curves of the original PNA-DNA duplexes with T_m in the order of 60 °C remained in the system since the cross-linking reaction was not 100% complete. These results supported the denaturing PAGE results that the cross-linked product with PNA(T) was formed in the case of DNA-A5, C5, and G5 but not DNA-T5.
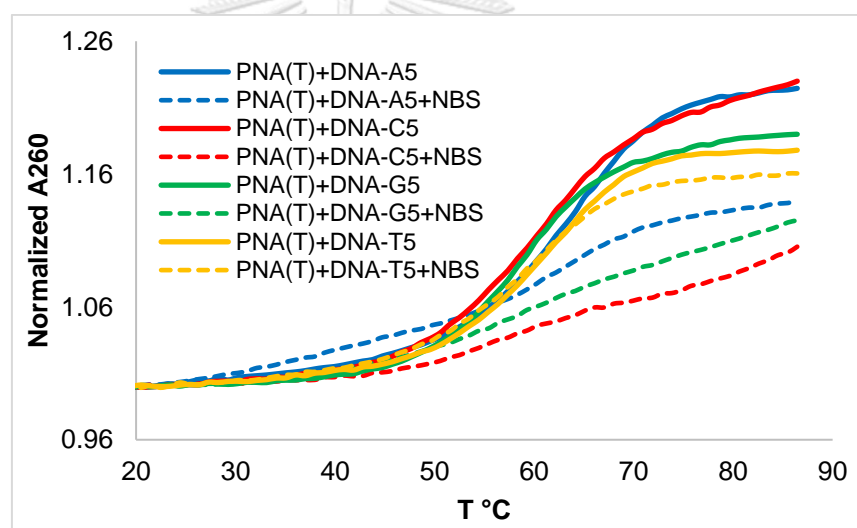
Figure 3.17 Melting curves of the duplexes of PNA(T) with different bases at the overhang (DNA-A5, DNA-C5, DNA-G5, DNA-T5] before (solid line) and after (dash line) the NBS addition (2 equiv., 37 °C, 1 h). Conditions: [DNA] = [PNA] = 1 μM, 100 mM NaCl, 10 mM sodium phosphate buffer, pH 7.0, heating rate 1.0 °C/min. The cross-linking was induced by addition of NBS (2 equiv.).

3.1.3.4 RP-HPLC

Next, the cross-linking of the PNA(T) targets was further studied by RP-HPLC to quantitatively determine the cross-linking efficiency. The duplexes were first formed by mixing the PNA strand (5 µM) with an excess of the DNA strand (7.5 µM). The DNA was deliberately added in excess to give a ratio of DNA: PNA at 1.5: 1 to ensure that the cross-linking efficiency was not determined by limiting amounts of the DNA. After annealing the PNA-DNA duplex by heating at 95 °C for 5 min followed by cooling at 0 °C (ice bath) for 1 h, 2 equiv. of freshly prepared stock solution NBS were added over 20 min period (1 equiv. every 10 min). The cross-linked products were generated almost immedieatly, within 10 min following the completion of the NBS addition. Based on the integration of signals from the HPLC chromatograms, the cross-linkings yield were calculated according to **Equation 1** where A_{ICL} is the peak area of the cross-linked product, ε_{ICL} is the molar extinction coefficient of the cross-linked product, which was estimated to be the sum of the molar extinction coefficients of the PNA and the DNA, A_{PNAsm} is the peak area of the starting PNA (without NBS), and ε_{PNA} is the molar extinction coefficient of the PNA.

Equation 1;
$$\% \text{ cross-linking yield} = \frac{A_{ICL}/\varepsilon_{ICL}}{A_{PNAsm}/\varepsilon_{PNA}} \times 100$$

After 20 min of the cross-linking reaction, the reaction mixtures were analyzed by RP-HPLC (**Figure 3.18**). In all cases, the PNA, DNA and cross-linking peaks were well-separated. Based on equation 1, the cross-linking efficiency measured at 20 min revealed that the DNA-C5 with cytosine overhangs could form the cross-linked product in the highest yield (71%) followed by the DNA-A5 (22%) and DNA-G5 (6%), respectively. Thus, the reactivity of the PNA(T) towards nucleobase can be ranked as C >> A > G and no cross-linking was observed for T. This is in good agreement with previously reported cross-linking reactions of furan-modified DNA as well as aegPNA probes,[59, 61, 62, 100, 112] with the exception that this is the first example to show that G can also participate in the cross-linking reaction. Moreover, the cross-linking

efficientcy was higher than the previous reports that showed only the small cross-linking band in the PAGE analysis (no % yield reported).[100]
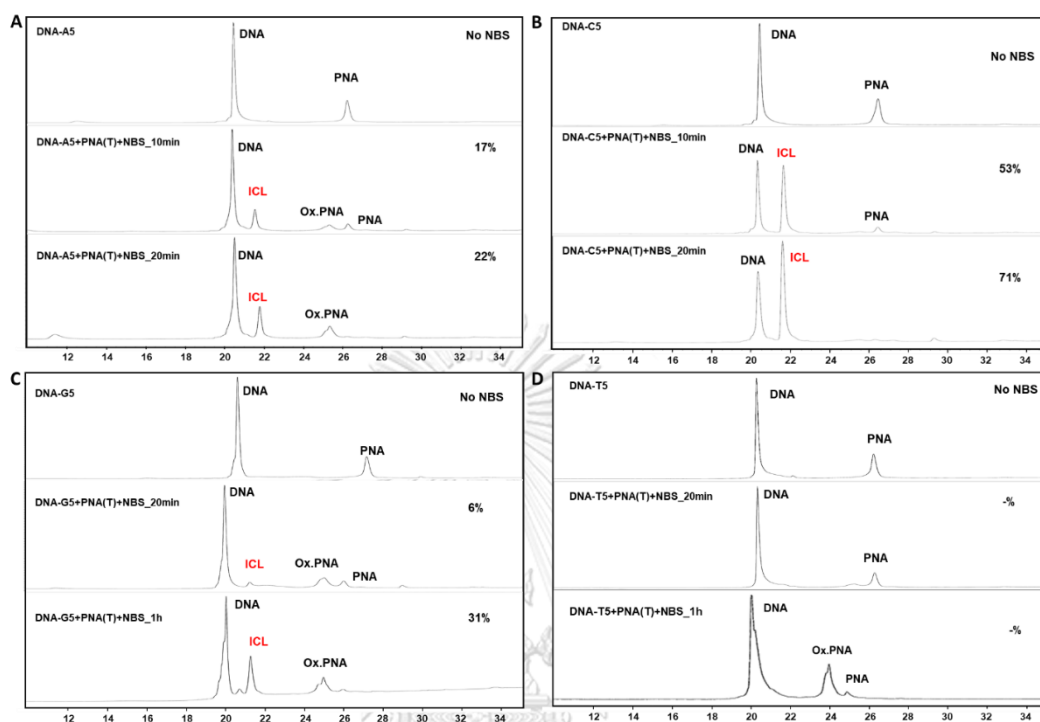
Figure 3.18 RP-HPLC chromatograms of cross-linking reactions between PNA(T) and its DNA complements carrying different overhang DNAs: DNA-A5 **A**; DNA-C5 **B**; DNA-G5 **C**; and DNA-T5 **D**, following oxidation with 2 equiv. NBS at 37 °C, 20 min to 1 h. After addition of NBS, the oxidized PNA (Ox. PNA) and cross-linked products were formed. Conditions:5%-80% MeCN in 1mM TEAA buffer, [DNA]:[PNA] = 7.5 μM:5 μM.

When the cross-linking reactions were allowed to proceed for longer time, the cross-linking efficiency also increased as shown in **Figure 3.18**. The products were analyzed at 10 min, 20 min and 1 h following the initial addition of the NBS to compare the reactivity of the PNA probe. The results revealed that the cross-linked product between the PNA(T) and DNA-A5 was formed in 17% at 10 min after the NBS addition (1 equiv.) (**Figure 3.18A**). The cross-linking slowly increased to 22% after 20 min of the reaction (immediately after the addition of the 2nd equiv. of NBS). In the case of DNA-C5, the reaction immediately produced the cross-linked product in 53% yield after the first 10 min after the NBS addition (1 equiv. NBS). The cross-linking

yield rapidly increased up to 71% after 20 min (2 equiv. NBS) (**Figure 3.18B**). For the DNA-G5, no cross-linked product was formed after the first 10 min. When the reaction time was extended to 20 min, a very small yield (6%) of the cross-linked product were produced. When the reaction of DNA-G5 was allowed to proceed until 1 h (from the initial addition of NBS, total = 2 equiv.), the cross-linking yield increased to 31% (**Figure 3.18C**). Although the cross-linking reaction of furan with G is slow, this is the first-time reporting that guanine can also participate in such cross-linking process with furan-modified probes. For the DNA targets with thymine overhangs, DNA-T5, no cross-linked product could be observed (**Figure 3.18D**). This is consistent with the proposed chemistry of the cross-linking (**Figure 3.1**) since thymine does not contain any exocyclic amino group that can participate in the cross-linking reaction with furan.

Table 3.2 Summary of cross-linking efficiency (% yield, as determined by RP-HPLC) of PNA(T) towards DNA-A5, DNA-C5, DNA-G5, and DNA-T5

condition	% cross-link		
	10 min[a]	20 min[b]	1 h[b]
PNA(T)+DNA-A5	17	22	N.D.
PNA(T)+DNA-C5	53	71	N.D.
PNA(T)+DNA-G5	N.D.	6	31
PNA(T)+DNA-T5	N.D.	-	-

[a]after addition of the first equiv. of NBS

[b]the 2nd equiv. of NBS was added after 10 min and waited for another 10 or 40 min.

N.D. = not determined; - = no cross-linked product peak was observed.

Conditions: 5%-80% MeCN in 1mM TEAA buffer, [DNA]:[PNA] = 7.5 µM:5 µM

3.1.4 Effect of position of C in the overhang region

K*-TCACTAGATG-O-**F**

5'-AGTGATCTAC-**X**-3'

X: **T4C1** = CTTTT; **T4C2** = TCTTT;
T4C3 = TTCTT; **T4C4** = TTTCT;
T4C5 = TTTTC; **com** = nothing

F = furan

O = -NHCH$_2$CH$_2$OCH$_2$CH$_2$OCH$_2$CO-(a hydrophilic linker)

K* = *N,N*-dimethyllysine

The PNA sequence was written from C to N direction.
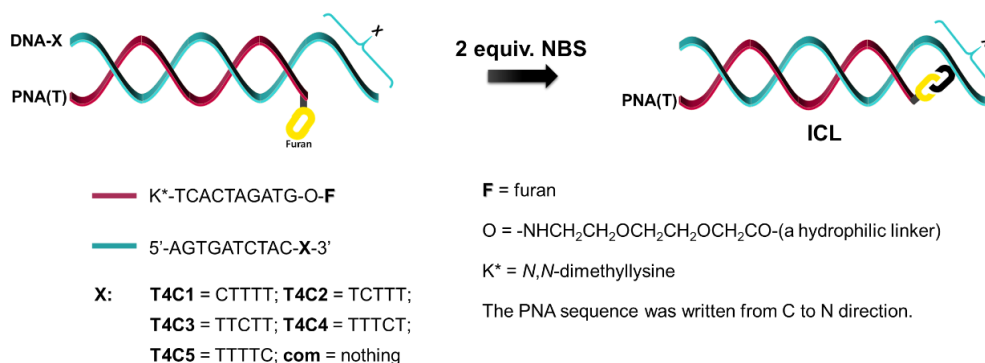
Figure 3.19 A schematic representation of the cross-linking reaction between the PNA(T) probe and DNA targets with various T4C1 overhangs

After confirming that the PNA(T) could undergo cross-linking formation with DNA-C5 which contains a five-cytosine overhang, the next study was performed to investigate the preferred position of the cytosine that form the cross-link. Since it is known from the previous experiments that T is unable to cross-link with the furan acpcPNA probe. We designed the DNA sequences in this experiment by replacing four cytosines in the overhang region with four thymine residues at different positions to form T4C1 overhangs in DNA-T4C1 – DNA-T4C5 instead of the C5 overhang in the DNA-C5. The cross-linking reaction was investigated by RP-HPLC and melting temperature experiments. In the melting temperature studies (**Table 3.3**), the T_m of the PNA(T) with DNA targets with different T4C1 overhangs were not different in the absence of NBS. However, in the presence of NBS, the melting temperatures were higher than non-activated condition (ΔT_m +10 to +14 °C). The results indicated that the cross-linking products can form with cytosine at any locations in the overhang region.

Table 3.3 Melting temperature (T_m) of PNA(T) with different T4C1 DNA targets

DNA target	sequence (5'-3')	T_m (°C) before cross-link	T_m (°C) after cross-link	%cross-link (20 min)
DNA-T4C1	AGTGATCTACCTTTT	60	73	38
DNA-T4C2	AGTGATCTACTCTTT	61	75	29
DNA-T4C3	AGTGATCTACTTCTT	61	75	26
DNA-T4C4	AGTGATCTACTTTCT	61	71	16
DNA-T4C5	AGTGATCTACTTTTC	61	71	11

Conditions: [DNA] = [PNA] = 1 µM, 100 mM NaCl, 10 mM sodium phosphate buffer, pH 7.0, heating rate 1.0 °C/min

The reactivity of the PNA(T) towards different T4C1 DNAs was further determined by RP-HPLC according to the protocols and conditions mentioned earlier. After 20 min of the NBS activation reaction, the cross-linked products were generated in with all T4C1 DNAs but the yield of the cross-linked product was different. Lower cross-linking yield was observed when the C-base was placed further away from the furan modification site on the PNA strand as shown in **Figure 3.20**. This could be explained by the flexibility of the overhang region that may allow folding back allowing the cross-linking reaction to occur regardless of the position of the C base.
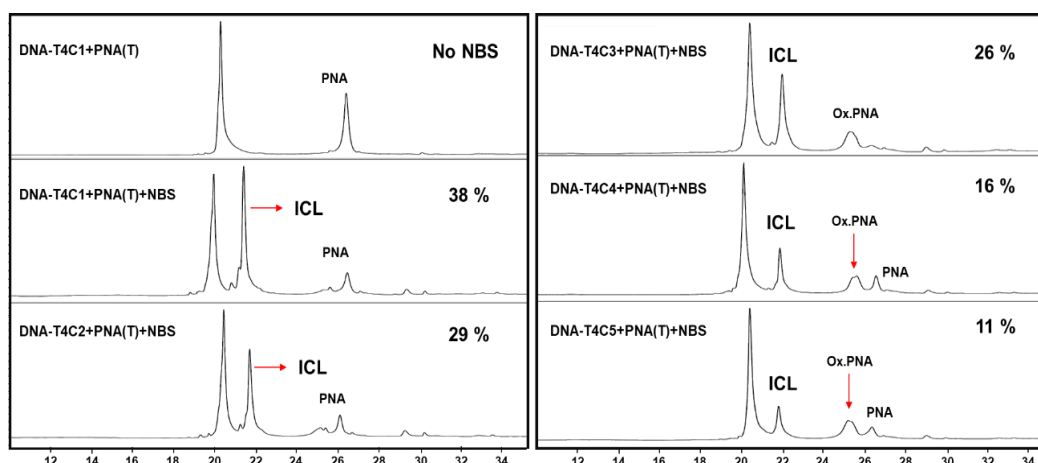
Figure 3.20 RP-HPLC chromatograms of the cross-linking reaction PNA(T) towards various T4C1 overhang target, Conditions: 5%-80% MeCN in 1mM TEAA buffer, [DNA]:[PNA] = 7.5 μM:5 μM.

3.1.5 Cross-linking study of PNA(T) with mismatched DNA targets
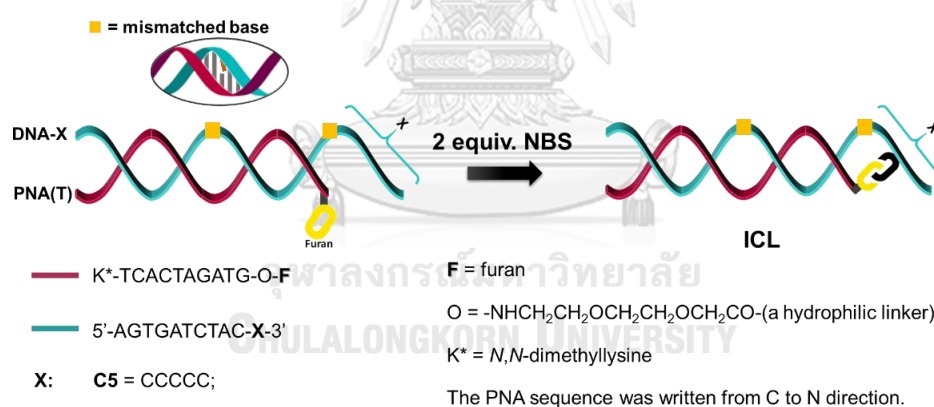
Figure 3.21 A schematic representation of the cross-linking reaction of PNA(T) with single base and double base mismatched DNA

This part of the research focuses on the study of the specificity of the cross-linking reactions between the PNA(T) probe and complementary vs mismatched DNA sequences. The mismatched DNA sequences were designed to be in the middle and at the end of the DNA strand. The mismatched DNA sequences with the mismatch base placed in the middle of the DNA strand include the sequences DNA-sm1 to DNA-sm3. Three additional DNA sequences carrying a 3'-end mismatched base close

to the furan modification (DNA-sm4 to DNA-sm6) were also included. The efficiencies of the cross-linking reactions were evaluated by RP-HPLC, and the melting temperature experiments were carried out under the same conditions as for the previous experiments. **Table 3.4** shows the melting temperature of the PNA(T) with mismatched DNA sequences before and after the NBS activation. Prior to the NBS activation, the melting temperature of the mismatched duplexes were dramatically decreased when compared to the fully complementary DNA (C5) (from 61 °C down to 35 °C). These results are in line with the generally observed high specificity of the acpcPNA probes.[99] When the cross-linking reaction was performed with these single base mismatched DNAs, the cross-linking formation was probably formed as shown in the increasing of the melting temperature after the NBS activation. The cross-linking products of the mismatched sequences increased the duplex stability up to 77 °C or >95 °C in some cases.

Table 3.4 Melting temperature (T_m) of PNA(T) with mismatch sequences

name	sequences (5'-3')	T_m (°C) with PNA(T)	T_m (°C) with PNA(T) after cross-linking	ΔT_m (°C)
DNA-sm1	AGTGAACTACCCCCC	37	82	45
DNA-sm2	AGTGACCTACCCCCC	35	78	42
DNA-sm3	AGTGAGCTACCCCCC	35	79	44
DNA-sm4	AGTGATCTAACCCCC	52	>95	>43
DNA-sm5	AGTGATCTAGCCCCC	48	>95	>47
DNA-sm6	AGTGATCTATCCCCC	49	>95	>45
DNA-dm1	AGTGATCTTGCCCCC	33	82	49
DNA-dm2	AGTGTTGTACCCCCC	33	-	-
DNA-com	AGTGATCTACCCCCC	61	>95	34

- = no melting curve

Conditions: [DNA] = [PNA] = 1 µM, 100 mM NaCl, 10 mM sodium phosphate buffer, pH 7.0, heating rate 1.0 °C /min

Figure 3.22 Melting curves of the duplexes of PNA(T) and single-base mismatched DNA targets (DNA-sm1 to DNA-sm5) before (solid line) and after (dash line) the NBS addition (2 equiv., 37 °C, 1 h). Conditions: [DNA] = [PNA] = 1 µM, 100 mM NaCl, 10 mM sodium phosphate buffer, pH 7.0, heating rate 1.0 °C/min.

The sensitivity of the PNA(T) probe towards mismatch were next studied in a more quantitative fashion by RP-HPLC. After the NBS activation for 20 min, the products from the cross-linking reaction were analyzed and the results are reported in **Figure 3.23**. Consistent with the T_m results, the cross-link products were clearly observable in all single base mismatch sequences, but in the lower efficiency (24% - 45%) when compared to the complementary sequence. The incorporation of one mismatched base was apparently unable to completely suppress the duplex formation (as confirmed by T_m experiments) and thus the cross-linking reaction could still occur, albeit in lower efficiencies when compared to the fully complementary duplex.
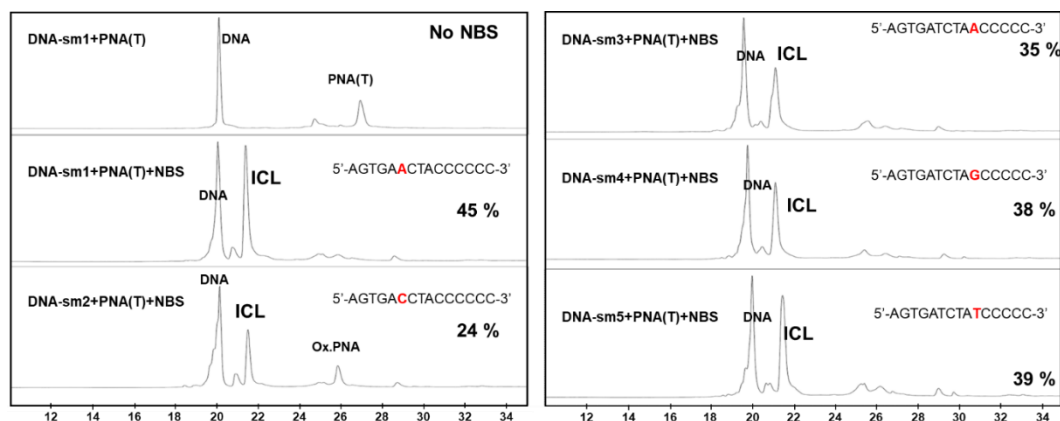
Figure 3.23 RP-HPLC chromatograms of cross-linking reactions between PNA(T) and single base mismatch-sequences (DNA-sm1 to DNA-sm5), following oxidation. Ox. PNA denotes the oxidized PNA probe after the NBS treatment. Conditions: [DNA]:[PNA] = 7.5 µM:5 µM; with 2 equiv. NBS at 37 °C, 20 min, total volume = 100 µL.

Next, we continued the study further with the double base mismatched sequences to investigate the effect on the cross-linking efficiency. The double mismatched bases were replaced at the end and the middle of the DNA strand [(DNA-dm1) and (DNA-dm2), respectively]. Melting temperature studies revealed that when the double mismatch was present at the end of the DNA strand (DNA-dm1), the cross-linking reaction was still possible as shown by the increase of the melting temperature after the NBS addition (from 33 °C to 82 °C). In contrast, when the double mismatch was placed in the middle of the strand, no cross-linking was formed, and the melting temperature could not be detected. Furthermore, the RP-HPLC experiments were performed to analyze the cross-linking products. The results shown in **Figure 3.24** are fully consistent with the T_m results. The cross-linking product of PNA(T) and DNA-dm1 was formed in 18% yield, while no cross-linking between PNA(T) and DNA-dm2 could be observed. From these results, it could be concluded that the PNA-DNA duplex stability influenced the cross-linking efficiency. The stronger the duplex stability, the higher the cross-linking efficiency. The double base mismatch inserted in the middle of the DNA strand destabilized the duplex more than when they were placed at the end of the strand.
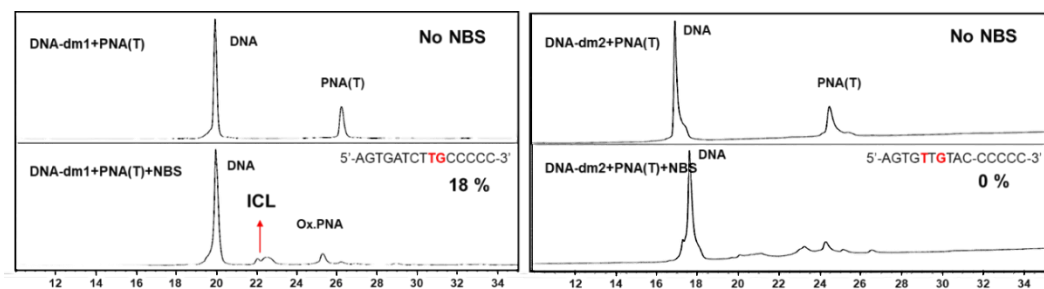
Figure 3.24 RP-HPLC chromatograms of cross-linking reactions between PNA(T) and double base mismatch-sequences (DNA-dm1 to DNA-dm2), following oxidation. Ox. PNA denotes the oxidized PNA probe after the NBS treatment. Conditions: [DNA]:[PNA] = 7.5 µM:5 µM; with 2 equiv. NBS at 37 °C, 20 min, total volume = 100 µL.

The cross-linking experiments performed so far could be summarized as follows: 1) The furan placed at the end of the PNA strand in PNA(T) forms cross-links efficiently only with the nucleobases that do not participate in Watson-Crick base pairing. 2) PNA(T) preferred to form cross-links with nucleobases in the order of C>A>>G. 3) This is the first time that guanine was found to participate in a cross-linking reaction with furan. 4) The cross-linking efficiency of PNA(T) with DNA targets carrying T4C1 overhangs revealed that the further away the C-base position the less the cross-linking efficiency. 5) The cross-linking reaction was sensitive to the base mismatch sequences. 6) The cross-linking reaction could not occur in the case of PNA probe PNA(I) which was modified with furan in the middle of the strand.

3.1.6 Internal cross-linking with base-inserted DNA targets
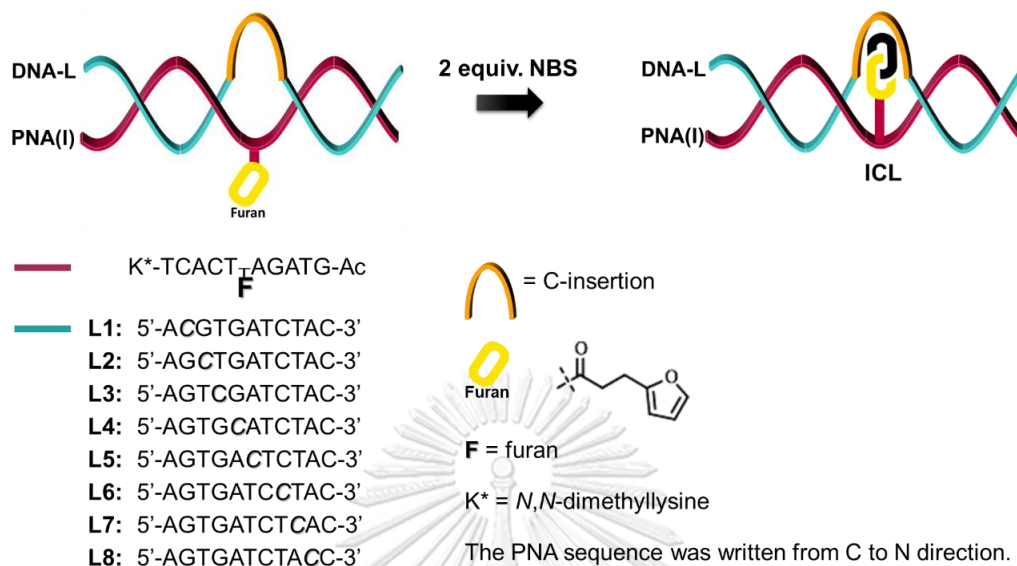
Figure 3.25 A schematic illustration of the cross-linking reaction between PNA(I) and DNA targets with a single C-insertion along the DNA strand (DNA-L1 to DNA-L8).The base-inserted position is indicated by the italicized letter.)

This part of the research presents the cross-linking study of the acpcPNA modified with furan at an internal position of the PNA strand [PNA(I)]. According to previous experiments, cross-linking of this internal probe with its complementary DNA was unsuccessful. The reason might be due to the opposing DNA nucleobases were involved in the Watson-Click base pairing with the PNA and were therefore unavailable for the cross-linking. We proposed that the cross-linking reaction could occur if there is an unpaired nucleobase bearing an exocyclic amino group located near the furan moiety. We designed the DNA sequences with a C-insertion at different position (DNA-L1 to DNA-L8) along the DNA strand to provide a potential site for the cross-link. It was expected that C-insertion could loop out of the duplex resulting in the formation of a bulge. Such bulged structures have been reported in the context of DNA duplexes.[103, 113] RP-HPLC and melting temperature experiments were next performed to investigate the ability of PNA(I) to cross-link with DNA-L1 to DNA-L8 carrying the C-insertion.

Table 3.5 Melting temperature (T_m) of PNA(I) with different position of C-insertion

name	sequence (5'-3')	T_m (°C) with PNA(I)	T_m (°C) with PNA(I) after NBS activation	ΔT_m (°C)
DNA-L1	ACGTGATCTAC	44	39	-5
DNA-L2	AGCTGATCTAC	37	33	-4
DNA-L3	AGTCGATCTAC	30	-ᵃ	-
DNA-L4	AGTGCATCTAC	29	-	-
DNA-L5	AGTGACTCTAC	25	<20	-
DNA-L6	AGTGATCCTAC	33	33, 75	+42
DNA-L7	AGTGATCTCAC	-	-	-
DNA-L8	AGTGATCTACC	47	44	-3

-ᵃ = no melting curves

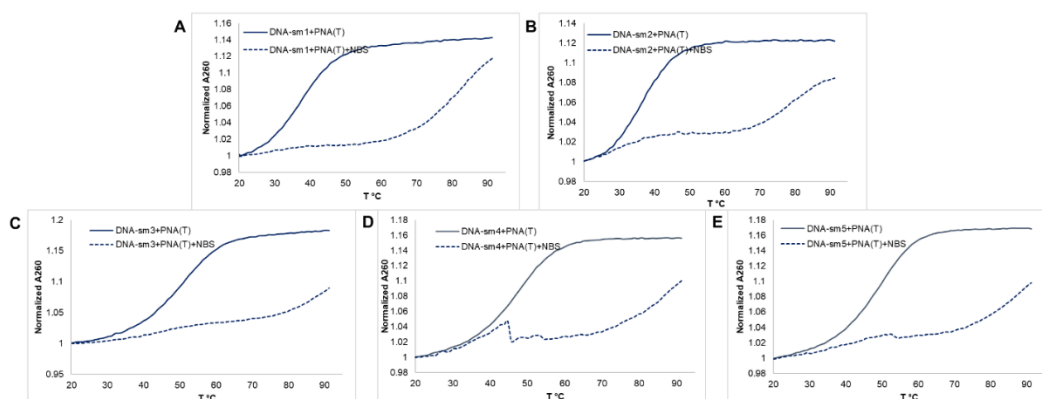Conditions: [DNA] = [PNA] = 1 µM, 100 mM NaCl, 10 mM sodium phosphate buffer, pH 7.0, heating rate 1.0 °C/min

The cross-linking reactions with DNA targets carrying a C-insertion along the DNA sequence was firstly investigated by melting temperature experiments. The T_m at before and after the NBS addition are shown in **Table 3.5.** The presence of a single C-insertion on the DNA strand substantially decreased the PNA-DNA duplex stability when compared to the fully complementary PNA-DNA duplex (61 °C). The T_m was as low as 25°C when the C-insertion was present in the middle of the strand as shown in DNA-L5. The reduction in the T_m was even larger than in the case of mismatched sequences. After the addition of NBS, the change in the melting profile which is indicative of cross-linking formation was observed only in the case of DNA-L6 as shown by an increase of the T_m up to 75 °C (ΔT_m = +42). When the location of the C-insertion was shifted only one base away, no cross-linked product was observed. The results indicated that the cross-linking reaction of the PNA(I) was possible if there is an unpaired nucleobase target close to the furan moiety, thus confirming the original hypothesis.
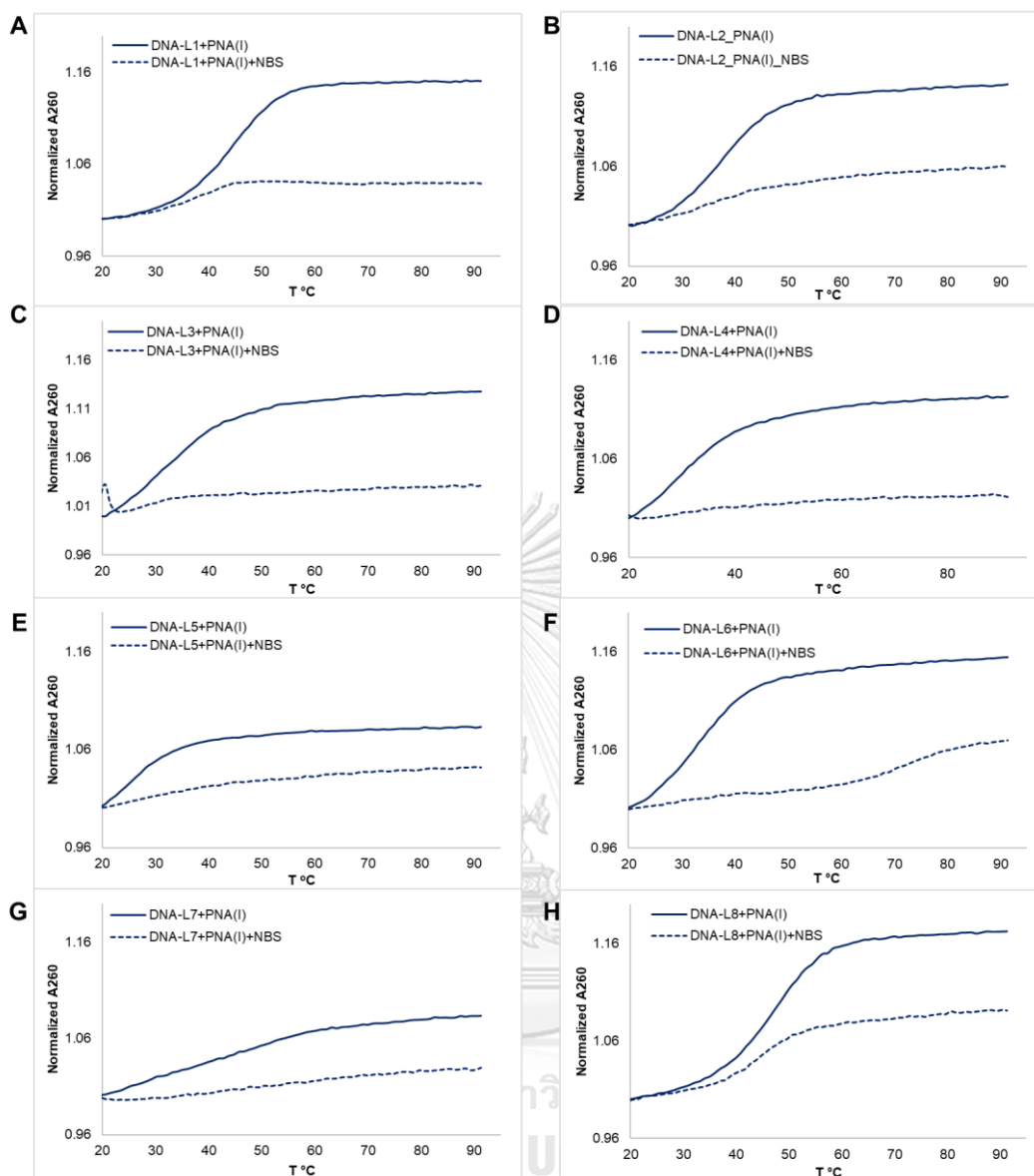
Figure 3.26 Melting curves of the duplexes of PNA(I) with C-insertion DNA targets (DNAL1-DNAL8) before (solid line) and after (dash line) the NBS addition (2 equiv., 37 °C, 1 h).Conditions: [DNA] = [PNA] = 1 μM, 100 mM NaCl, 10 mM sodium phosphate buffer, pH 7.0, heating rate 1.0 °C/min. The cross-linking was induced by addition of NBS (2 equiv.).

In addition, the study of C-insertion was studied further by RP-HPLC. The results as shown in **Figure 3.27** indicated the cross-linking formation only in the case of PNA(I) and DNA-L6 in 13% yield following the NBS activation. Only the oxidized form of the PNA (Ox. PNA) was observed with other DNA sequences. These results show that the cross-linking reaction can occur only when the C-insertion was located

close to the furan moiety as in DNA-L6. Overall, the results support our hypothesis that the important factors for the cross-linking formation is the presence of an unpaired nucleobase with a free exocyclic amino group close to the furan part.
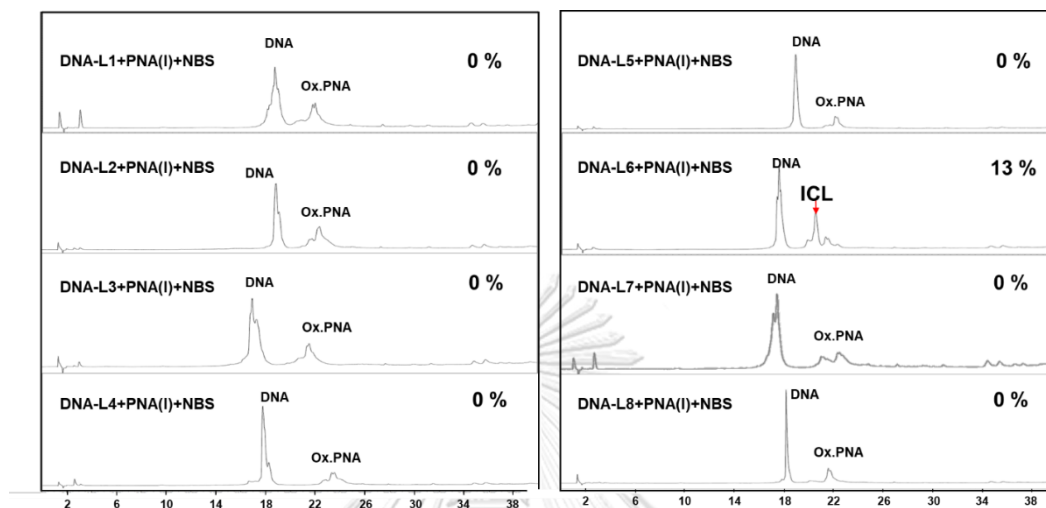
Figure 3.27 RP-HPLC chromatograms of cross-linking reactions between PNA(T) with various C-insertion targets (DNAL1-DNAL8), following oxidation. Ox. PNA denotes the oxidized PNA probe after the NBS treatment. Conditions: [DNA]:[PNA] = 7.5 µM:5 µM; with 2 equiv. NBS at 37 °C, 20 min, total volume = 100 µL.

Moreover, we investigated the specificity of the cross-linking reaction of DNA-L6 further with another mismatched sequence. A single-base mismatched sequence (**Table 3.1**) of DNA-L6 (DNA-L6sm) was introduced as the target to investigate the specificity of the PNA(I) probe. The cross-linking reaction was performed as usual and the products were analyzed by RP-HPLC as shown in **Figure 3.28.** As expected, no cross-linking product was observed with the mismatched DNA target, thus confirming the specificity of the cross-linking by this PNA probe.
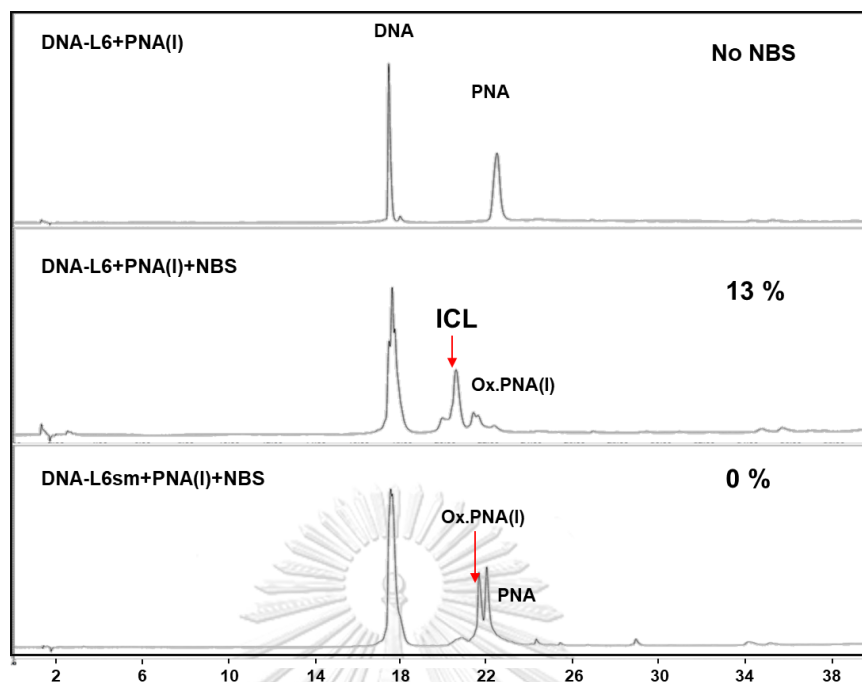
Figure 3.28 RP-HPLC chromatograms of cross-linking reactions between PNA(T) with the DNA-L6 and its mismatched sequences (DNA-L6sm), following oxidation. Ox. PNA denotes the oxidized PNA probe after the NBS treatment. Conditions: [DNA]:[PNA] = 7.5 µM:5 µM; with 2 equiv. NBS at 37 °C, 20 min, total volume = 100 µL.

In the last experiment, we studied further with the cross-linking reaction of PNA(I) with different base insertions including one C-insertion (DNA-L5, DNA-L6), two C-insertions (DNA-L2C) and A-insertion (DNA-LA). The reactions were analyzed by denaturing PAGE as shown in **Figure 3.29**. The results revealed that when the C or A bases in DNA-L5 and DNA-LA, respectively, was shifted away from DNA-L6 only one position, no cross-linked product was formed as shown in lane 3 and lane 5, respectively. This result indicated that when the C or A base was far away from the furan moiety just only one base away, the cross-linking reaction could not occur. Thus, the presence of free exocyclic nucleobase that was close to the furan part is necessary for the cross-linking reaction. However, the cross-linked products were observed in lane 7 and lane 9 which were the DNA targets carrying one C-insertion (DNA-L6) and two C-insertions (DNA-L2C) close to the furan modification site. These results support the previous conclusion that PNA(I) could participate in the cross-

linking reaction if there is one or more free C-base close to the furan moiety. In the future, the structure of the C-insertion duplexes should be studied by techniques such as X-ray crystal structure to obtain further insights into the structure and orientation of the furan in the PNA-DNA duplex.
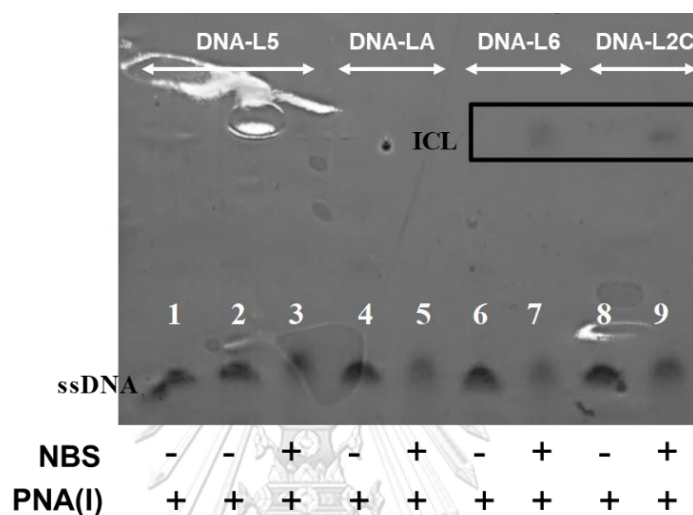
Figure 3.29 Denaturing PAGE results of internally furan-modified PNA [PNA(I)] with different base insertions; C-insertion in DNA-L5, DNA-L6; two-C-insertion in DNA-L2C; A-insertion in DNA-LA in presence (+) and absence (−) of NBS. Conditions: [DNA] = [PNA] = 20 nmol, 100 mM NaCl, 10 mM sodium phosphate buffer, pH 7.0; 2 equiv. NBS, 37 °C, 20 min; total volume = 10 μL.

3.1.7 The fate of PNA probes following NBS activation

As noticed earlier, there was a significant drop in the hyperchromicity in the melting curves observed in the case of the hybrids between PNA probes and mismatched (**Figure 3.22**) or C-insertion (**Figure 3.26**) DNA targets following the NBS activation. Moreover, it was found that the PNA peak area in the RP-HPLC chromatogram was decreased after NBS addition in the case of double-mismatched (**Figure 3.24**) and C-insertion (**Figure 3.27**) DNA targets even though they did not form cross-link products. We hypothesized that the furan PNA probe could decompose upon NBS activation. The MALDI-TOF MS trace of the PNA(T) probe after NBS addition was investigated under three different conditions consisting of PNA(T) alone (**Figure 3.30**), PNA(T) with DNA-T5 (**Figure 3.31**), and PNA(T) with DNA-dm2 (**Figure 3.32**).
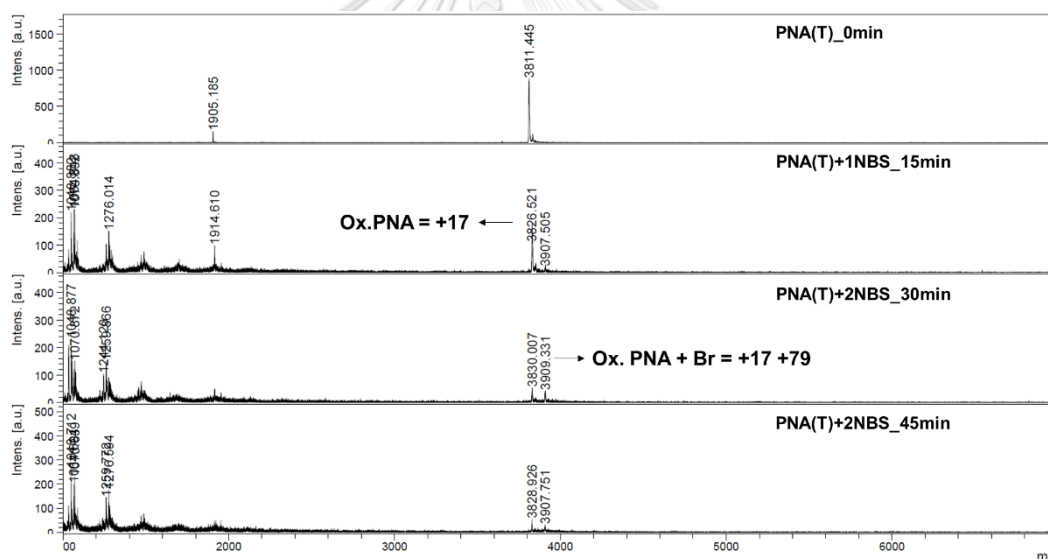
Figure 3.30 MALDI-TOF mass spectra of the PNA(T) before and after treatment with NBS (1 equiv. was added, wait for 15 min, then add another 1 equiv.).The time was taken from the addition of the first equiv. of NBS.

The results revealed that after addition of 1 equiv. of NBS for 15 min, the PNA(T) was completely oxidized to generate the reactive aldehyde species (Ox. PNA) which had a +17 Da mass difference. In the case of PNA alone, the degradation was almost complete in 45 min. However, the MALDI-TOF MS analysis result of the PNA(T) that was activated in the presence of DNA-T5, which could form a stable duplex with the PNA probe but could not form the cross-linked product, showed

very small amount of the degradation even after 2 h (**Figure 3.30**). Furthermore, the result of the PNA(T) with DNA-dm2 (double mismatched sequences) revealed that after 2 equiv. of NBS was added the degradation took place and was complete during the same 2 h period. The brominated intermediate (B in **Figure 3.33**) was also detected under this condition. According to these results, it was concluded that the degradation of the activated furan PNA probe occurred more readily when the activated PNA probe was free, i.e. did not form a base pair with the opposite DNA strand, regardless of the ability to form cross-links. As a result, the degradation process is proposed to be the self-destruction of the activated PNA probe through the reaction between the activated furan and the PNA bases, which could be prevented by the base-pairing. A similar behavior was observed in the context of furan-modified RNA probes by Carrette and coworkers in 2014.[61] The less stable the duplex, the higher the degradation rate. These results could be used to explain the drop in hyperchromicity in the melting temperature and the decrease in the peak area of PNA in the RP-HPLC chromatogram in the case of targets that could not form stable duplexes with the PNA probes.
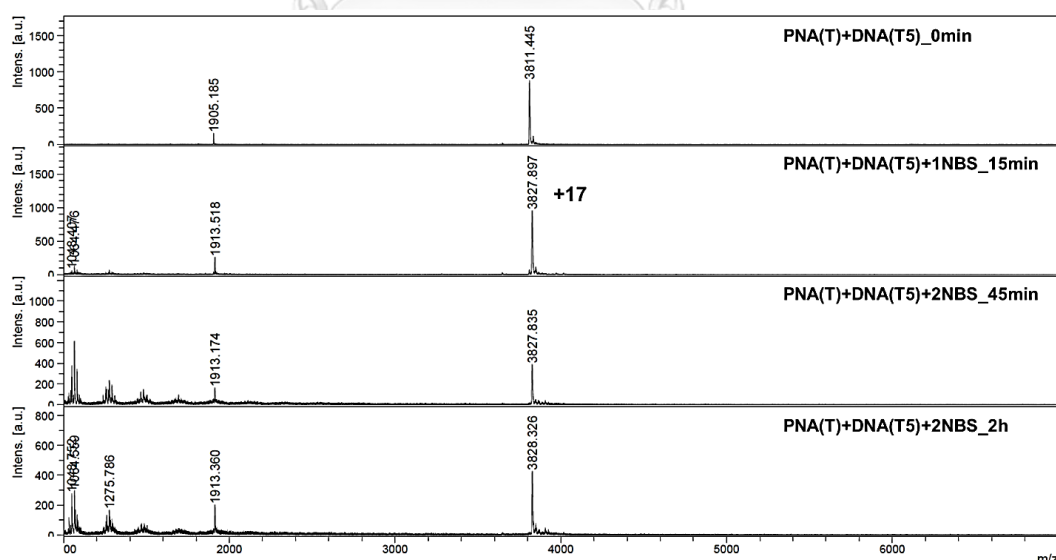
Figure 3.31 MALDI-TOF mass spectra of a mixture of PNA(T) and DNA-T5 before and after treatment with NBS (1 equiv. was added, wait for 15 min, then add another 1 equiv.). The time was taken from the addition of the first equiv. of NBS.
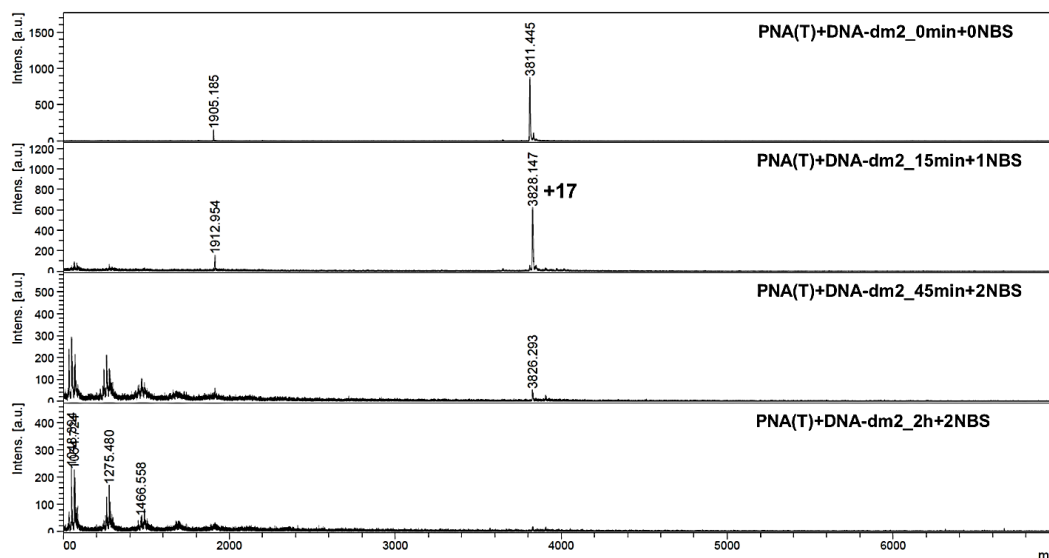
Figure 3.32 MALDI-TOF mass spectra of a mixture of PNA(T) and DNA-dm2 before and after treatment with NBS (1 equiv. was added, wait for 15 min, then add another 1 equiv.). The time was taken from the addition of the first equiv. of NBS.
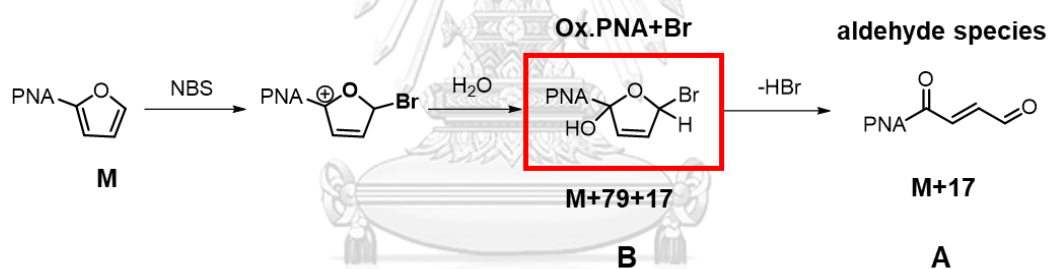
Figure 3.33 Mechanism of furan oxidation reaction upon NBS oxidation to generated reactive aldehyde species (A) or brominated side products (B).

3.1.8 Conclusion

In conclusion, the work carried out in this section involved the cross-linking studies of furan-modified acpcPNA at the end (terminal) and the middle (internal) of the strand. The synthesis of furan-modified PNA was achieved by a convenient and efficient post-synthetic method, which could be adapted for the preparation of both internally and termimally furan-modified PNA probes. The result revealed that the terminal PNA probe could undergo cross-linking reactions with the C-base in the overhang region of the complementary DNA (DNA-C5) with a high yield (71%). Under the same condition, the internally labeled PNA probe failed to cross-link with complementary DNA (DNA-com). This could be explained as a result of the formation of the Watson-Crick base pairing of the DNA bases, making them unavailable to react with the activated furan species. To confirm the hypothesis and to improve the cross-linking efficiency of the internal probe, C-insertion was introduced to the DNA targets. The result showed that the cross-linked product could indeed form with the DNA-L6 that contain an unpaired C-base located close to the furan moiety. These results indicate that the essential factors for the successful cross-linking process was the ability to form a stable duplex between the PNA probe and the DNA target, as well as the availability of the nucleobase that was located nearby the activated furan moiety. In addition, it was found the activated PNA probe in the free form (not forming a duplex with its DNA target, regardless of the cross-linking ability) rapidly degraded which might be beneficial to reduce the side reactions associated with non-specific "off-target" cross-linking reactions.

SECTION II

3.2 FURAN-MODIFIED acpcPNA VS FURAN-MODIFIED aegPNA

This section of the thesis focuses on the comparison of the cross-linking efficiency between aegPNA and acpcPNA. The cross-linking studies are divided into two main parts. The first part focuses on the cross-linking studies of the PNA probes bearing the furan modification at the end of the strand (terminal modification). The second part involves the cross-linking studies of the PNA probes carrying the furan modification in the middle of the strand (internal modification). Three different furan building blocks were incorporated into the PNA strand at the terminal study which compose of the f (furan attached to a flexible O-linker), F (abasic furan building block) and T (thymine+furan building block) as shown in **Figure 3.34A**. In the case of internal modification study, only two building blocks (F and T) were included in the study (**Figure 3.34B**). *N*-bromosuccinimide (NBS) was used as an activator in all studies.

The cross-linking studies are performed by denaturing PAGE, RP-HPLC and MALDI-TOF MS. All experiments in this section, with the exception of the synthesis of acpcPNA probes, were performed at Organic and Biomimetic Chemistry Research Group, Department of Organic and Macromolecular Chemistry, Ghent University, Belgium under supervision of Prof. Annemieke Madder. The RP-HPLC conditions used in this section follow the previous works from Madder's group.[100, 101] The gels are strained by SYBR gold dye before taking the photograph under UV light at 365 nm.

A

DNA: 5'-AGATCATGCCC-**X**-3'
PNA: H$_2$N-rrr-TCTAGTACGGG-**M**

M: Rf   RF   RT

X: **A1 =** A   **A5 =** AAAAA
         C1 = C   **C5 =** CCCCC
         G1 = G   **G5 =** GGGGG
         T1 = T   **T5 =** TTTTT

R: regPNA
P: acpcPNA
The PNA sequence was written from C to N direction.

M: Pf   PF   PT

B

DNA: 5'-GCAGATC-**X Y**-GCCCGGC-3'
PNA: H$_2$N-rrr-TCTAG-**M P**-CGGG-Ac

Regular PNA

X: the facing nucleobase of **M** building block
Y: the previous DNA nucleobase
P: the nucleobase before **M** building block

acpc PNA

R: regPNA
P: acpcPNA
The PNA sequence was written from C to N direction.

F   T

F   T

Figure 3.34 Furan building blocks and PNA/ DNA sequences in terminal (A) and internal (B) modified PNA, R: aegPNA, P: acpcPNA

3.2.1 Synthesis of furan-modified acpcPNA

The synthesis of acpcPNAs studied in section II were similar to the synthesis in section I, with an additional special abasic F-monomer (Fmoc-D-proline-OPfp) and arginine as a solubility enhancer instead of lysine. Normally, lysine was inserted as the C-terminal residue on the resin, but in these sequences three arginine residues were used instead to facilitate the post-synthetic modification because the arginine side chain would not react with the activated furan building block under the post-synthetic modification condition, hence the protection step was deemed unnecessary.

The synthesis was started with the loading of Fmoc-Arg(Pbf)-OH as the first residue on TentaGel S RAM Resin followed by successive incorporation of two additional arginine residues. The solid phase peptide synthesis cycle (Fmoc-deprotection, coupling, capping) was performed as usual. Concerning the synthesis of the PNA carrying the furan modification at the terminal position, the PNA precursor (Fmoc-GGGCATGATCT-rrr) was synthesized on the solid support as follows. In the case of Pf-series, a hydrophilic linker (Fmoc-AEEA-OH) was added as the last residue. In the PF-series, the precursor was coupled with the abasic monomer (Fmoc-D-proline-OPfp) followed by an APC spacer for the furan modification. In the PT-series, the synthesis was similar to the PF-series, but the T-monomer (Fmoc-T-OPfp) was used instead of the abasic monomer. In the latter two cases, the N-terminal NH_2 group of the APC spacer was capped by acetylation whilst still on the solid support. Next the terminal Fmoc group, the Tfa protecting group of the APC residue, and the nucleobase side chain protecting groups were removed by ammonia treatment, and the PNA were cleaved from the solid support as usual. However, the TFA cleavage step required at least 2 h to ensure complete removal of the Pbf protecting group of the arginine residues. The incompletely removed Pbf protecting group could be identified by MALDI-TOF MS, which showed the mass increase of +252.3 Da from the expected product. The PNA precursors were post-synthetically modified after cleavage from the solid support with Pfp-activated 3-(2-furyl) propionic acid as described in section I.

Internal furan-modified acpcPNA sequences were synthesized in the same manner with PNA(I) as described in section I. Two acpcPNA monomers (Fmoc-D-proline-OPfp for the F-series, Fmoc-T-OPfp for T-series) were inserted separately in the middle of the PNA strand followed by incorporation of an APC spacer to provide a handle for the furan modification. After that, four nucleobase monomers were separately incorporated to obtain four sequences of each monomer (4 types of nucleobase for F and another 4 types of nucleobase for T). The final sequences are: Ac-GGGC-PM-GATC-rrr-NH$_2$ where P is one of the four nucleobases (A, T, C, G), and M is the F or T monomer (see the full sequences in **Table 3.6**). After obtaining the desired sequences, the *N*-terminal Fmoc group was removed, and the free amino terminus was capped by acetylation. Next, the deprotection of the nucleobase and APC residue was performed followed by the cleavage step as usual. The post-synthetic modification was carried out on the cleaved PNA in solution phase according to the procedure described in section I. However, the post-synthetic modification time was extended to 1 - 2 h to ensure complete reactions. The poor solubility of the PNAs in DMF might be the reason of this long reaction time.

3.2.2 Synthesis of furan-modified aegPNA

The objective of this part was to compare the cross-linking efficiency of acpcPNA and aegPNA side-by-side. Thus, the same PNA sequences were required. Three furan-modified aegPNA sequences (Rf, RF, RT) were synthesized for the study in this section (see the sequence in **Table 3.6**). The synthesis was carried out by Fmoc-SPPS method on ChemMatrix Rink Amide resin. Three arginine residues were firstly loaded onto the resin to improve the PNA solubility as described earlier in the case of acpcPNA. The synthesis cycle, consisting of Fmoc-deprotection, coupling, and capping, was repeated until the desired PNA sequences were obtained. At the *N*-terminus, three different monomers were separately added as the last residue. All furan-modified aegPNA monomers were synthesized by Dr. Alex Manicardi (Ghent University). For the Rf-series, a hydrophilic linker (Fmoc-AEEA-OH) was inserted prior to the maleimide-protected furfuryl propionic acid residue (FurDA-OH). In the RF-

series, the PNA probe was modified by the protected abasic furan monomer (Fmoc-FurDA-OH) and for the RT-series, a thymine+furan monomer (Fmoc-Fur-T-OH) was added as the last residue. Note that in the aegPNA series, all furan modifications were performed on the solid support employing the maleimide protected furan monomers. Interestingly, the thymine+furan monomer had been successfully incorporated into the aegPNA sequence without the furan protection. After that, the PNAs were cleaved from the solid support by a treatment of TFA/m-cresol 9:1 for the Rf and RF sequences. Due to the presence of free furan in the RT PNA sequence, a TFA/m-cresol/thioanisole 8:1:1 solution was used to cleave the PNA to prevent the alkylation side reaction of the furan moiety caused by the cations generated from the Bhoc-protecting group. The crude PNAs were purified by RP-HPLC and characterized by MALDI-TOF MS.

Table 3.6 PNA sequences studied in section II

sequence (N to C direction)	code
f-O-GGGCATGATCT-rrr-NH$_2$	Rf
Ac-F*-GGGCATGATCT-rrr-NH$_2$	RF
Ac-T*-GGGCATGATCT-rrr-NH$_2$	RT
f-O-GGGCATGATCT-rrr-NH$_2$	Pf
Ac-F*-GGGCATGATCT-rrr-NH$_2$	PF
Ac-T*-GGGCATGATCT-rrr-NH$_2$	PT
Ac-GGGC**AF**GATCT-rrr-NH$_2$	P-AF
Ac-GGGC**CF**GATCT-rrr-NH$_2$	P-CF
Ac-GGGC**GF**GATCT-rrr-NH$_2$	P-GF
Ac-GGGC**TF**GATCT-rrr-NH$_2$	P-TF
Ac-GGGC**AT**GATCT-rrr-NH$_2$	P-AT
Ac-GGGC**CT**GATCT-rrr-NH$_2$	P-CT
Ac-GGGC**GT**GATCT-rrr-NH$_2$	P-GT
Ac-GGGC**TT**GATCT-rrr-NH$_2$	P-TT

Table 3.7 Target DNA list used in section II

DNA Sequence (5'- 3')	code
AGATCATGCCCA	A1
AGATCATGCCCC	C1
AGATCATGCCCG	G1
AGATCATGCCCT	T1
AGATCATGCCCAAAAA	A5
AGATCATGCCCCCCCC	C5
AGATCATGCCCGGGGG	G5
AGATCATGCCCTTTTT	T5
<u>TACG</u>CATGCCCA	EA1
<u>TACG</u>CATGCCCC	EC1
<u>TACG</u>CATGCCCG	EG1
<u>TACG</u>CATGCCCT	ET1
<u>TACG</u>CATGCCCAAAAA	EA5
<u>TACG</u>CATGCCCCCCCC	EC5
<u>TACG</u>CATGCCCGGGGG	EG5
<u>TACG</u>CATGCCCTTTTT	ET5
<u>TACG</u>CATGCCCCTTTT	ET4C1
<u>TACG</u>CATGCCCTCTTT	ET4C2
<u>TACG</u>CATGCCCTTCTT	ET4C3
<u>TACG</u>CATGCCCTTTCT	ET4C4
<u>TACG</u>CATGCCCTTTTC	ET4C5
GCAGATCAAGCCCGG C	AA
GCAGATCACGCCCGGC	AC
GCAGATCAGGCCCGGC	AG
GCAGATCATGCCCGGC	AT
GCAGATCCAGCCCGGC	CA
GCAGATCCCGCCCGGC	CC
GCAGATCCGGCCCGGC	CG
GCAGATCCTGCCCGGC	CT

DNA Sequence (5'- 3')	code
GCAGATCGAGCCCGGC	GA
GCAGATCGCGCCCGGC	GC
GCAGATCGGGCCCGGC	GG
GCAGATCGTGCCCGGC	GT
GCAGATCTAGCCCGGC	TA
GCAGATCTCGCCCGGC	TC
GCAGATCTGGCCCGGC	TG
GCAGATCTTGCCCGGC	TT

3.2.3 Cross-linking studies of terminally furan-modified PNA with fully complementary DNA targets

This part focuses on the cross-linking study of furan-modified PNAs at the terminal position of the strand. The furan-modified PNA sequences employed in this study are shown in **Table 3.6**. Three different furan building blocks (f, F, T) were incorporated into the aegPNA (R) and the acpcPNA (P) probes. The sequences of the complementary DNA targets are presented in **Table 3.7**. These DNA sequences were divided into two main groups. The first group was the fully complementary sequences carrying an unpaired base (one-base overhang) at the 3'-end (DNA-A1, DNA-C1, DNA-G1, and DNA-T1). The second group was the fully complementary sequences with five unpaired bases (five-base overhang) at the 3'-end of the DNA strand (DNA-A5, DNA-C5, DNA-G5, and DNA-T5).

3.2.3.1 Denaturing PAGE
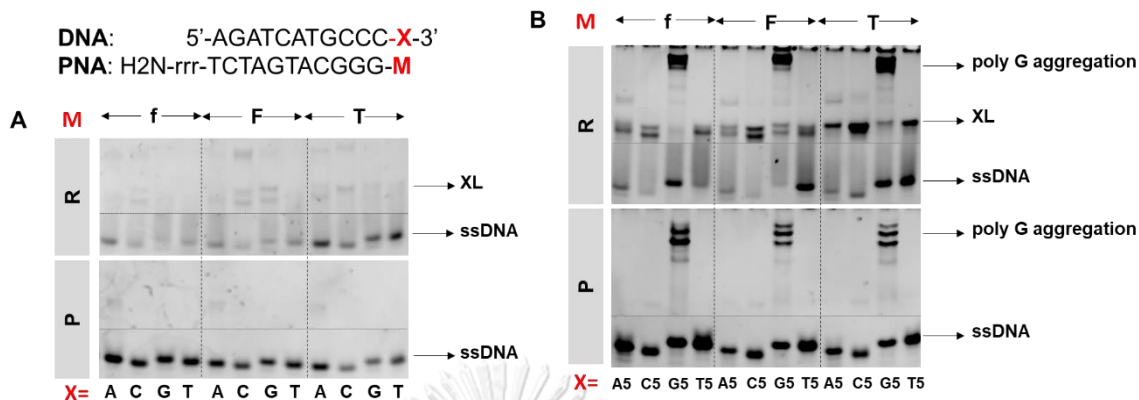
Figure 3.35 Denaturing PAGE results of the terminally modified aegPNA (R) and acpcPNA (P) towards fully complementary DNA targets with one base at the overhang (A) and five bases at the overhang targets (B) in presence of NBS (4 equiv.). Conditions: [DNA] = [PNA] = 5 µM in the total volume = 100 µL, 4 equiv. NBS (1 equiv. every 15 min), for 2h. The experiment was performed in 20% polyacrylamide gel containing 7 M urea at 230 V for 1 h and 40 min. XL: cross-linked product, ssDNA: single stranded DNA

In the preliminary cross-linking studies, the two PNAs with three different furan modifications were tested with various target DNAs. The reactions were performed at 5 µM concentration of PNA probe and DNA concentration in 10 mM phosphate buffer (pH 7.4) containing 100 mM NaCl. The reactions were incubated at 25 °C before adding 4 equiv. of NBS (1 equiv. for every 15 min) and shaken in Eppendorf thermomixer for 1 h. The cross-linked products were analyzed by denaturing PAGE. The results of both types of PNAs with one-base overhang DNA targets (**Figure 3.35A**) and five-base overhang DNA targets (**Figure 3.35B**) are shown.

As shown in **Figure 3.35A**, the results of the aegPNA (R) showed very faint bands at the expected positions of both the cross-linked products and the single strand DNAs (ssDNA) bands were still clearly visible. For acpcPNA, no cross-linking band was observed with all types of furan modification and the ssDNA bands were quite prominent. Similar results were observed with the DNA targets carrying five-

base overhangs (**Figure 3.35B)**. Some of the cross-linking bands were evidenced in the aegPNA but not acpcPNA. Moreover, while the single-stranded DNA bands (the fast-moving bands) seemed to disappear upon NBS addition, the intensity of the cross-linked product bands were lower than expected. It was hypothesized that the SYBR-Gold dye may not stain small DNAs and their hybrids or cross-linked products with aegPNA/acpcPNA well enough to allow the detection by PAGE.[114, 115] Thus, the cross-linking reactions were analyzed by RP-HPLC to asset the formation of the cross-linked products instead.

3.2.3.2 RP-HPLC

In this part, the RP-HPLC analyses were performed on the reaction products from the cross-linking experiments before and after the NBS addition. The study was performed on the same set of terminally furan-modified aegPNA with three different furan building blocks (Rf, RF, and RT) and DNA target with one-base overhang (A) and five-base overhang (B). The reactions were performed under the same conditions as for the PAGE experiments, with the aliquots being sampled for the analysis by RP-HPLC at 2 h following the activation. The results are shown in **Figures 3.36** - **3.38**. Some of the cross-linking products were collected and characterized by MALDI-TOF MS.

The cross-linking percentage was calculated from the DNA conversion. This calculation is different from the cross-linking yield in the section I because in several cases the peak area of the PNA could not be measured accurately due to the broadening of the aegPNA peaks. In some experiments, it was not possible to observe the PNA peak and the cross-linked product's peak behaved differently among different PNAs. In addition, the concentration of PNA and DNA used were equal (1:1) which does not allow the same calculation as in the section I. Thus, the percentage of DNA conversion was calculated according to **Equation 2** where A_{DNA-A} is the peak area of DNA after NBS addition for 2 h, A_{DNA-B} is the peak area of the starting DNA (without NBS). The calculated cross-linking yield are presented together with the HPLC chromatrogram for each condition.

Equation 2; % conversion of DNA$=\dfrac{A_{DNA\text{-}A}}{A_{DNA\text{-}B}} \times 100$

a) terminally furan-modified aegPNA

Figure 3.36 RP-HPLC showing the cross-linking experiments of the terminal furan-modified aegPNA with f-building block (Rf) with one base at the overhang (A) and five bases at the overhang targets (B) before (blue line) and after NBS addition (red line); Ox. PNA: oxidized PNA probe; XL: cross-linked product; -: no cross-linked product

Figure 3.37 RP-HPLC showing the cross-linking experiments of the terminal furan-modified aegPNA with F-building block (RF) with one base at the overhang (A) and five bases at the overhang targets (B) before (blue line) and after NBS addition (red line); Ox. PNA: oxidized PNA probe; XL: cross-linked product; -: no cross-linked product
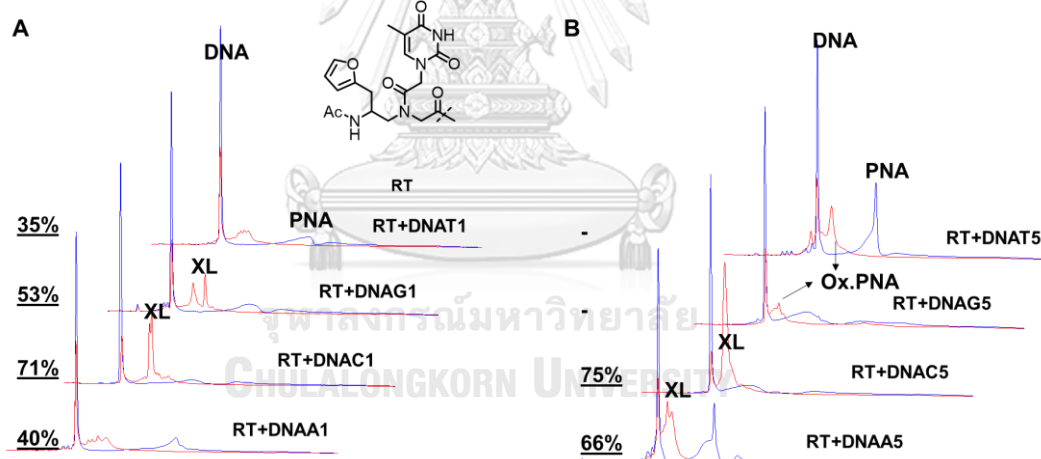
Figure 3.38 RP-HPLC showing the cross-linking experiments of the terminal furan-modified aegPNA with T-building block (RT) with one base at the overhang (A) and five bases at the overhang targets (B) before (blue line) and after NBS addition (red line); Ox. PNA: oxidized PNA probe; XL: cross-linked product; -: no cross-linked product

The RP-HPLC results revealed that after NBS addition the cross-linked products as well as the oxidized specie (Ox. PNA) of the furan-modified PNA probes were generated. In the case of DNA carrying one-base overhang, the cross-linking

experiments can be summarized as follows. Firstly, the aegPNA probe carrying the f-building block (Rf) with the longest linker can cross-link with DNA-C1 and DNA-T1. In the latter case, the f-building block cross-linked with the C-base adjacent to the T-base, and not the T-base itself, according to MALDI-TOF MS results (vide infra). Secondly, the F-building block (RF) which is an abasic building block, was found to cross-link with all one-base overhang DNA targets including A, C, G and T (via the adjacent C-base). Lastly, the T-building block (RT), a thymine-PNA monomer with a furan modification, was found to cross-link with the one-base overhang DNA targets carrying A, C and G. The MALDI-TOF MS analysis of the F-PNA linked to DNA-C1 (RF+C1) presented in **Figure 3.39** was performed to confirm the cross-linking. The MALDI-TOF MS data revealed that the cross-linked product (RF+C1) was observed at 7369.02 Da. The mass was larger than calculated data which is 7337.15 Da by +32 Da, which might be attributed to the potassium ion adduct.

Figure 3.39 MALDI-TOF mass spectra of a cross-linked product (RF+C1) of aegPNA with F-building block and DNA-C1 after treatment with NBS (4 equiv., 2h)

In the case of DNA targets with five-base overhangs, the cross-linking results follow the same trend as the one-base overhangs, but the cross-linked products peaks were more complex and difficult to interpret, possibly due to the formation of multiple cross-linked species. The results can be concluded as follows. Firstly, the aegPNA carrying the terminal f-building block was found to cross-link with only DNA-C5 and DNA-T5 (via the C adjacent to the T, see section **3.2.5**). Secondly, the aegPNA with the terminal F-building block can form cross-links with DNA-A5, DNA-C5 and DNA-G5. Moreover, the complex between the oxidized PNA (Ox. PNA) and the DNA (as shown by MALDI-TOF MS in section **3.2.5**) was also observed as broad peak (**Figure 3.37B** and **Figure 3.38**). Lastly, the aegPNA with the terminal T-building block was found to cross-link with DNA-A5 and DNA-C5 and the complexes between the Ox. PNA and DNA (as shown by MALDI-TOF MS in section **3.2.5**) were also observed in the case of DNA-G5 and DNA-T5 (**Figure 3.38B**).
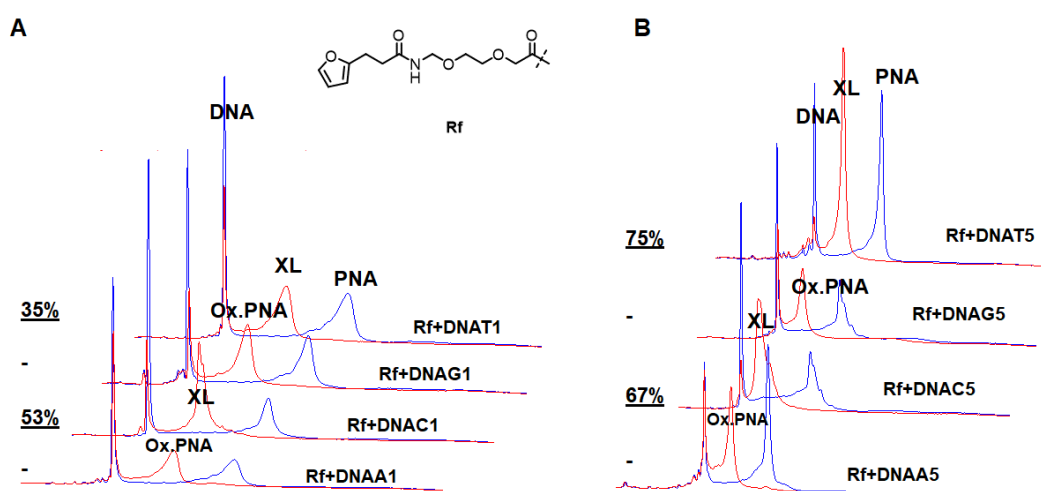
b) terminally furan-modified acpcPNA
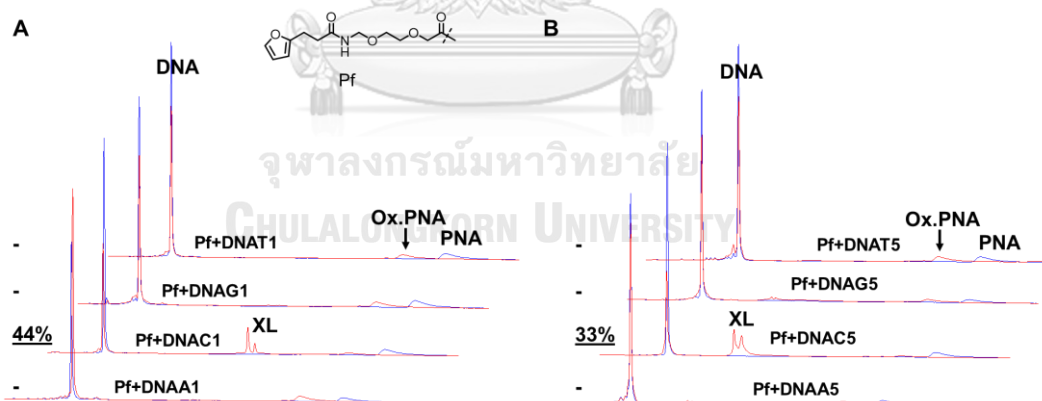
Figure 3.40 RP-HPLC showing the cross-linking experiments of the terminal furan-modified acpcPNA with f-building block (Pf) with one base at the overhang (A) and five bases at the overhang targets (B) before (blue line) and after NBS addition (red line); Ox. PNA: oxidized PNA probe; XL: cross-linked product; -: no cross-linked product
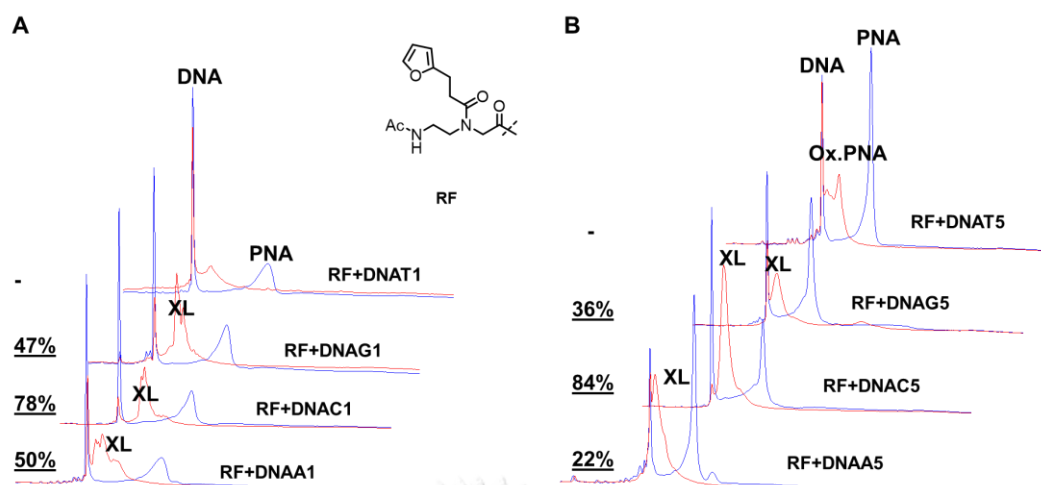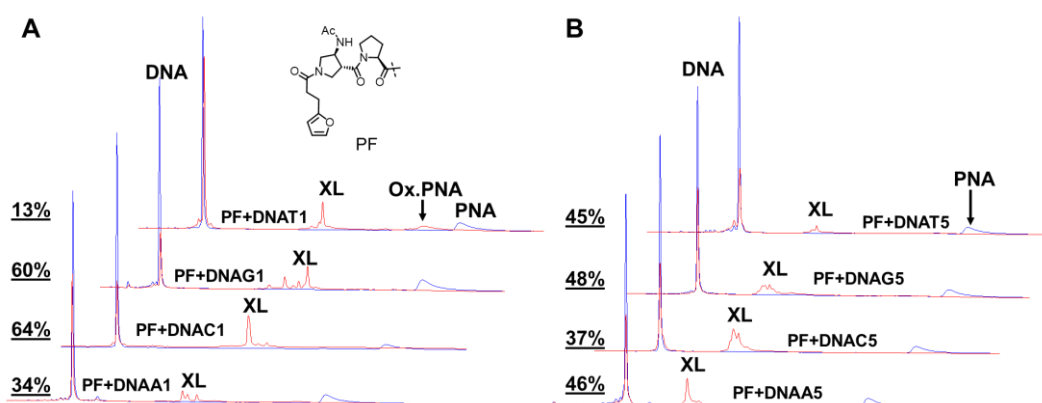
Figure 3.41 RP-HPLC showing the cross-linking experiments of the terminal furan-modified acpcPNA with F-building block (PF) with one base at the overhang (A) and five bases at the overhang targets (B) before (blue line) and after NBS addition (red line); Ox. PNA: oxidized PNA probe; XL: cross-linked product; -: no cross-linked product
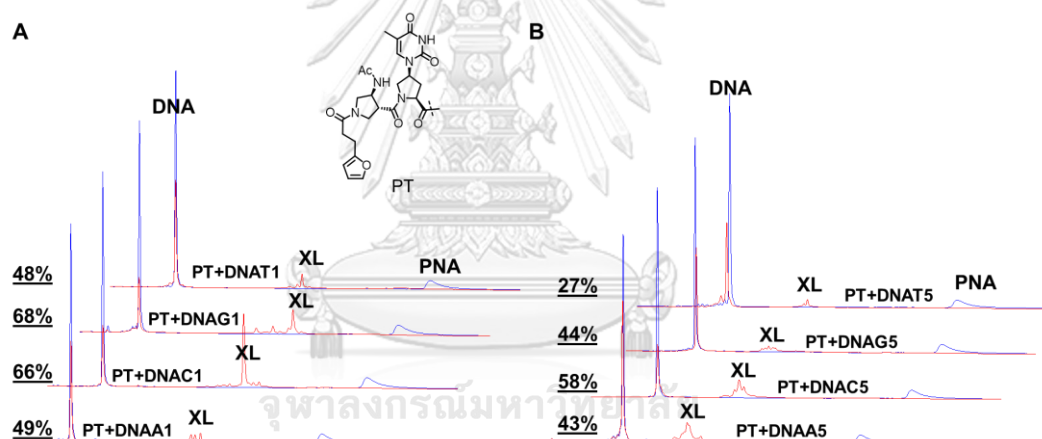
Figure 3.42 RP-HPLC showing the cross-linking experiments of the terminal furan-modified acpcPNA with T-building block (PT) with one base at the overhang (A) and five bases at the overhang targets (B) before (blue line) and after NBS addition (red line); Ox. PNA: oxidized PNA probe; XL: cross-linked product; -: no cross-linked product

Similar cross-linking and RP-HPLC experiments were performed with terminally furan-modified acpcPNAs Pf, PF and PT bearing the f-, F- and T-monomers, respectively. According to the HPLC results (**Figures 3.40** - **3.42**), the acpcPNA peaks were observed as small bumps or very broad peaks even though when no NBS was

present. This behavior was different from aegPNA and might be explained by the more hydrophobic nature of acpcPNA that make it more difficult to elute from the column under the same condition as aegPNA. However, the cross-linked product peaks could be clearly visible, and the calculation of the cross-linking efficiency according to the **Equation 2** is still valid here. In the case of DNA targets with one-base overhang, the results can be summarized as follows. AcpcPNA with the f-monomer (Pf) selectively cross-linked with C similar to the aegPNA probe. Similar results were observed with aegPNA with the f-monomer (Rf). However, in this case, additional cross-link formation was observed between Rf and DNA-T. The rigidity of the acpcPNA structure and the high stability of the acpcPNA-DNA duplex may preclude the available C to react with the NBS-activated furan moiety as in the case of aegPNA. AcpcPNA with the F-monomer (PF) can cross-link with A, C and G. The cross-linking reaction was confirmed by MALDI-TOF MS of PF+C1 as shown in **Figure 3.43** that revealed the mass of cross-linked product 8236.71 Da ($+K^+$). In addition, acpcPNA with the F-building block could also undergo some cross-linking reaction with DNA-T (via the C base adjacent to the T base, see section **3.2.5**). This unexpected result might involve the opening of the last base pair and it was somehow positioned close to the activated furan moiety. Note that the furan in the case of the F-building block was located closer to the terminal base pair than the f-building block.

Lastly, acpcPNA with the T-monomer (PT) can cross-link with A, C, and G similar to aegPNA. Again, DNA-T1 and DNA-T5 can also participate in the cross-linking reaction via the adjacent C base similar to the F-building block (see section **3.2.5**). AcpcPNA with the T-building block showed higher cross-linking efficiencies (43-68%) than the F-building block (13-64%) with most DNA targets. These results could be explained by the combination of the duplex formation and the stacking interaction. Even though the T-base cannot form a base pair in these cases (no A on the DNA strand located opposite to the T-building block), the T-building block could form a base stacking with the nucleobase at the overhang region. This stacking interaction might favorably position the base to react with the furan moiety to generate the cross-linked product.
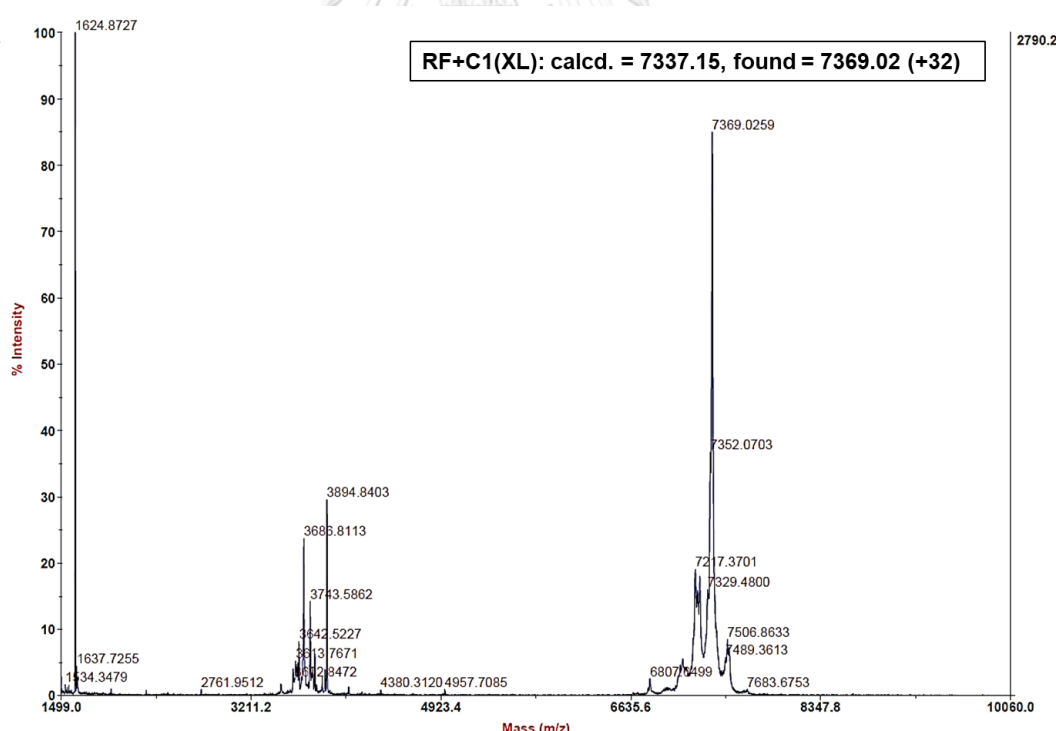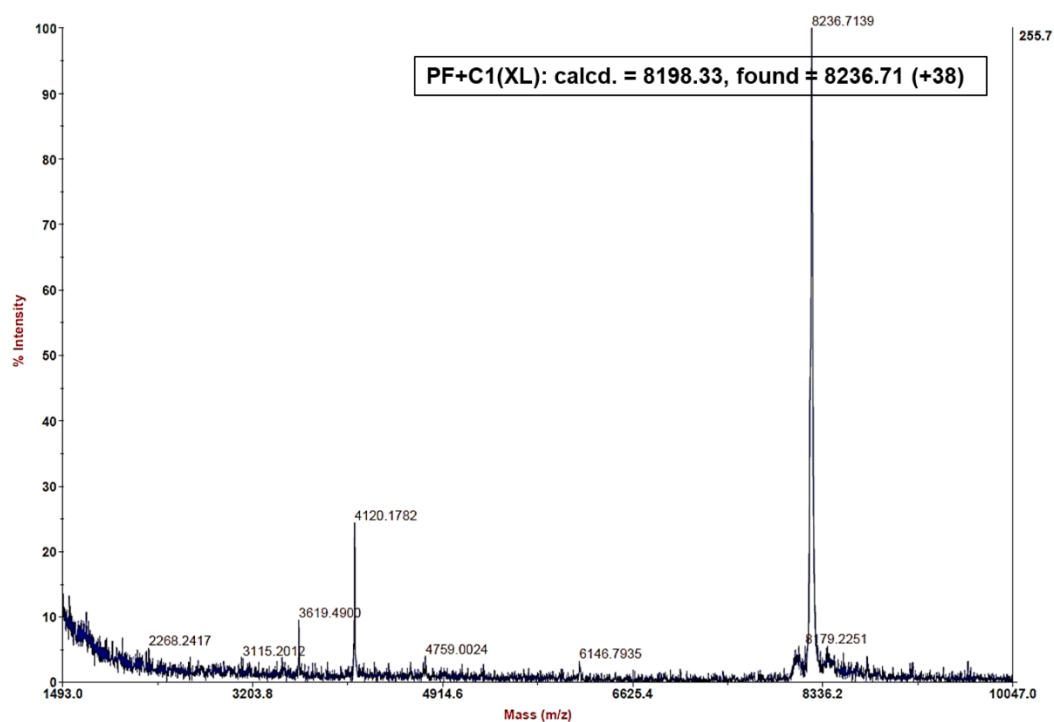
Figure 3.43 MALDI-TOF mass spectra of a cross-linked product (PF+C1) of acpcPNA with F-building block and DNA-C1 after treatment with NBS (4 equiv., 2h)

The overall cross-linking efficiencies of both acpcPNA and aegPNA with fully complementary DNA targets are summarized as shown in **Table 3.8**. The results revealed that C-base is the most favorable base to react with the furan-modified PNAs in all kinds of building blocks with 44-78% efficiencies for DNA-C1 and 33-84% efficiencies for DNA-C5. Only one extra base at the overhang region was sufficient to react with the furan on the PNA strand. The f-building block of acpcPNA (same as the terminal modification in section I) showed the best selectivity towards C-base (44% for DNA-C1 and 33% for DNA-C5). Similar selectivity was observed in aegPNA, but additional cross-linking with T1 and T5 was also observed. The F and T building blocks in both acpcPNA and aegPNA showed no selectivity toward A, C, and G base, with a marginal preference for C over other nucleobases, and the cross-linking efficiencies were generally higher than the f-building block.

Table 3.8 Summary of the cross-linking yield (%) of the terminal cross-linking reaction with fully complementary DNA targets

PNA	Furan building block	% cross-link with DNA[a,b]							
		A1	C1	G1	T1	A5	C5	G5	T5
R (aegPNA)	f	-	53	-	35	-	67	-	75
	F	50	78	47	-	22	84	36	-
	T	40	71	53	35	66	75	-	-
P (acpcPNA)	f	-	44	-	-	-	33	-	-
	F	34	64	60	13	46	37	48	45
	T	49	66	68	48	43	58	44	27

[a] based on RP-HPLC analysis calculated from **Equation 2**

[b] -: no cross-linked product formed

It should be noted that the cross-linking yields of acpcPNA here were somewhat smaller than in section I despite the same type of furan-modification was employed [Pf in section II vs PNA(T) in section I]. There are several possible reasons to explain the discrepancy. Firstly, the sequences and type of C-terminal modification (lysine vs arginine) were different which might affect the PNA:DNA stability, resulting in different cross-linking behavior. Secondly, the way the cross-linking experiment was performed was different. The use of 4 equiv. NBS in section II instead of 2 equiv. in section I might induce additional side reactions with acpcPNA probes or DNA targets. Finally, the conditions to evaluate the cross-linking results are different such as the HPLC conditions and the way the cross-linking efficiency was calculated. The different conditions that was used in section I and section I is summarized in **Table 3.9**.

Table 3.9 Summary of the conditions used in section I and section II

parameters	section I	section II
DNA: PNA concentration	7.5 µM:5 µM	5 µM:5 µM
incubation temperature of cross-linking reaction	37 ℃	25 ℃
denaturing the sample at 95 ℃ before adding NBS	yes	no
waiting period before the detection	immediate	vary
HPLC column temperature	rt	60 ℃
HPLC mobile phase	TEAA buffer (0.1 M, pH 7) and acetonitrile (5-80% MeCN over 30 min)	TEAA buffer (0.1 M, pH 7) with 5% acetonitrile and acetonitrile (0-100% MeCN over 12 min)
equiv. NBS	2	4
cross-linking reaction time	30 min (HPLC)	over 1 h

Despite these shortcomings, it can be concluded from the results that both aegPNA and acpcPNA modified with all types of furan at the terminal positions could undergo cross-linking reactions with DNA. The RP-HPLC experiments also confirms that the cross-linked products were formed, despite the absence or very weak bands were observed in the denaturing PAGE experiments. This was attributed to the poor ability of the SYBR gold dye to bind to short DNA and PNA-DNA duplexes. Therefore, the results from PAGE with SYBR gold staining should be interpreted with care.

3.2.4 Cross-linking studies of terminally furan-modified aegPNA and acpcPNA with partially complementary DNA targets
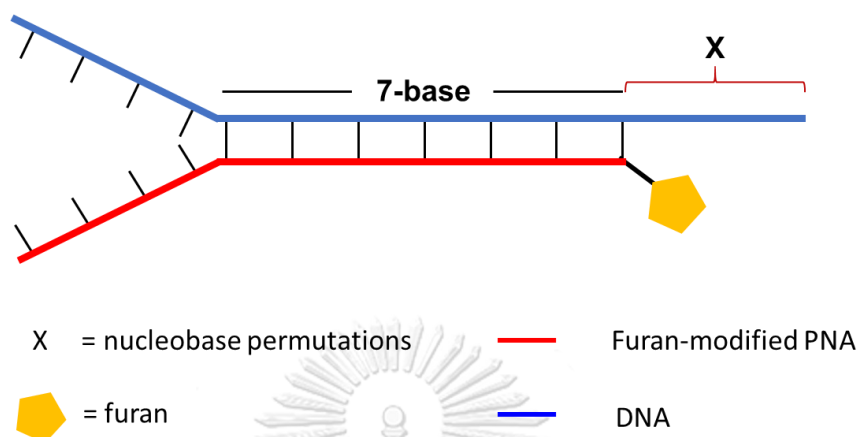
Figure 3.44 An illustration of the duplex formation of terminally furan-modified PNA with 7-base complementary DNA targets

This part focuses on the cross-linking studies of aegPNA and acpcPNA towards partial complementary DNA targets (containing only 7-base complementary region on the DNA strand) with the sequences shown in **Table 3.7.** The number of the complementary bases was reduced to seven bases to provide unpaired DNA sites that can bind better with the SYBR gold staining dye to improve the visualization as shown in **Figure 3.44**. Two groups of the DNA targets were studied as in the previous experiments with the fully complementary DNA targets. These include DNA-EA1 to DNA-ET1 bearing one-base overhang, and DNA-EA5 to DNA-ET5 bearing five-base overhang. Due to the limited quantities of the acpcPNA bearing the f-monomer were available, only the aegPNA and acpcPNA bearing the F- and T-monomers were compared.

3.2.4.1 Denaturing PAGE

DNA: 5'-<u>TACG</u>CATGCCC-**X**-3'
PNA: H2N-rrr-TCTAGTACGGG-**M**
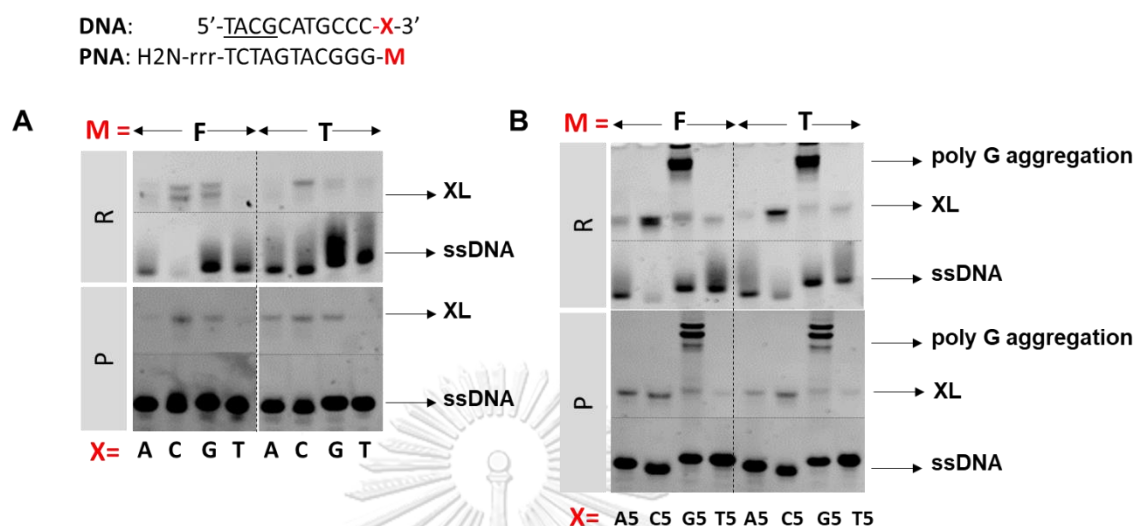
Figure 3.45 Denaturing PAGE results of the terminally modified aegPNA (R) and acpcPNA (P) towards 7-mer complementary DNA targets with one base overhang (A) and five bases overhang (B) in presence of NBS (4 equiv.). Conditions: [DNA] = [PNA] = 5 µM in the total volume = 100 µL, 4 equiv. NBS (1 equiv. every 15 min), for 2 h. The experiment was performed in 20% polyacrylamide gel containing 7 M urea at 230 V for 1 h and 40 min. XL: cross-linked product, ssDNA: single stranded DNA

According to the hypothesis of the inability of the SYBR gold dye to intercalate into the PNA-DNA duplex resulting in the poor visualization of the cross-linked adducts in the previous PAGE analysis, the experiments were repeated with the new DNA targets carrying only 7-mer complementary region (DNAEA1-DNAET5) to verify this. The experiments were performed only with the F and T PNAs under the similar condition as in the previous experiments. The denaturing PAGE results of the PNAs towards the 7-mer complementary DNA targets with one-base overhang and five-base overhang are shown in **Figures 3.45A** and **3.45B**, respectively.

In agreement with the hypothesis, the intensity of the bands in both the DNA targets were increased (**Figure 3.45**) compared to the previous results (**Figure 3.35**). In the case of DNA targets with one-base overhang, the aegPNA with the F-building

block (RF) preferred to cross-link with DNA-EA1, DNA-EC1 and DNA-EG1. The same trend was observed in the acpcPNA bearing the F-monomer of (PF). The aegPNA carrying the T-building block (RT) was found to selectively cross-link with DNA-EC1 while the acpcPNA PT showed no selectivity towards different nucleobases with the exception of T (DNA-EA1, EC1, EG1). Guanine base was found to participate in the cross-linking reaction with both terminally furan-modified aegPNA and acpcPNA.

In the case of DNA targets carrying five-base overhang, the results revealed that the F-building block of both aegPNA and acpcPNA could form the cross-linked product with all DNA targets [DNA-EA5, EC5, EG5 and ET5 (via the adjacent C-base)]. The T-building block selectively cross-linked with C only in both PNA systems while the fully complementary DNA targets preferred to cross-link with A and C. The aggregation of guanine was found in the free DNA as shown by the slowest moving bands in all samples containing DNA-G5 (**Figure 3.45B**). In the case of cross-linking with thymine, MALDI-TOF MS revealed that the thymine-linked product was formed with the last complementary cytosine base rather than thymine itself (see in section **3.2.5**).

3.2.4.2 HPLC

a) terminally furan-modified aegPNA

Figure 3.46 RP-HPLC showing the cross-linking experiments of the terminal furan-modified aegPNA with F-building block (RF) towards the 7-base complementary DNA targets with one base overhang (A) and five bases at the overhang targets (B) before (blue line) and after NBS addition (red line); Ox. PNA: oxidized PNA probe; XL: cross-linked product
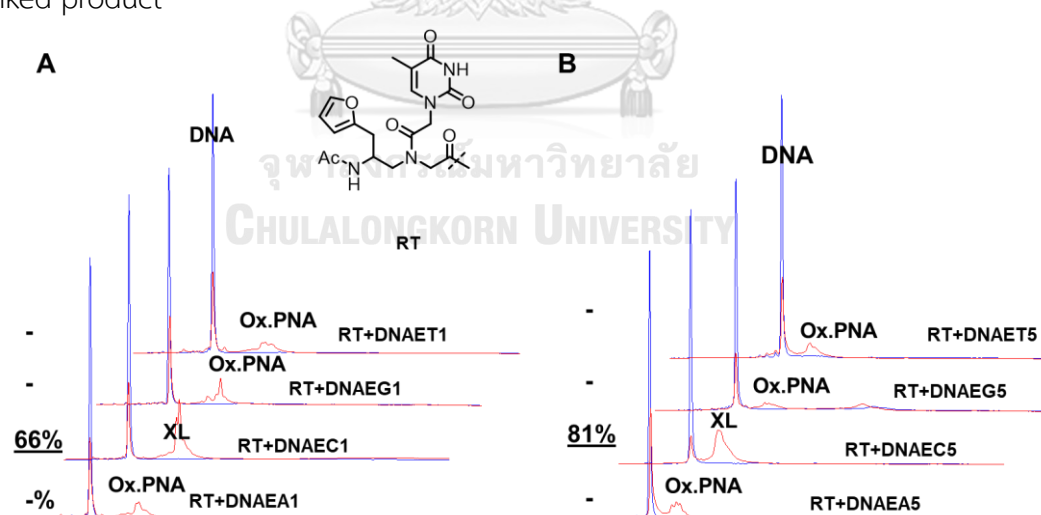
Figure 3.47 RP-HPLC showing the cross-linking experiments of the terminal furan-modified aegPNA with T-building block (RT) towards 7-base complementary DNA targets with one base overhang (A) and five bases overhang targets (B) before (blue line) and after NBS addition (red line); Ox. PNA: oxidized PNA probe; XL: cross-linked product

The cross-linking reactions of the same set of terminally furan-modified aegPNAs and 7-base complementary DNA targets were next investigated by RP-HPLC (see summary data in **Table 3.9**). The HPLC chromatogram of the aegPNA with the F-building block (**Figure 3.46**) showed the formation of cross-linked products with all DNA base permutations in both one base and five bases overhangs. The results were in line with the fully complementary targets, but the efficiency was higher (51-91% vs 22-84%). Concerning the aegPNA with the T-building block (**Figure 3.47**), the cross-linking results showed a good selectivity towards cytosine in both one base (66%) and five bases overhangs (81%). This is in sharp contrast to the fully complementary DNA targets, which could form cross-links with other nucleobases as well. Moreover, the cross-linking efficiency of the 7-base complementary targets (66-81%) were higher than fully complementary one (35-75%). Most notably, in the case of aegPNA carrying the F-monomer with DNA-EC1, the cross-linking efficiency was increased up to 91% when compared to the fully complementary target (78%). These results counterintuitively suggested that the destabilization of the duplex is beneficial for the cross-linking reaction. The imperfect duplex might affect the PNA-DNA duplex orientation which allows the furan moiety in the suitable position to react with the overhanging nucleobase.

b) terminally furan-modified acpcPNA
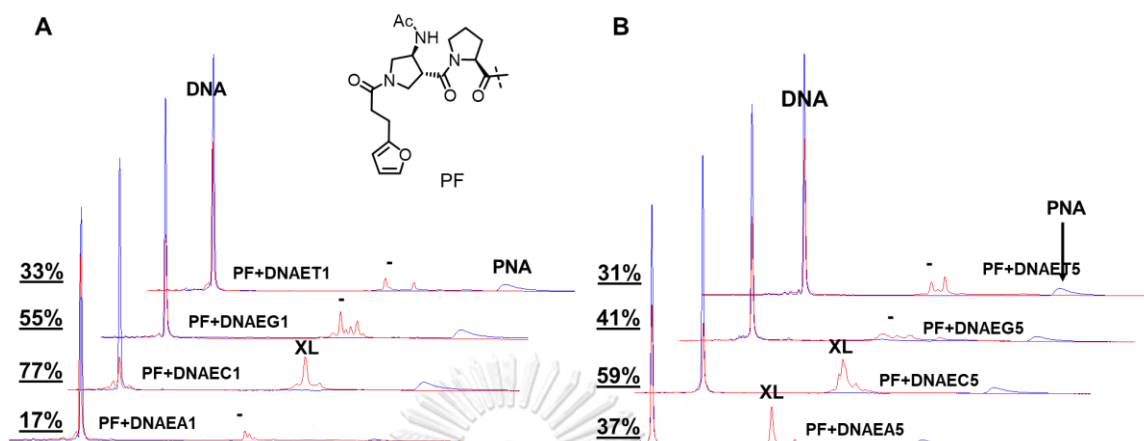
Figure 3.48 RP-HPLC showing the cross-linking experiments of the terminal furan-modified acpcPNA with F-building block (PF) towards 7-base complementary DNA targets with one base overhang (A) and five bases overhang targets (B) before (blue line) and after NBS addition (red line); Ox. PNA: oxidized PNA probe; XL: cross-linked product
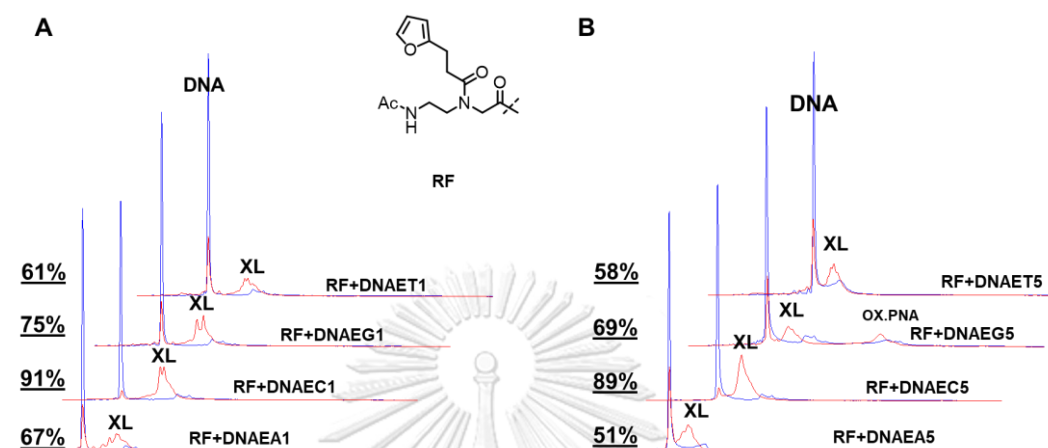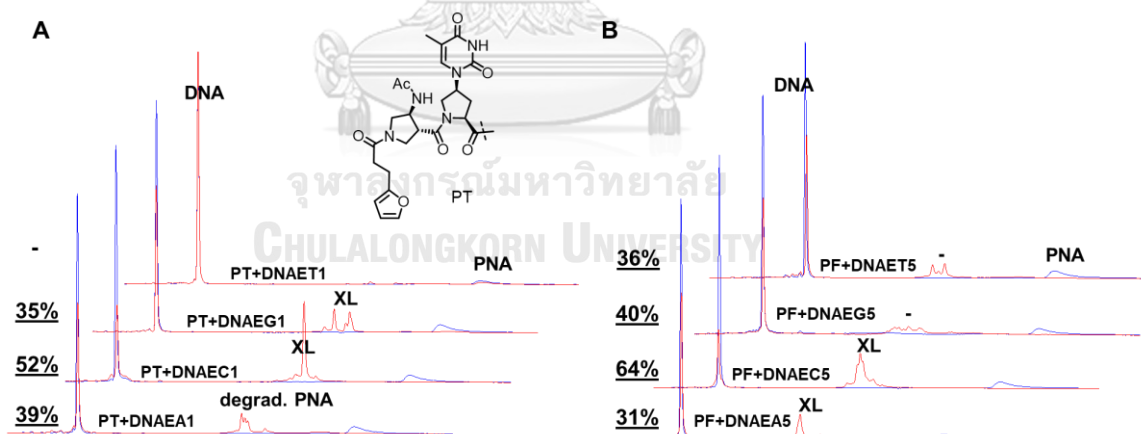

Figure 3.49 RP-HPLC showing the cross-linking experiments of the terminal furan-modified acpcPNA with T-building block (PT) towards 7-base complementary DNA targets with one base overhang (A) and five bases overhang targets (B) before (blue line) and after NBS addition (red line); Ox. PNA: oxidized PNA probe; XL: cross-linked product; -: no cross-linked product

Table 3.10 Summary of the cross-linking yield of the terminal cross-linking reaction with 7-base complementary DNA targets

PNA	Furan building block	% cross-link with DNA[a,b]							
		EA1	EC1	EG1	ET1	EA5	EC5	EG5	ET5
R (aegPNA)	F	67	91	75	61	51	89	69	58
	T	-[b]	66	-	-	-	81	-	-
T (acpcPNA)	F	17	77	55	33	37	59	41	31
	T	39	52	35	-	31	64	40	36

[a] based on RP-HPLC analysis calculated from **Equation 2**

[b] -: no cross-linked product formed

Similar experiments were also performed with the terminally furan-modified acpcPNA probes PF and PT. The results could be summarized as follows. The results of the acpcPNA carrying the *N*-terminal F-building block (**Figure 3.48**) show the same trend as the fully complementary DNA targets. Cross-linking was observed in target DNA with all base permutations (DNA-EA1, EC1, EG1, ET1 (via the adjacent C-base)) with both one base and five-base at the overhang targets. The highest cross-linking efficiency of acpcPNA probes was observed between the PF probe and the DNA-EC1 target (77%). The efficiency increased from 64% in the case of complementary target DNA-C1. These results are also consistent with the aegPNA cases above.

For the acpcPNA carrying the T-building block, the cross-linking occurred with most nucleobases with a slight preference for C over other nucleobases (**Figure 3.49**). The cross-linking efficiencies for the F-building block (17-77%) were somewhat higher than the fully complementary DNA targets (13-64%) similar to the cases of aegPNA probes. On the other hand, the cross-linking efficiencies of the T-building block were similar in for the fully complementary and for the 7-base complementary targets (43-68% vs 31-64%). This might be explained by the stacking interaction of the nucleobase with the thymine of the T-building block that locally stabilize the duplex and thus positioned the furan building block that allows the cross-linking

reaction regardless of the overall duplex stability. Small degradation peaks were observed in the HPLC chromatogram for both aegPNA and acpcPNA cross-links with 7-base complementary targets. This could be due to the less stable PNA:DNA duplex systems resulting in the degradation of the oxidized PNA probe as previously shown in section I (**3.1.7**).

In conclusion, the PAGE experiments revealed that both aegPNA and acpcPNA probes carrying F and T building blocks at the *N*-termini were able to form cross-links with short DNA targets with one- and five-base overhangs. With the exception of the aegPNA probe with T-building block (RT), which showed preference for cross-linking only with C, other probes (RF, PF, PT) could form the cross-linked products with all nucleobases (A, C, and G) indiscriminately. The reduction of the complementary region in the DNA target to only 7-base not only improved the visualization of the PAGE but also unexpectedly improved the cross-linking efficiencies for both types of PNA probes. Also, the cross-linking efficiency of acpcPNA were lower than aegPNA probes for both the F and T series similar to the studies with fully complementary DNA targets.

3.2.5 Cross-linked product identification by MALDI-TOF MS

The detection of the cross-linking product by MALDI-TOF MS was rather difficult. The method of sample preparation and the matrix optimization are reported in the appendix **(Figure A26 - A27)**. It was found that the sample preparation by dropping the sample as the first layer and then followed by dropping of the matrix (2,5-DHB) as the second layer gave the best results. According to the literature,[100] the cross-linked product mass could be detected only when the sample was free from salt and its concentration must be sufficient (5 µM). In some cases, the cleavage of the *N*-glycosidic bond between the nucleobase A and the sugar backbone of the DNA strand under the MALDI-TOF analysis conditions was observed as shown in **Figure 3.50**. Manicardi and coworkers[100] revealed that the mass of a furan-modified aegPNA cross-linked with adenine in the DNA strand was 131 Da larger than the starting PNA. This corresponded to the adenine+PNA adduct resulting from the *N*-glycosidic bond cleavage explained above. Hence, we could take advantage from this

N-glycosidic bond cleavage to identify the base that cross-linked to the PNA probe. According to the mechanism shown in **Figure 3.50**, the mass increase would be 107.1, 131.1 and 148.1 Da for cytosine, adenine and guanine, respectively. For thymine, if the cross-link occurs, the mass increase would be 122 Da.
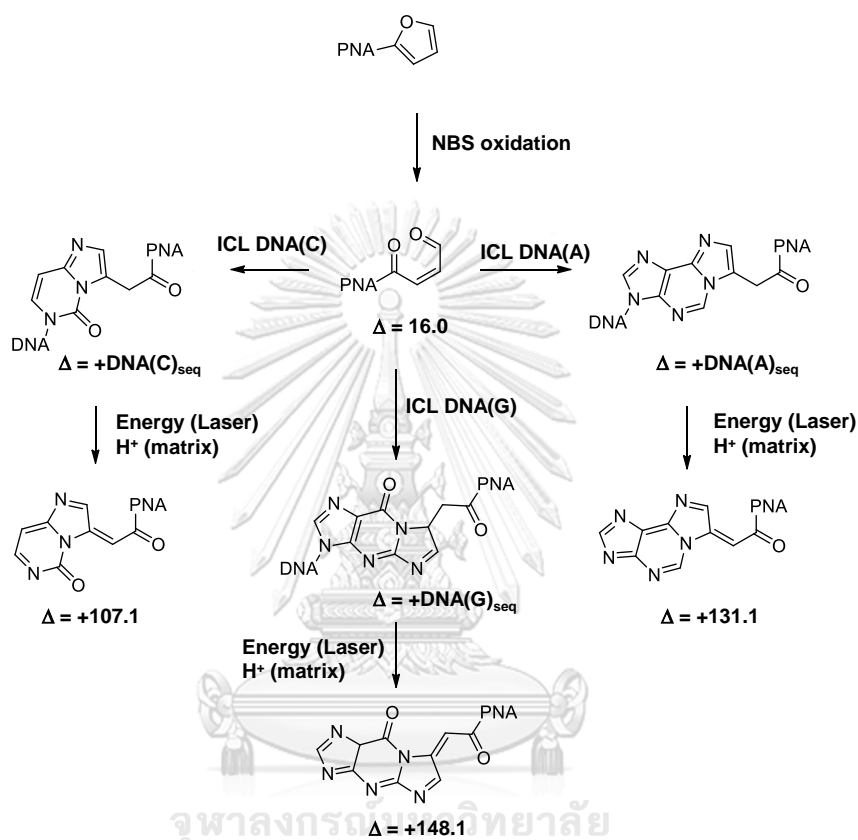
Figure 3.50 The proposed of cross-linked product degradation fragmentation products generating by MALDI-TOF laser and the matrix.

a) Cross-linking with DNA-T

Sine thymine does not contain exocyclic amino group that can react with the furan moiety, it is notable to observe some cross-links formation with the DNA-ET5 targets. It was proposed that cross-linked product observed in the case of DNA-T should result from the reaction between the activated furan probe at the C rather than the T itself. According to the MALDI-TOF MS analysis of the cross-link product between RF-PNA and DNA-ET5 which carried five T-base at the overhang region, the cross-linked product peak as 8583.55 Da (K⁺ ion adduct) in **Figure 3.51A** was

observed, together with the degradation fragmentation products as 3896.76 Da. For the similar reaction with DNA-ET1 which carried only one T-base at the overhang region, the cross-linked product peak was very small but a peak at 3897.29 Da similar to that derived from DNA-ET5 was prominent as shown in **Figure 3.51B**. The degradation product was assigned to the RF-PNA linked to cytosine (+107.1 Da) rather than thymine (+122 Da).
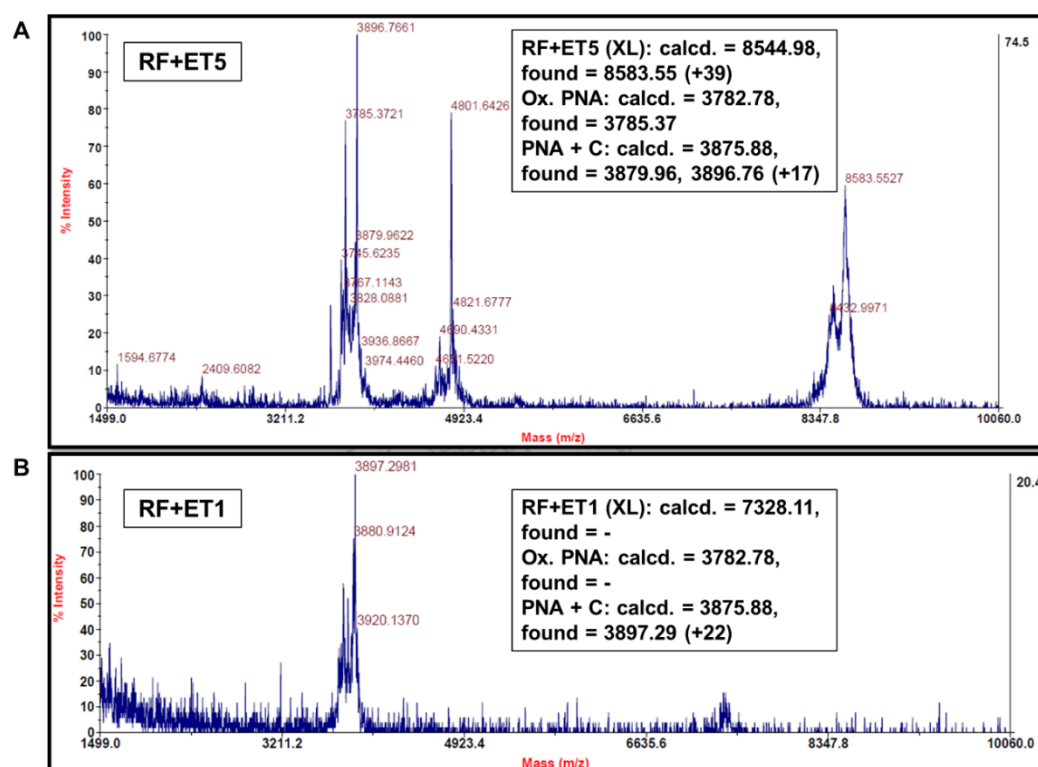
Figure 3.51 MALDI-TOF mass spectra of cross-linked product of the aegPNA with F-building block and DNA (T5) (A) or with DNA (T1) (B) after treatment with NBS (4 equiv., 2h)

b) cross-linking with DNA-G

To confirm the participation of G participating in the cross-linking reaction, the cross-linked products between both PNA systems (RF and PT) and the guanine DNA (DNA-EG1) were analyzed by MALDI-TOF MS. The crude reaction products were purified by RP-HPLC to isolate the pure cross-linked products before characterization by MALDI-TOF MS. Very clear mass signals that corresponded to the linkage of the guanine-DNA with both furan-modified aegPNA and acpcPNA probes were observed as shown in **Figure 3.52**. For the cross-linking product RF-EG1, the signal was observed as 7,389.88 Da. This was different from the calculated mass (7,353.18 Da) by 37 *m/z* units, which might be attributed to the adduct with potassium ion (M+K$^+$). In the case of PT-EG1 cross-link (calculated mass = 8314.43 Da), the difference was of 23 *m/z* units (8337.79 Da) due to the formation of sodium ion adduct (M+Na$^+$). Another evidence to support the identity of the cross-linked products was the decomposition peak of the cross-linked products due to the cleavage of the glycosidic bond (**Figure 3.50**) when they were exposed to the laser light during the ionization process as previously reported by Manicardi et al.[100] The degradation product of the aegPNA linked with only base G was found in 3937.64 (calculated mass = 3915.91 Da). This was different from the calculated mass by 22 *m/z* units, which might be attributed to the formation of sodium ion adduct (M+Na$^+$) as shown in **Figure 3.52A**. A similar result was observed in the acpcPNA linked with G. The degradation product mass was detected at 4881.92 Da which was in good agreement with the calculated mass (4877.16 Da) (**Figure 3.52B**). The experiments thus clearly confirmed, and revealed for the first time, that the cross-linking of furan-modified probe with G is possible.
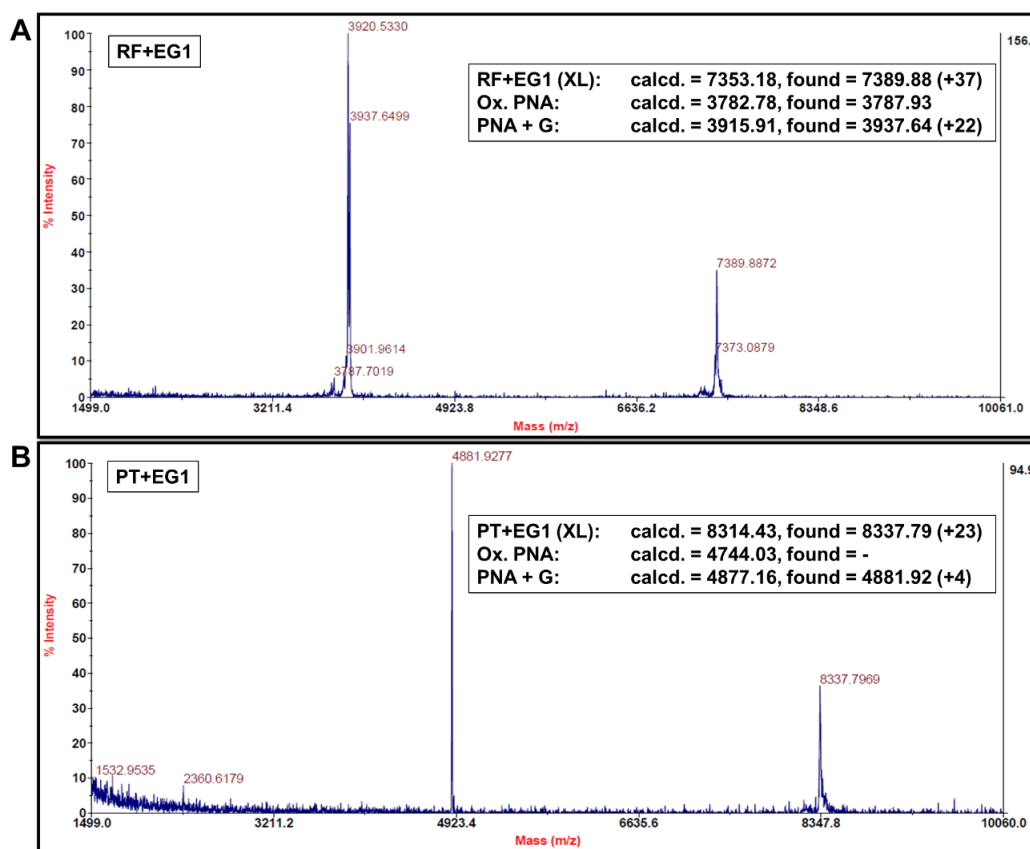
Figure 3.52 MALDI-TOF mass spectra of the cross-linked product of Guanine with (A) aegPNA (RFEG1) and (B) acpcPNA (PTEG1) after treatment with NBS (4 equiv., 2h)

c) PNA-DNA complex

To reveal the identity of the broad peaks found in the RP-HPLC chromatogram in the same region as the cross-linked product (**Figure 3.36A** - **B**, **Figure 3.37B**, **Figure 3.38B** in section **3.2.3.2**), the peaks were collected and investigated by MALDI-TOF MS. The results revealed that no mass peak of the cross-linked product was found in the expected region. Only the mass of the Ox. PNA and DNA mass were present as shown in the case of RF+G1 (**Figure 3.54A**) and RT+T5 (**Figure 3.54B**). Moreover, the degradation product was not observed. These results could confirm that the broad peak found in the RP-HPLC in the same region of the cross-linked product may not always be the cross-linked product. They are most likely to be the complex of the Ox.PNA and DNA, and thus it is suggested that the identity of the cross-linked product should always be confirmed by MALDI-TOF MS.
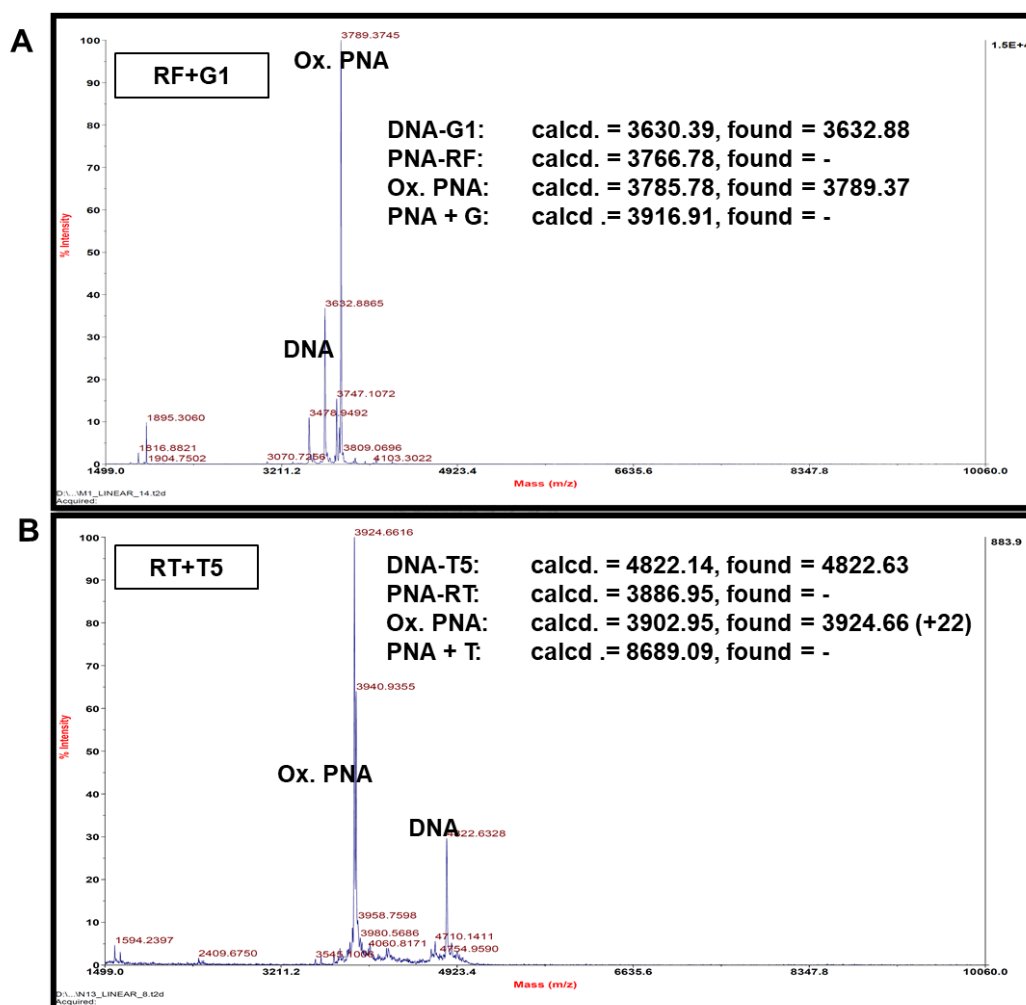
Figure 3.53 MALDI-TOF mass spectra of the PNA-DNA complex of RF+G1 (A) and RT+T5 (B) after treatment with NBS (4 equiv., 2h)

3.2.6 Effect of the C-position at the overhang of the strand

The next experiment was performed to study the effect of the position of C in the five-base overhang towards the cross-linking efficiency of furan-modified PNAs at the terminal position. Four thymine bases were incorporated to substitute the cytosine in the overhang region of the target DNA with a short (7-base) complementary region, leaving only one C available, to give DNAET4C1-DNAET4C5 (**Table 3.7**). PAGE and RP-HPLC analyses were performed to investigate the cross-linking efficiency. The results are presented in **Figures 3.54** - **3.55**. Based on the HPLC data, the cross-linking results revealed that the formation of the cross-linked products was similar among different DNA targets for the aegPNA probe, RF, therefore the position of the C base does not affect the cross-linking efficiency in this case. This is not surprising due to the flexibility of the furan part caused by the abasic building block. The same is true for the T-building block, the cross-inking formation were not different among different positions of the C base. The result could be explained in terms of flexibility as for the F-building block. Because the T-building block in this T4C1 system could not form any base pairing with the nucleobase at the overhang region so that it was free to react with any C position.

However, in the case of acpcPNA probes, the cross-linking results from the PAGE were not consistent with the result in the HPLC for both F and T building blocks. In the PAGE's analysis of the acpcPNA with the F-building block revealed that the further away C position the lower cross-linking efficiency (**Figure 3.54B**). While the opposite trend was observed in the case of T-building block (**Figure 3.55B**). This is in contrary to the HPLC results of both F and T building blocks, which showed similar cross-linking efficiencies regardless of the position of the base C. Nevertheless, the HPLC results were deemed to be more reliable than the PAGE for the reasons explained in section **3.2.3**.

Figure 3.54 RP-HPLC showing the cross-linking experiments of the terminal furan-modified aegPNA (A) and acpcPNA (B) with F-building block towards T4C1 overhang DNA targets (DNAET4C1-DNAET4C5) before (blue line) and after NBS addition (red line), equipped with their PAGE analysis in the presence of NBS. Conditions: [DNA] = [PNA] = 5 µM in the total volume = 100 µL, 4 equiv. NBS (1 equiv. every 15 min), for 2h. The experiment was performed in 20% polyacrylamide gel containing 7 M urea at 230 V for 1 h and 40 min. XL: cross-linked product, ssDNA: single stranded DNA

Table 3.11 Summary of the cross-linking yield of the terminal cross-linking reaction with 7-base complementary DNA targets with variable T position

PNA	Furan building block	% cross-link with DNA[a]				
		T4C1	T4C2	T4C3	T4C4	T4C5
R	F	74	80	76	81	76
(aegPNA)	T	70	68	64	68	70
T	F	65	50	65	61	66
(acPcPNA)	T	55	58	55	57	56

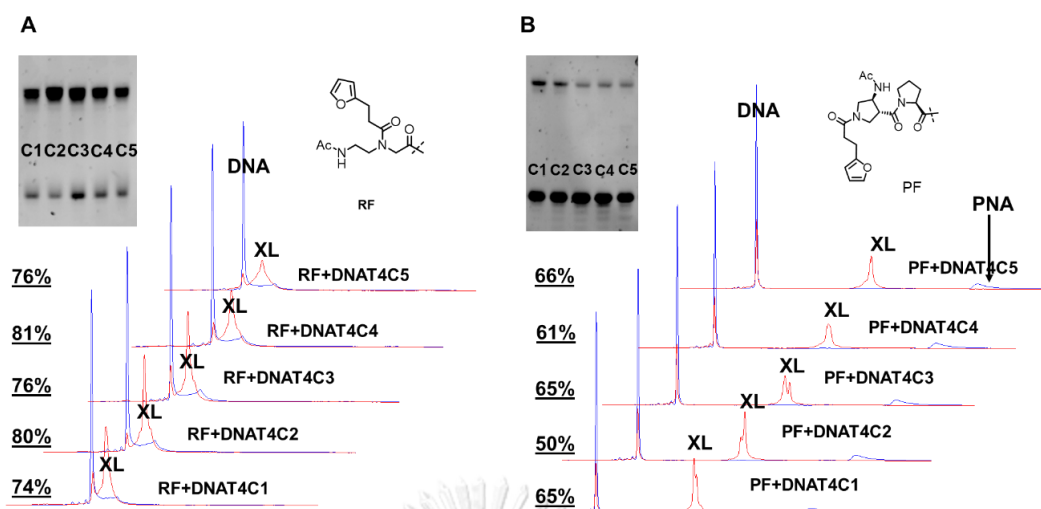[a] based on RP-HPLC analysis calculated from **Equation 2**
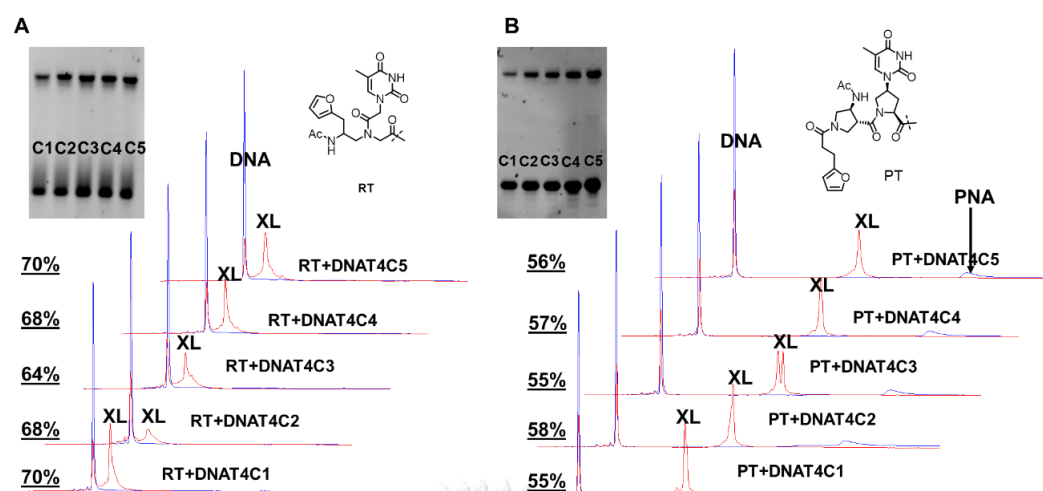
Figure 3.55 RP-HPLC showing the cross-linking experiments of the terminal furan-modified aegPNA (A) and acpcPNA (B) with T-building block towards T4C1 overhang DNA targets (DNAET4C1-DNAET4C5) before (blue line) and after NBS addition (red line), equipped with their PAGE analysis in the presence of NBS. Conditions: [DNA] = [PNA] = 5 μM in the total volume = 100 μL, 4 equiv. NBS (1 equiv. every 15 min), for 2h. The experiment was performed in 20% polyacrylamide gel containing 7 M urea at 230 V for 1 h and 40 min. XL: cross-linked product, ssDNA: single stranded DNA

From the results, it could be concluded that both terminally furan-labeled aegPNA and acpcPNA formed cross-links with base overhangs in a non-position-dependent fashion, probably due to the flexibility of the overhanging DNA part that might fold back to allow any bases in that region to react with the activate furan. Consistent with previous results, the furan-modified aegPNA seem to form cross links more efficiently than acpcPNA. The difference was more pronounced in the case of the PNAs carrying the F-monomer (aegPNA: 74-81%; acpcPNA: 50-66%) than the T-monomer (aegPNA: 64-70%; acpcPNA: 55-58%), possibly due to the stacking effect of the T monomer as explained earlier (section **3.2.3.2**).

3.2.7 Cross-linking studies of internally furan-modified acpcPNA

The work in this part involves the cross-linking study of PNA modified with furan in the middle of the strand. Two building blocks, consisting of **F** and **T** as shown in **Figure 3.58**, were inserted into the acpcPNA with the sequences P-**XF** and P-**XT** (X = A, T, C, G) as shown in **Table 3.6**. The combination of the two furan building blocks (designated as **M**) and the nucleobase adjacent to the modification site towards the *N*-terminal position (designated as **P**) created four PNA sequences for each furan building block (F and T). The sequences of DNA involved are shown in **Table 3.7**. In the F-series, the furan was placed on the backbone next to an abasic site. Hence any DNA bases opposite to the furan modification would not form a base pair with the PNA base. In the T-series, the furan was incorporated on the backbone next to the T monomer. Hence, a complementary base pair would form only if the opposite nucleobase in the DNA strand is A. Complementary DNA targets with a longer sequence (16-mer) than the PNA sequence (11-mer) was designed in order to facilitate the detection by SYBR gold staining. The cross-linking reactions of the 8 PNA probes were performed with all 16 possible base permutations at the X-Y position of the DNA targets as shown in **Table 3.7** (AA, AC, AG, AT, CA, CC, ..., GA, GC, ..., TA, TC, ...).

DNA: 5'-GCAGATC-**X Y**-GCCCGGC-3'
PNA: H$_2$N-rrr-TCTAG-**M P**-CGGG-Ac

X: the facing nucleobase of **M** building block
Y: the previous DNA nucleobase
P: the nucleobase before **M** building block
M: furan modified acpcPNA building block
The PNA sequence was written from C to N direction.

F   **T**
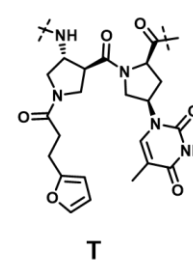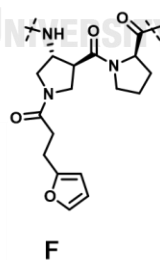
Figure 3.56 Furan building blocks and PNA/ DNA sequences in internal modified PNA

DNA: 5'-GCAGATC-**XY**-GCCCGGC-3'
PNA: H$_2$N-rrr-TCTAG-**MP-**CGGG-Ac

Code   PMXY   M facing X
                     P facing Y

Example

AFAC

5' ⟶ 3'
DNA: 5'-GCAGATC-**AC**-GCCCGGC-3'
PNA: H$_2$N-rrr-TCTAG-**F A-**CGGG-Ac
                     C ⟵ N
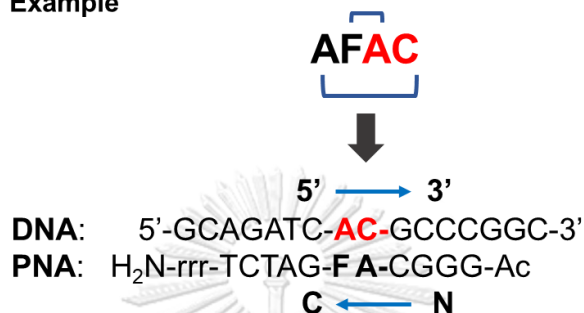
Figure 3.57 An illustration to show how to interpret the PNA-DNA combination code (**PM**XY) used in this section

     Figure 3.59 explains how to interpret the code that performed in this section. The PNA sequences are written from C to N direction and the DNA sequences are written from 5' to 3' direction. The **P** on the PNA strands refers to the nucleobase preceding the furan monomer and **M** refers to the furan monomer (F or T). X and Y refer to the base permutations on the DNA strand opposite to **M** and **P**, respectively. The code was designated as **PM**XY, whereby the P base in the PNA strand is pairing with the Y of the DNA strand and the **M** base in the PNA strand is facing the X base in the DNA strand. For example, **AF**AC means the base A in the PNA strand is facing the base C on the DNA strand and the F-monomer on the PNA strand is facing the base A on the DNA strand.

3.2.7.1 RP-HPLC

The cross-linking reactions were performed as described in section **2.2.4**. The 15 µL aliquots sampled from the reactions were directly analyzed by RP-HPLC to investigate the cross-linking reactions. Unfortunately, in the HPLC chromatograms only the DNA peaks were detected for all reactions, even in the absence of the NBS whereupon both DNA and PNA peaks should have been observed. In the presence of NBS, no cross-linked products nor PNA peaks were observed in the HPLC chromatograms as shown in the **Figure 3.58**. However, the reduction of the DNA intensity was clearly evidenced after NBS addition that might be attributed to the formation of the cross-linked products. Moreover, the crude reactions of the same sample for the HPLC experiments were analysed by MALDI-TOF MS further. Representative results from the **CF**GC and **CF**TC pairs (exhibiting highest band's intensity in PAGE experiment in **Figure 3.59**) showed no cross-linked products in the expected region. However, the single-nucleobase adduct with PNA previously attributed to the degraded cross-linked products were indeed detected. This indirectly implied that the PNA had formed the cross-link with cytosine as shown in **Figure 3.58**. Thus, the HPLC technique might not be suitable for studying the cross-linking of the internal furan-modified acpcPNA. The fact that no PNA band was observed even before the NBS addition suggests that the HPLC condition might not be optimal, or the PNA concentration might not be correct, although the latter was unlikely since the concentration was determined by UV spectrophotometry as in previous experiments. Unfortunately, due to the limited amounts of furan-modified acpcPNA available and the time constraint during the COVID-19 outbreak in Belgium, it was not possible to repeat the HPLC experiments.
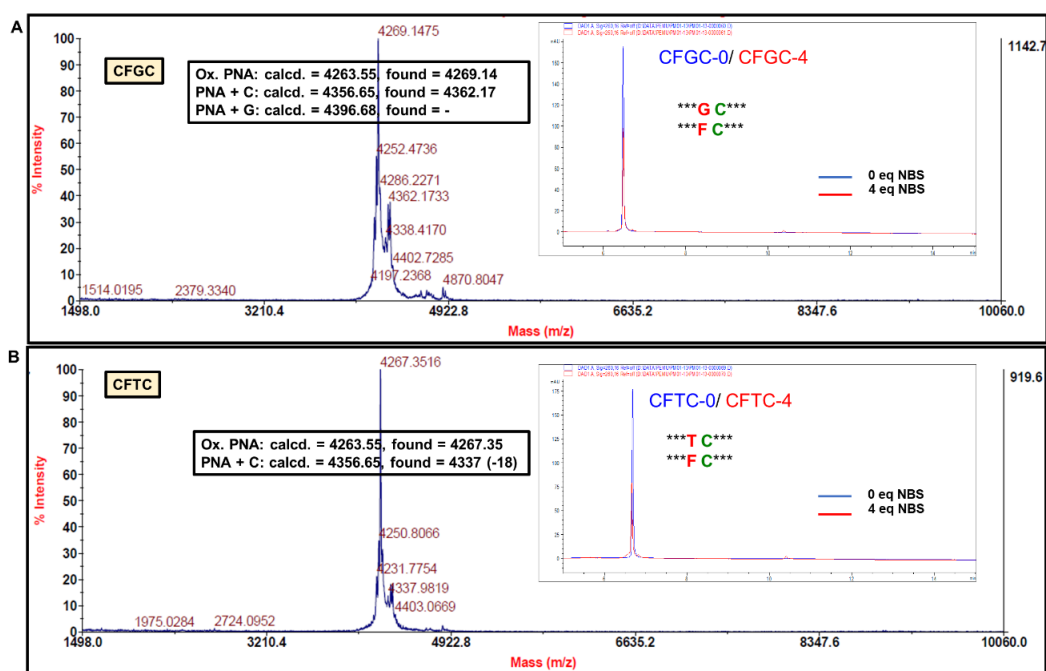
Figure 3.58 The MALDI-TOF MS of the cross-linked product of acpcPNA with F-building block with DNA-GC (A) and DNA-TC (B), equipped with their RP-HPLC chromatograms at before (blue line) and after NBS addition (red line); Ox. PNA: oxidized PNA probe; XL: cross-linked product

3.2.7.2 Denaturing PAGE

a) F-monomer

Since the cross-linked products of the internal furan-modified acpcPNA could not be detected by RP-HPLC, the denaturing PAGE experiments was next performed to investigate the cross-linking reactions. The same samples prepared for the RP-HPLC experiments were used for the PAGE analyses. **Figure 3.59** shows the PAGE experiments of acpcPNA with F-building block. The rows refer to the 16 DNA permutations (XY). The column represents the base combinations in the PNA with the furan building block (**PM**).
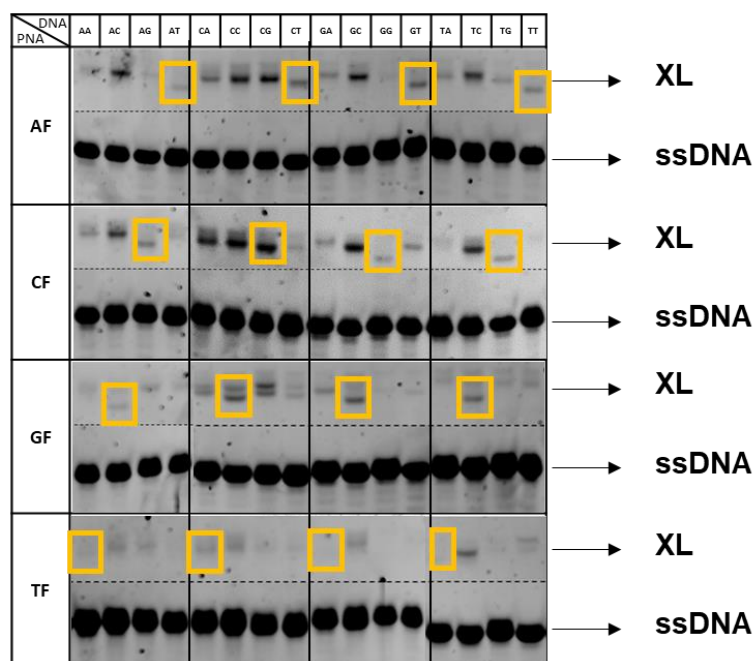
Figure 3.59 The denaturing PAGE results of the cross-linking reactions of acpcPNA carrying F-building block (**AF**, **CF**, **GF**, and **TF**) and 16 DNA permutation sequences. The cross-linking bands of the fully matched DNA sequences (ignoring the abasic site) are marked by the yellow boxes. Conditions: [DNA] = [PNA] = 5 µM in the total volume = 100 µL, 4 equiv. NBS (1 equiv. every 15 min), for 2h. The experiment was performed in 20% polyacrylamide gel containing 7 M urea at 230 V for 1 h and 40 min. XL: cross-linked product, ssDNA: single stranded DNA

Because of the absence of a nucleobase, the F-building block in acpcPNA cannot pair with the opposite nucleobase in the DNA strand. Nevertheless, the denaturing PAGE results in **Figure 3.59** showed the cross-linked product bands with several DNA targets. As with previous PAGE experiments, these cross-linking bands appear as slow-moving bands relative to the ssDNA bands. The PAGE studies confirmed at least qualitatively that the cross-linking was indeed formed in many cases. The AF PNA sequence (Ac-GGGC**AF**GATCT-rrr-NH$_2$), which carries carrying the base A adjacent to the furan modification site towards the *N*-terminal direction, formed the cross-linked products when the base C was present on the DNA strand either at the position X (opposite to the furan building block) or Y (opposite to the

base adjacent to the furan building block, which is A in this case). The base adenine and guanine also participated in the cross-linking reaction with the AF sequence, but the cross-linking efficiency was lower than with C as shown by the less intense cross-linking bands in the gel. In the case of the CF sequence (Ac-GGGC**CF**GATCT-rrr-NH$_2$) carrying the base C adjacent to the furan modification site, the strong cross-linking bands were observed in several PNA-DNA combinations. The same trend previously noted for the AF sequence was observed, but the cross-linking efficiency was even higher. For the GF sequence (Ac-GGGC**GF**GATCT-rrr-NH$_2$), the cross-linking reactions were less efficient and occurred only when the DNA base at the positions X or Y was cytosine, but not adenine or guanine. Finally, the TF sequence (Ac-GGGC**TF**GATCT-rrr-NH$_2$) only formed the cross-linked products with the DNA sequences which had the base C opposite to the base T in the PNA strand. In all cases, when the complementary base pairing could form (marked by yellow boxes in **Figure 3.59**), the bands appeared at the slightly different position from the cross-linked products which might represent PNA-DNA duplexes. It can be concluded that in the case of F-building block, the cross-linking reaction could occur most efficiently when there is a cytosine base present at the opposite position or nearby the furan part. Some cross-linking reactions was also observed with A and G, albeit with less efficiencies. This is generally possible since the F-building block is an abasic building block that could not form hydrogen bonds with any nucleobase. Thus, whenever a nucleobase with an exocyclic amino group was located opposite to this furan building block or adjacent to it, the cross-linking reaction would occur regardless of the ability to form perfect base pairings between the PNA and DNA strands.

b) T-monomer

In the case of acpcPNA with the T-building block, the cross-linking results of the four acpcPNA sequences (AT, CT, GT, and TT) with the same 16 DNA sequences as in the previous experiments are presented in **Figure 3.60**. The results revealed that the cross-linking reaction was much less efficient (i.e., more selective) for the T-building block when compared to the F-building block. For the AT sequence (Ac-GGGC**AT**GATCT-rrr-NH$_2$), the cross-linking reaction did not occur when adenine was present in the DNA strand at the opposite position to the T-building block (AA, AC, AG, AT), regardless of the identity of the adjacent base. In these cases, the T-building block would form a stable base pair with the opposite A base on the DNA strand. On the other hand, when the base opposite to the T-building block was cytosine (as in CA, CC, CG, CT), a mismatched pair was formed and thus the cross-linking reaction could occur, albeit in much less efficiency when compared to the F-building block. The accessibility of the target nucleobases in these cases, even though being unpaired, is evidently more difficult than in the case of the F-building block with a complete abasic site. The cross-linked products were also formed with the DNA target carrying the cytosine base opposite to the A in the PNA strand as in the DNA-GC and DNA-TC sequences. The cross-linking reactions of the CT sequence (Ac-GGGC**CT**GATCT-rrr-NH$_2$) were possible only with a few DNA targets including DNA-CC, DNA-GC and DNA-TC. All of these sequences carry one or more cytosines located close to the T-building block. The GT and TT sequences seemed to be unable to participate in the cross-linking reaction, as shown by the presence of very faint bands or complete absence of the cross-linked products. Nevertheless, the results fully confirm that the presence of an unpaired base located close to the furan modification site is essential for the cross-linking formation, and that C is the most susceptible target for the furan cross-linking. The results from this and the previous experiments also suggest that the identity of the adjacent base to the furan cross-linker also affect the cross-linking efficiency, with A and C were preferred over T and G.
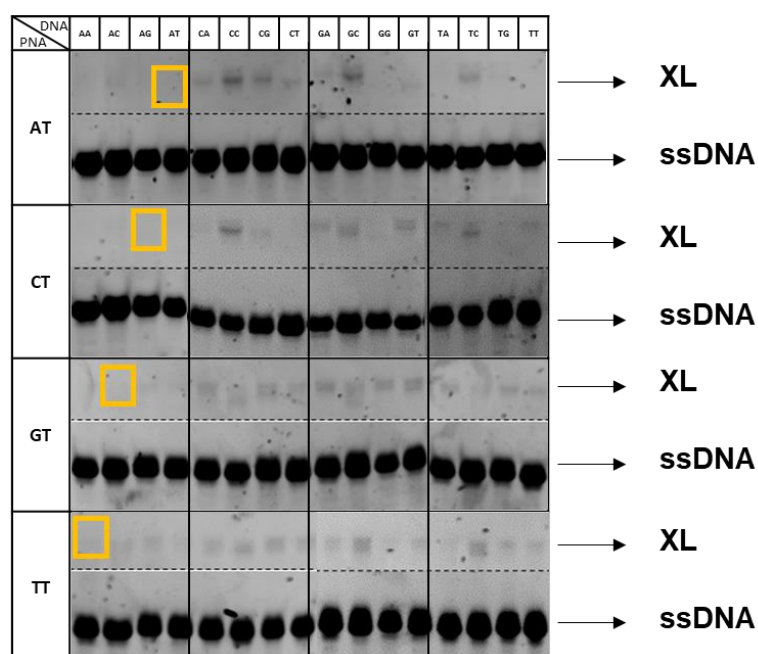
Figure 3.60 The denaturing PAGE results of the cross-linking reactions of acpcPNA carrying F-building block (**AT**, **CT**, **GT**, and **TT**) and 16 DNA permutation sequences. The cross-linking bands of the fully matched DNA sequences are marked by the yellow boxes. Conditions: [DNA] = [PNA] = 5 µM in the total volume = 100 µL, 4 equiv. NBS (1 equiv. every 15 min), for 2h. The experiment was performed in 20% polyacrylamide gel containing 7 M urea at 230 V for 1 h and 40 min. XL: cross-linked product, ssDNA: single stranded DNA

To further present the information in a more quantitative fashion, the cross-linking efficiencies were determined from the relative intensities of the DNA and the cross-linked bands according to **Equation 3**, where the intensity of the starting DNA is measured at the starting DNA band (data not shown) before adding NBS by using ImageJ. The same area as measuring in the starting DNA is used to measure the intensity of the cross-linked product.

Equation 3; $\quad$ % cross-link $= \dfrac{\text{intensity of cross-linked product}}{\text{intensity of starting DNA}} \times 100$

Table 3.12 Summary of the cross-linking yield of the internal cross-linking reaction

PNA	% cross-link with DNA[a,b]															
	AA	AC	AG	AT	CA	CC	CG	CT	GA	GC	GG	GT	TA	TC	TG	TT
AF	-[b]	28	-	-	14	35	37	-	-	37	-	-	-	30	-	-
CF	-	36	-	-	54	63	58	-	-	42	-	-	-	35	-	-
GF	-	-	-	-	-	26	27	-	-	17	-	-	-	16	-	-
TF	-	-	-	-	-	-	-	-	-	-	-	-	-	24	-	-
AT	-	-	-	-	-	25	18	-	-	22	-	-	-	-	-	-
CT	-	-	-	-	-	15	24	-	-	16	-	15	-	5	-	-
GT	-	-	-	-	-	-	-	-	7	-	-	9	-	-	-	-
TT	-	-	-	-	-	-	-	-	-	6	-	-	-	-	-	-

[a] based on PAGE analysis calculated from **Equation 3** by using ImageJ

[b] estimated lower limit for the determination of cross-linked efficiency is 5%, - = <5 %

According to the results from the PAGE analysis in **Table 3.11**, the acpcPNA with F-building block, which was unable to form stable base pairs with the opposite nucleobase in the DNA strand, provided a higher cross-linking efficiency than the T-building block. The highest cross-linking yield at 63% was observed in the case of **CF**CC system. While the highest cross-linking efficiency of the T-building block was only 25%, and was observed with the **AT**CC system, whereby no base pairing was formed. Although the results can at best interpretable in semi-quantitative fashion, they fully supported the important conclusion from previous experiments that the accessibility of the nucleobase is essential for the cross-linking reaction. The presence of unpaired nucleobases according to this design could therefore improve the cross-linking efficiency of the internally modified acpcPNA.

3.2.8 Conclusion

In conclusion, the work carried out in this section involved the cross-linking studies of furan-modified PNA at the end (aegPNA vs acpcPNA) and the middle of the strand (acpcPNA only). The work was divided into two main parts, consisting of the studies of the PNAs modified with furan at the terminal and internal positions.

In the case of terminally furan-modified PNA, both aegPNA and acpcPNA were compared in the context of three different furan building blocks (f, F, and T). The DNA targets include those carrying one- and five-base overhangs (A, T, C and G). The initial results from denaturing PAGE were difficult to interpret because of the poor binding of the SYBR gold dye to the PNA-DNA duplexes or the cross-linked products. The RP-HPLC and MALDI-TOF MS were next implemented to investigate the cross-linking reactions. The results showed that all three furan building blocks (f, F, and T) of the aegPNA and acpcPNA could undergo the cross-linking reactions when placed at the terminal positions. All furan building blocks on the PNA strand readily form cross-linked products with the DNA targets, albeit with different efficiencies among different DNA targets and different PNA probes, as well as different type of modifications.

According to the structure of the furan monomer in **Figure 3.61**, only the f-building block of both acpcPNA and aegPNA are exactly the same, but other monomers are slightly different. Thus, a fair comparison of the cross-linking efficiency could only be made for this f-building block. Concerning the F-building block, furan moiety of the aegPNA was attached on the backbone as the same position of the nucleobase. In the case of acpcPNA, the furan moiety was modified on the APC-spacer part which was different from the F-building block in the aegPNA. These different structures might provide a different furan-orientation when they formed their duplex with DNA target. The T-building block of acpcPNA was almost the same as the F-building block except for the presence of an extra T-base which might affect the selectivity of the cross-linking reaction. However, aegPNA with the T-building block was different from its F-building block and different from other furan building

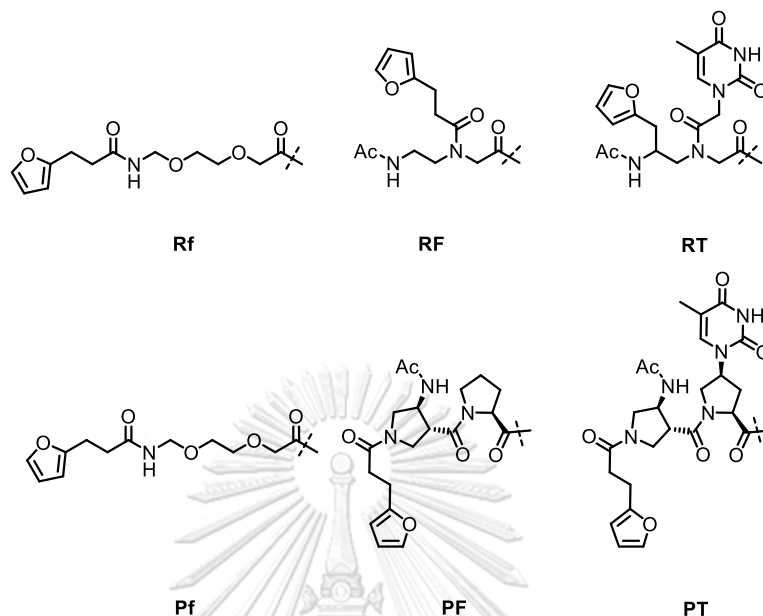blocks. Because the linker of the furan moiety was very short and directly attached on the aegPNA backbone.

Figure 3.61 Structures of the three furan building blocks used to modified on aegPNA (R) and acpcPNA (P)

For the cross-linking studies of terminally furan-modified PNAs (section **3.2.3**), the cross-linking efficiencies towards different nucleobases depend on the type of that allows the furan part to react with nucleobases that are further away. At the same time, the longer spacer might allow the folding back of the furan part to react with the closer nucleobase. The results revealed that aegPNA with the f-building block could undergo the cross-linking reaction with DNA-C and DNA-T (via adjacent C-base) in both one-base and five-base at the overhang region. While acpcPNA with the same f-building block selectively cross-liked with only DNA-C with both one and five base(s) at the overhang region. The f-modification in acpcPNA was the same as the one employed in section I and the results from this and from the experiments in section I were quite consistent.

Concerning the F-building block, the absence of the nucleobase on the furan-modified monomer suggest that it should cross-link with all nucleobases with exocyclic amino groups indiscriminately. Also, the flexibility of the linker to the F

monomer should be less than the f-monomer. Thus, both aegPNA and acpcPNA were found to cross-link with A, C, and G (as well as T, via the adjacent T base in the case of acpcPNA).

In the case of T-building block which differ from the F-building block by only the presence of an additional T-base, the cross-linking reaction was expected to be disfavored when there was adenine as the complementary DNA base. There was only one instance where this is true – when the target is DNA-A. Nonetheless, it appeared that both aegPNA and acpcPNA with the T-building block can form cross-links with most DNA targets similar to the F-building block. Evidently, the terminal A base might not efficiently pair with the T-building block otherwise the cross-link would not form. Also, the cross-linking involving the T-building block showed marginally higher efficiency than the F-building block. The results might be attributed to the ability of the T-base to form a base stack without forming a base pair with the target nucleobase. This in turns organizes the furan moiety in the T-building block, leading to a more efficient cross-linking reaction.

In all cases, aegPNA appears to show better cross-linking efficiency than acpcPNA, but not much. The selectivity towards different nucleobases are also slightly different. For the F- and T-building blocks, the difference in the selectivity and the cross-linking efficiency cannot be attributed to the different structures of the two types of PNA alone since the F and T building blocks in the acpcPNA and aegPNA were also different. Nevertheless, all the results suggested that the structure of the furan on the PNA strand affected the cross-linking efficiency and selectivity. Thus, the optimizations of the length of the linker, the position of the furan on the PNA strand, and the position of the furan moiety in the duplex are expected to improve the cross-link efficiency and selectivity further.

Regarding to the internally-furan-modified PNA, the cross-linking reactions were evaluated by RP-HPLC and denaturing PAGE. Only acpcPNA with the F and T building blocks were included in the studies. Unfortunately, HPLC could not detect the acpcPNA peaks and their cross-linked products. Nevertheless, the cross-linking

reactions were detectable, albeit in a more qualitative fashion, by denaturing PAGE and MALDI-TOF MS. The results revealed that the cross-linking reactions could occur when a free C-base was present at the opposite or adjacent position to the furan building block (F and T). The F-building block generally showed higher cross-linking efficiencies comparing to T-building block due to the inability of the abasic F-building block to form base pairing with the opposite nucleobase. The acpcPNAs with T building block failed to show any cross-linking when the fully complementary target DNA in all sequence contexts, which confirmed the previous results in section I. This also confirms that the presence of unpaired nucleobase (C>A>>G) in the DNA strand located near the furan site on the PNA strand is essential for the successful cross-linking.

The overall lower cross-inking efficiency of acpcPNA might be the result of the higher duplex stability of acpcPNA when compared to aegPNA, making the tighter base pair that in fact hindering rather than improving the cross-link formation. Moreover, the behavior of the two PNAs might be different, and each might require different optimal conditions for both the cross-linking and detection. Further optimization of these conditions could have been performed. These include: determining the cross-linking reaction, especially in the case of internal cross-link, from the same condition as in Ghent University but using the HPLC condition developed at Chulalongkorn University; determining the PAGE's result by UV shadowing instead of SYBR gold which is difficult to quantitate; optimizing the cross-linking reactions by changing the equivalents of NBS, temperature, incubation time and the way the duplex is being prepared. Unfortunately, the work in this section was not completed according to the originally proposed plan due to the global pandemic situation of COVID-19.

When compared with other cross-linkable PNA probes, the present furan-modified acpcPNA probes are still efficient in the context of PNA-DNA cross-linking. There are only two published papers from the Madder's group about the furan-modified aegPNA to cross-link with DNA.[100, 101] The cross-linking efficiency was very low according to the PAGE's result of internally furan-modified aegPNA and no

%yield has been reported. The cross-linking involving terminally furan-modified PNA probes has not been reported yet. A few studies of the aegPNA with other cross-linking agents have been reported. For instance, Nagatsugi's group have been studied the modification of 4-amino-6-oxo-2-vinylpyrimidine (AOVP)[116] and 2-amino-6-vinylpurine (AVP)[117] at the middle of the aegPNA strand used as a cross-linkable probes. Both cross-linking agents induced the complementary base to cross-link via H-bond to lock conformation of the base to react with. The AOVP selectively cross-linked with T-base (using the oxygen atom at C4 position) in 42% yield. In the same manner, the AVP selectively cross-linked with T-base of the DNA target in 52% yield and U-base of RNA target in 62% yield. Another example was the phenylselenide-modified thymine as a cross-linking monomer to modify on the aegPNA at the middle of the strand targeting complementary DNA.[118] The cross-link occurred by an oxidation reaction of phenylselenide molecule by 350 nm light and/or $NaIO_4$. The recognition of the base paring between the phenylselenide-modified thymine and A-base providing the selectivity of this probe towards A (31%) rather than C (16%). The nucleobase target for different type of cross-linking is distinctively different. furan modified PNA is the method of choices to selectively cross-link with C > A >>G. While the AOVP and AVP modified PNA are suitable to selectively cross-link with T or U base. On the other hand, the phenylselenide-modified PNA is selectively cross-link with A > C. Thus, these different cross-linking strategies are complementary rather than competing.

CHAPTER IV

CONCLUSIONS

The dissertation aimed to develop a new activate-on-demand cross-linkable probe that could react with the nucleobase of the DNA strand via a covalent bond. Furan was selected as a modifier on the acpcPNA strand, which can be activated by treatment with NBS. In section I, the cross-linking studies were performed on the furan-modified acpcPNA targeting complementary DNA. The furan modifier was incorporated into acpcPNA in the middle [PNA(I)] and the end [PNA(T)] of the strand to study the effect of the furan position towards cross-linking efficiency. The PNA(T) could undergo cross-linking reaction with the C-base in the overhang region of the DNA target (DNA-C5) in a very high yield (71%) followed by the DNA-A5 (22%) and DNA-G5 (6%), respectively. Thus, the reactivity of the PNA(T) towards nucleobase can be ranked as C >> A > G, and no cross-linking was observed for T. Under the same conditions, the PNA(I) could not undergo cross-linking reaction with the same set of DNA targets. The result could be explained by the formation of the Watson-Crick base pairing of all DNA bases in the case of internally furan-modified PNA, making them unavailable to react with the activated furan species. The insertion of a C base in the DNA strand was performed to verify the hypothesis that the availability of unpaired nucleobase can improve the cross-linking efficiency. As expected, the cross-linked product could form with the DNA-L6 that carries a C-insertion close to the furan modification site. It was observed that the activated furan-modified PNA probe could rapidly degrade itself, unless when it forms a stable duplex with a complementary DNA strand regardless of the ability to form cross-linked products. This phenomenon might be beneficial to reduce the side reactions due to non-specific cross-linking reactions.

In section II, the cross-linking studies were carried out with both acpcPNA and aegPNA in order to compare their cross-linking efficiencies side-by-side. The studies were performed with furan modification both at the end (acpcPNA and aegPNA) and the middle (acpcPNA only) of the PNA strand. Three furan-modified building blocks (f,

F, T) were used for the terminal modification, and two furan-modified building blocks (F, T) were employed for the internal modification. The DNA targets for the terminal-modified PNA probes include those carrying one- and five-base overhangs (A, T, C, and G). The denaturing PAGE results were difficult to interpret due to the poor binding of the SYBR gold dye to the PNA-DNA duplexes and/or the cross-linked products. The RP-HPLC and MALDI-TOF were next implemented to investigate the cross-linking reactions. The results revealed that all three furan building blocks of both acpcPNA and aegPNA could undergo cross-linking reaction with the complementary DNA. It was found that aegPNA with the f-building block could undergo the cross-linking reaction with DNA-C and DNA-T (via adjacent C-base). While the acpcPNA with the same f-building block selectively cross-liked only with DNA-C. The abasic building block F and the thymine-containing building block T in aegPNA and acpcPNA could undergo cross-linking reaction with all DNA bases carrying exocyclic amino groups including A, C, and G. In all cases, the cross-linking efficiency of the aegPNA was higher than acpcPNA. The overall lower cross-linking efficiency of acpcPNA might be the result of the higher duplex stability of acpcPNA when compared to aegPNA, making the tighter base pair that hindering rather than improving the cross-link formation. To solve the poor binding of SYBR-gold, 7-base complementary DNA targets were included as targets for the cross-linking experiments. It was found that decreasing the number of complementary DNA bases not only improved the visualization of the PNA/DNA by PAGE but also improved the cross-linking efficiency. This unexpected result could be explained by the formation of the imperfect duplex which might affect the PNA-DNA duplex orientation which allows the furan moiety in the suitable position to react with the overhanging nucleobase.

Concerning the furan-modification at the middle of the strand, only acpcPNA probes with the F (abasic furan) and T (T+furan) building blocks were included in the studies. Unfortunately, the RP-HPLC chromatogram did not show the cross-linked product and the acpcPNA. The PAGE and MALDI-TOF MS were next implemented to investigate the cross-linking reactions. The results revealed that both the F and T building blocks could undergo the cross-linking reaction most efficiently when there

was a free C-base present at or adjacent to the opposite position of the furan building block. It was found that the abasic F-building block showed higher cross-linking efficiencies comparing to the T-building block. The results could be explained by the inability of the abasic F-building block to form a base pairing with the opposite nucleobase, making the nucleobase available to react with the furan building block. In the case of the T-building block, its cross-linking reaction failed with the fully complementary DNA target in all sequence contexts, but the cross-linking was possible when there was one or more C-mismatched at or adjacent to the T building block. These results confirmed that the essential factor for the successful cross-linking reaction was the available nucleobase (C>A>>G) in the DNA strand located near the furan site on the PNA strand.

This dissertation has successfully demonstrated the synthesis and cross-linking behavior of furan-modified acpcPNA probes. Comparison of the terminal and internal modifications revealed the reactivity patterns that lead to the conclusion that both the ability to form a stable duplex and the availability of the unpaired DNA nucleobase with an exocyclic amino group (A, C, G) were essential for the cross-linking reaction, which provides further insights into the requirements for an efficient ICL formation. Under the optimal conditions, reasonably good cross-linking efficiencies (up to 71%) were observed, and C was the base that most readily formed cross-linked with the activated furan-modified PNA probes.

.

REFERENCES

1.   Setlow, R. B., Cyclobutane-Type Pyrimidine Dimers in Polynucleotides. *Science* **1966,** *153*, 379-386.

2.   Deans, A. J.; West, S. C., DNA Interstrand Crosslink Repair and Cancer. *Nat. Rev. Cancer* **2011,** *11*, 467-480.

3.   Gilman, A.; Philips, F. S., The Biological Actions and Therapeutic Applications of the B-Chloroethyl Amines and Sulfides. *Science* **1946,** *103*, 409-436.

4.   Huang, Y.; Li, L., DNA Crosslinking Damage and Cancer - a Tale of Friend and Foe. *Transl Cancer Res* **2013,** *2*, 144-154.

5.   Freifelder, D.; Davison, P. F.; Geiduschek, E. P., Damage by Visible Light to the Acridine Orange-DNA Complex. *Biophys. J.* **1961,** *1*, 389-400.

6.   Smith, K. C.; Hodgkins, B.; O'Leary, M. E., The Biological Importance of Ultraviolet Light Induced DNA-protein Crosslinks in Escherichia coli 15 TAU. *Biochim. Biophys. Acta, Nucleic Acids Protein Synth.* **1966,** *114*, 1-15.

7.   Quievryn, G.; Zhitkovich, A., Loss of DNA–protein Crosslinks from Formaldehyde-exposed Cells Occurs Through Spontaneous Hydrolysis and an Active Repair Process Linked to Proteosome Function. *Carcinogenesis* **2000,** *21*, 1573-1580.

8.   Chakrabarti, S. K.; Bai, C.; Subramanian, K. S., DNA–Protein Crosslinks Induced by Nickel Compounds in Isolated Rat Renal Cortical Cells and Its Antagonism by Specific Amino Acids and Magnesium Ion. *Toxicol. Appl. Pharmacol.* **1999,** *154*, 245-255.

9.   Al-Nabulsi, I.; Wheeler, K. T., Temperature Dependence of Radiation-Induced DNA-Protein Crosslinks Formed under Hypoxic Conditions. *Radiat. Res.* **1997,** *148*, 568-574.

10.   Noll, D. M.; Mason, T. M.; Miller, P. S., Formation and Repair of Interstrand Cross-Links in DNA. *Chem. Rev.* **2006,** *106*, 277-301.

11.   Chen, W.; Balakrishnan, K.; Kuang, Y.; Han, Y.; Fu, M.; Gandhi, V.; Peng, X., Reactive Oxygen Species (ROS) Inducible DNA Cross-Linking Agents and Their Effect on Cancer Cells and Normal Lymphocytes. *J. Med. Chem.* **2014,** *57*, 4498-4510.

12.   Wu, J.; Huang, R.; Wang, T.; Zhao, X.; Zhang, W.; Weng, X.; Tian, T.; Zhou, X.,

Fluoride as an Inducible DNA Cross-linking Agent for New Antitumor Prodrug. *Org. Biomol. Chem.* **2013,** *11,* 2365-2369.

13.   Kuang, Y.; Balakrishnan, K.; Gandhi, V.; Peng, X., Hydrogen Peroxide Inducible DNA Cross-Linking Agents: Targeted Anticancer Prodrugs. *J. Am. Chem. Soc.* **2011,** *133,* 19278-19281.

14.   Sheng, C.; Xiaohua, P., Exploiting Endogenous Cellular Process to Generate Quinone Methides In Vivo. *Curr. Org. Chem.* **2014,** *18,* 70-85.

15.   Peng, X.; Greenberg, M. M., Facile SNP Detection Using Bifunctional, Cross-Linking Oligonucleotide Probes. *Nucleic Acids Res.* **2008,** *36,* e31.

16.   Peng, X.; Pigli, Y. Z.; Rice, P. A.; Greenberg, M. M., Protein Binding Has a Large Effect on Radical Mediated DNA Damage. *J. Am. Chem. Soc.* **2008,** *130,* 12890-12891.

17.   Haque, M. M.; Sun, H.; Liu, S.; Wang, Y.; Peng, X., Photoswitchable Formation of a DNA Interstrand Cross-Link by a Coumarin-Modified Nucleotide. *Angew. Chem. Int. Ed.* **2014,** *53,* 7001-7005.

18.   Kashida, H.; Doi, T.; Sakakibara, T.; Hayashi, T.; Asanuma, H., p-Stilbazole Moieties as Artificial Base Pairs for Photo-Cross-Linking of DNA Duplex. *J. Am. Chem. Soc.* **2013,** *135,* 7960-7966.

19.   Rusling, D. A.; Nandhakumar, I. S.; Brown, T.; Fox, K. R., Triplex-directed Covalent Cross-linking of a DNA Nanostructure. *Chem. Commun.* **2012,** *48,* 9592-9594.

20.   Rajendran, A.; Endo, M.; Katsuda, Y.; Hidaka, K.; Sugiyama, H., Photo-Cross-Linking-Assisted Thermal Stability of DNA Origami Structures and Its Application for Higher-Temperature Self-Assembly. *J. Am. Chem. Soc.* **2011,** *133,* 14488-14491.

21.   Povirk, L. F.; Shuker, D. E., DNA Damage and Mutagenesis Induced by Nitrogen Mustards. *Mutat. Res. Rev. Genet. Toxicol.* **1994,** *318,* 205-226.

22.   Rai, K. R.; Peterson, B. l.; Appelbaum, F. R.; Kolitz, J.; Elias, L.; Shepherd, L.; Hines, J.; Threatte, G. A.; Larson, R. A.; Cheson, B. D.; Schiffer, C. A., Fludarabine Compared with Chlorambucil as Primary Therapy for Chronic Lymphocytic Leukemia. *N. Engl. J. Med.* **2000,** *343,* 1750-1757.

23.   Bayraktar, U. D.; Bashir, Q.; Qazilbash, M.; Champlin, R. E.; Ciurea, S. O., Fifty Years of Melphalan Use in Hematopoietic Stem Cell Transplantation. *Biol. Blood Marrow Transplant.* **2013,** *19,* 344-356.

24.   Leoni, L. M.; Bailey, B.; Reifert, J.; Reifert, J.; Bendall, H. H.; Zeller, R. W.; Zeller, R. W.; Corbeil, J.; Elliott, G.; Niemeyer, C. C., Bendamustine (Treanda) Displays a Distinct Pattern of Cytotoxicity and Unique Mechanistic Features Compared with other Alkylating Agents. *Clin. Cancer Res.* **2008,** *14*, 309-317.

25.   Newton, H. B., CHAPTER 2 - Clinical Pharmacology of Brain Tumor Chemotherapy. In *Handbook of Brain Tumor Chemotherapy*, Newton, H. B., Ed. Academic Press: San Diego, 2006; pp 21-43.

26.   Florea, A.-M.; Büsselberg, D., Cisplatin as an Anti-Tumor Drug: Cellular Mechanisms of Activity, Drug Resistance and Induced Side Effects. *Cancers* **2011,** *3*, 1351–1371.

27.   Serrano-Pérez, J. J.; Merchán, M.; Serrano-Andrés, L., Photoreactivity of Furocoumarins and DNA in PUVA Therapy: Formation of Psoralen–Thymine Adducts. *J. Phys. Chem. B* **2008,** *112*, 14002-14010.

28.   Querfeld, C.; Rosen, S. T.; Kuzel, T. M.; Kirby, K. A.; Roenigk, H. H.; Prinz, B. M.; Guitart, J., Long-term Follow-up of Patients with Eearly-stage Cutaneous T-cell Lymphoma Who Achieved Complete Remission with Psoralen Plus UV-A Monotherapy. *Arch. Dermatol.* **2005,** *141*, 305-311.

29.   Khosrow Momtaz, T.; Fitzpatrick, T. B., The Benefits and Risks of Long-Term Puva Photochemotherapy. *Dermatol. Clin.* **1998,** *16*, 227-234.

30.   Verweij, J.; Pinedo, H. M., Mitomycin C: Mechanism of Action, Usefulness and Limitations. *Anticancer Drugs* **1990,** *1*, 5-13.

31.   Tomasz, M., Mitomycin C: Small, Fast and Deadly (but very Selective). *Chem. Biol.* **1995,** *2*, 575-579.

32.   Nagatsugi, F.; Kawasaki, T.; Usui, D.; Maeda, M.; Sasaki, S., Highly Efficient and Selective Cross-Linking to Cytidine Based on a New Strategy for Auto-Activation within a Duplex. *J. Am. Chem. Soc.* **1999,** *121*, 6753-6754.

33.   Imoto, S.; Chikuni, T.; Kansui, H.; Kunieda, T.; Nagatsugi, F., Fast DNA Interstrand Cross-linking Reaction by 6-vinylpurine. *Nucleosides Nucleotides Nucl. Acids* **2012,** *31*, 752-762.

34.   Hong, I. S.; Greenberg, M. M., DNA Interstrand Cross-Link Formation Initiated by Reaction between Singlet Oxygen and a Modified Nucleotide. *J. Am. Chem. Soc.* **2005,**

127, 10510-10511.

35.   Peng, X.; Hong, I. S.; Li, H.; Seidman, M. M.; Greenberg, M. M., Interstrand Cross-Link Formation in Duplex and Triplex DNA by Modified Pyrimidines. *J. Am. Chem. Soc.* **2008,** *130*, 10299-10306.

36.   Sloane, J. L.; Greenberg, M. M., Interstrand Cross-link and Bioconjugate Formation in RNA from a Modified Nucleotide. *J. Org. Chem.* **2014,** *79*, 9792-9798.

37.   Du, Y.; Weng, X.; Huang, J.; Zhang, D.; Ma, H.; Chen, D.; Zhou, X.; Constant, J.-F., Oligonucleotide-selenide Conjugate: Synthesis and Its Inducible Sequence-specific Alkylation of DNA. *Biorg. Med. Chem.* **2010,** *18*, 4149-4153.

38.   Cimino, G. D.; Shi, Y. B.; Hearst, J. E., Wavelength Dependence for the Photoreversal of a Psoralen-DNA Crosslink. *Biochemistry* **1986,** *25*, 3013-3020.

39.   Higuchi, M.; Yamayoshi, A.; Yamaguchi, T.; Iwase, R.; Yamaoka, T.; Kobori, A.; Murakami, A., Selective Photo-Cross-Linking of 2′-O-Psoralen-Conjugated Oligonucleotide with Rnas Having Point Mutations. *Nucleosides Nucleotides Nucl. Acids* **2007,** *26*, 277-290.

40.   Fujimoto, K.; Toyosato, K.; Nakamura, S.; Sakamoto, T., RNA Fluorescence in Situ Hybridization using 3-Cyanovinylcarbazole Modified Oligodeoxyribonucleotides as Photo-Cross-Linkable Probes. *Bioorg. Med. Chem. Lett.* **2016,** *26*, 5312-5314.

41.   Fujimoto, K.; Yamada, A.; Yoshimura, Y.; Tsukaguchi, T.; Sakamoto, T., Details of the Ultrafast DNA Photo-Cross-Linking Reaction of 3-Cyanovinylcarbazole Nucleoside: Cis–Trans Isomeric Effect and the Application for SNP-Based Genotyping. *J. Am. Chem. Soc.* **2013,** *135*, 16161-16167.

42.   Yoshimura, Y.; Ohtake, T.; Okada, H.; Fujimoto, K., A New Approach for Reversible RNA Photocrosslinking Reaction: Application to Sequence-Specific RNA Selection. *ChemBioChem* **2009,** *10*, 1473-1476.

43.   Fujimoto, K.; Sasago, S.; Mihara, J.; Nakamura, S., DNA Photo-cross-linking using Pyranocarbazole and Visible Light. *Org. Lett.* **2018,** *20*, 2802-2805.

44.   Doi, T.; Kawai, H.; Murayama, K.; Kashida, H.; Asanuma, H., Visible-Light-Triggered Cross-Linking of DNA Duplexes by Reversible [2+2] Photocycloaddition of Styrylpyrene. *Chem. Eur. J.* **2016,** *22*, 10533-10538.

45.   Murayama, K.; Yamano, Y.; Asanuma, H., 8-Pyrenylvinyl Adenine Controls

Reversible Duplex Formation between Serinol Nucleic Acid and RNA by [2 + 2] Photocycloaddition. *J. Am. Chem. Soc.* **2019,** *141,* 9485-9489.

46.	Yamano, Y.; Murayama, K.; Asanuma, H., Dual Crosslinking Photo-switches for Orthogonal Photo-control of Hybridization between Serinol Nucleic Acid and RNA. *Chem. Eur. J.* **2020,** doi:10.1002/chem.202003528.

47.	Sirica, A. E., Biliary Proliferation and Adaptation in Furan-Induced Rat Liver Injury and Carcinogenesis. *Toxicol. Pathol.* **1996,** *24,* 90-99.

48.	Grill, A. E.; Schmitt, T.; Gates, L. A.; Lu, D.; Bandyopadhyay, D.; Yuan, J.-M.; Murphy, S. E.; Peterson, L. A., Abundant Rodent Furan-Derived Urinary Metabolites are Associated with Tobacco Smoke Exposure in Humans. *Chem. Res. Toxicol.* **2015,** *28,* 1508-1516.

49.	Wang, E.; Chen, F.; Hu, X.; Yuan, Y., Protective Effects of Apigenin Against Furan-induced Toxicity in Mice. *Food Funct.* **2014,** *5,* 1804-1812.

50.	Pennati, F. Design and Synthesis of Furan Modified Peptide Nucleic Acids (PNAs) for Cross-linking to ssDNA. Università degli Studi di Milano - Facoltà di Scienze del Farmaco, 2019-2020.

51.	Duszka, K.; Clark, B. F. C.; Massino, F.; Barciszewski, J., Biological Activities of Kinetin. In *Herbal Drugs: Ethnomedicine to Modern Medicine*, Ramawat, K. G., Ed. Springer Berlin Heidelberg: Berlin, Heidelberg, 2009; pp 369-380.

52.	Peterson, L. A.; Cummings, M. E.; Vu, C. C.; Matter, B. A., Glutathione Trapping to Measure Microsomal Oxidation of Furan to *cis*-2-butene-1,4-dial. *Drug Metab. Disposition* **2005,** *33,* 1453.

53.	Byrns, M. C.; Vu, C. C.; Neidigh, J. W.; Abad, J.-L.; Jones, R. A.; Peterson, L. A., Detection of DNA Adducts Derived from the Reactive Metabolite of Furan, Cis-2-butene-1,4-dial. *Chem. Res. Toxicol.* **2006,** *19,* 414-420.

54.	Phillips, M. B.; Sullivan, M. M.; Villalta, P. W.; Peterson, L. A., Covalent Modification of Cytochrome C by Reactive Metabolites of Furan. *Chem. Res. Toxicol.* **2014,** *27,* 129-135.

55.	Halila, S.; Velasco, T.; Clercq, P. D.; Madder, A., Fine-Tuning Furan Toxicity: Fast and Quantitative DNA Interchain Cross-Link Formation Upon Selective Oxidation of a Furan Containing Oligonucleotide. *Chem. Commun.* **2005,** 936-938.

56.   Carrette, L. L. G.; Morii, T.; Madder, A., Toxicity Inspired Cross-Linking for Probing DNA–Peptide Interactions. *Bioconjugate Chem.* **2013,** *24,* 2008-2014.

57.   Op de Beeck, M.; Madder, A., Unprecedented C-Selective Interstrand Cross-Linking through in Situ Oxidation of Furan-Modified Oligodeoxynucleotides. *J. Am. Chem. Soc.* **2011,** *133,* 796-807.

58.   Jawalekar, A. M.; Op de Beeck, M.; van Delft, F. L.; Madder, A., Synthesis and Incorporation of a Furan-Modified Adenosine Building Block for DNA Interstrand Crosslinking. *Chem. Commun.* **2011,** *47,* 2796-2798.

59.   Stevens, K.; Madder, A., Furan-modified Oligonucleotides for Fast, High-yielding and Site-selective DNA Inter-strand Cross-linking with Non-modified Complements. *Nucleic Acids Res.* **2009,** *37,* 1555-1565.

60.   Carrette, L. L. G.; Gyssels, E.; De Laet, N.; Madder, A., Furan Oxidation Based Cross-Linking: A New Approach for the Study and Targeting of Nucleic Acid and Protein Interactions. *Chem. Commun.* **2016,** *52,* 1539-1554.

61.   Carrette, L. L. G.; Gyssels, E.; Loncke, J.; Madder, A., A Mildly Inducible and Selective Cross-Link Methodology for RNA Duplexes. *Org. Biomol. Chem.* **2014,** *12,* 931-935.

62.   Op de Beeck, M.; Madder, A., Sequence Specific DNA Cross-Linking Triggered by Visible Light. *J. Am. Chem. Soc.* **2012,** *134,* 10737-10740.

63.   Gyssels, E.; Carrette, L. L.; Vercruysse, E.; Stevens, K.; Madder, A., Triplex Crosslinking Through Furan Oxidation Requires Perturbation of the Structured Triple-Helix. *ChemBioChem* **2015,** *16,* 651-658.

64.   Veliz Montes, C.; Memczak, H.; Gyssels, E.; Torres, T.; Madder, A.; Schneider, R. J., Photoinduced Cross-Linking of Short Furan-Modified DNA on Surfaces. *Langmuir* **2017,** *33,* 1197-1201.

65.   Galyuk, E. N.; Lando, D. Y.; Egorova, V. P.; Dai, H.; Dosin, Y. M., Na_2CO_3 Influence on DNA Double Helix Stability: Strong Anion Destabilizing Effect. *J. Biomol. Struct. Dyn.* **2003,** *20,* 801-809.

66.   Kolb, H. C.; Finn, M. G.; Sharpless, K. B., Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angew. Chem. Int. Ed.* **2001,** *40,* 2004-2021.

67.   Kocalka, P.; El-Sagheer, A. H.; Brown, T., Rapid and Efficient DNA Strand Cross-

Linking by Click Chemistry. *ChemBioChem* **2008,** *9*, 1280-1285.

68.	Pujari, S. S.; Xiong, H.; Seela, F., Cross-Linked DNA Generated by "Bis-click" Reactions with Bis-functional Azides: Site Independent Ligation of Oligonucleotides via Nucleobase Alkynyl Chains. *J. Org. Chem.* **2010,** *75*, 8693-8696.

69.	Pujari, S. S.; Seela, F., Parallel Stranded DNA Stabilized with Internal Sugar Cross-Links: Synthesis and Click Ligation of Oligonucleotides Containing 2′-Propargylated Isoguanosine. *J. Org. Chem.* **2013,** *78*, 8545-8561.

70.	Pujari, S. S.; Seela, F., Cross-Linked DNA: Propargylated Ribonucleosides as "Click" Ligation Sites for Bifunctional Azides. *J. Org. Chem.* **2012,** *77*, 4460-4465.

71.	Shimada, K.; Crother, T. R.; Arditi, M., Chapter 10 - DNA Damage Responses in Atherosclerosis. In *Biological DNA Sensor*, Ishii, K. J.; Tang, C. K., Eds. Academic Press: Amsterdam, 2014; pp 231-253.

72.	Lehoczký, P.; McHugh, P. J.; Chovanec, M., DNA Interstrand Cross-Link Repair in Saccharomyces Cerevisiae. *FEMS Microbiol. Rev.* **2007,** *31*, 109-133.

73.	McHugh, P. J.; Spanswick, V. J.; Hartley, J. A., Repair of DNA Interstrand Crosslinks: Molecular Mechanisms and Clinical Relevance. *The Lancet Oncology* **2001,** *2*, 483-490.

74.	Niedernhofer, L. J.; Lalai, A. S.; Hoeijmakers, J. H. J., Fanconi Anemia (Cross)linked to DNA Repair. *Cell* **2005,** *123*, 1191-1198.

75.	Akkari, Y. M. N.; Bateman, R. L.; Reifsteck, C. A.; Olson, S. B.; Grompe, M., DNA Replication is Required to Elicit Cellular Responses to Psoralen-induced DNA Interstrand Cross-Links. *Mol. Cell. Biol.* **2000,** *20*, 8283-8289.

76.	Rothfuss, A.; Grompe, M., Repair Kinetics of Genomic Interstrand DNA Cross-Links: Evidence for DNA Double-Strand Break-Dependent Activation of the Fanconi Anemia/BRCA Pathway. *Mol. Cell. Biol.* **2004,** *24*, 123-134.

77.	Räschle, M.; Knipscheer, P.; Enoiu, M.; Angelov, T.; Sun, J.; Griffith, J. D.; Ellenberger, T. E.; Schärer, O. D.; Walter, J. C., Mechanism of Replication-Coupled DNA Interstrand Crosslink Repair. *Cell* **2008,** *134*, 969-980.

78.	Hashimoto, S.; Anai, H.; Hanada, K., Mechanisms of Interstrand DNA Crosslink Repair and Human Disorders. *Genes Environ.* **2016,** *38*, 9.

79.	Brulikova, L.; Hlavac, J.; Hradil, P., DNA Interstrand Cross-Linking Agents and their

Chemotherapeutic Potential. *Curr. Med. Chem.* **2012,** *19*, 364-385.

80.   Rycenga, H. B.; Long, D. T., The Evolving Role of DNA Inter-strand Crosslinks in Chemotherapy. *Curr. Opin. Pharmacol.* **2018,** *41*, 20-26.

81.   Kartalou, M.; Samson, L. D.; Essigmann, J. M., Cisplatin Adducts Inhibit 1,N6-Ethenoadenine Repair by Interacting with the Human 3-Methyladenine DNA Glycosylase. *Biochemistry* **2000,** *39*, 8032-8038.

82.   Wilson, D. M., 3rd; Seidman, M. M., A Novel Link to Base Excision Repair? *Trends Biochem. Sci* **2010,** *35*, 247-252.

83.   Robertson, A. B.; Klungland, A.; Rognes, T.; Leiros, I., DNA Repair in Mammalian Cells: Base Excision Repair: the Long and Short of It. *Cell. Mol. Life Sci.* **2009,** *66*, 981-993.

84.   Wang, P.; Liu, R.; Wu, X.; Ma, H.; Cao, X.; Zhou, P.; Zhang, J.; Weng, X.; Zhang, X.-L.; Qi, J.; Zhou, X.; Weng, L., A Potent, Water-Soluble and Photoinducible DNA Cross-Linking Agent. *J. Am. Chem. Soc.* **2003,** *125*, 1116-1117.

85.   Nielsen, P.; Egholm, M., An Introduction to Peptide Nucleic Acid. *Curr. Issues Mol. Biol.* **1999,** *1*, 89-104.

86.   Nielsen, P. E., Peptide Nucleic Acid. A Molecule with Two Identities. *Acc. Chem. Res.* **1999,** *32*, 624-630.

87.   Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Norden, B.; Nielsen, P. E., PNA Hybridizes to Complementary Oligonucleotides Obeying the Watson–Crick Hydrogen-bonding Rules. *Nature* **1993,** *365*, 566-568.

88.   Nielsen, P. E., Peptide Nucleic Acid: A Versatile Tool in Genetic Diagnostics and Molecular Biology. *Curr. Opin. Biotechnol.* **2001,** *12*, 16-20.

89.   Nielsen, P. E.; Haaima, G., Peptide Nucleic Acid (PNA). A DNA Mimic with a Pseudopeptide Backbone. *Chem. Soc. Rev.* **1997,** *26*, 73-78.

90.   Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O., Sequence-Selective Recognition of DNA by Strand Displacement with a Thymine-Substituted Polyamide. *Science* **1991,** *254*, 1497-1500.

91.   Sugiyama, T.; Kittaka, A., Chiral Peptide Nucleic Acids with a Substituent in the *N*-(2-aminoethy)glycine Backbone. *Molecules* **2012,** *18*, 287-310.

92.	Kitamatsu, M.; Shigeyasu, M.; Okada, T.; Sisido, M., Oxy-peptide Nucleic Acid with a Pyrrolidine Ring that is Configurationally Optimized for Hybridization with DNA. *Chem. Commun.* **2004**, 1208-1209.

93.	Vilaivan, T.; Suparpprom, C.; Harnyuttanakorn, P.; Lowe, G., Synthesis and Properties of Novel Pyrrolidinyl PNA Carrying β-amino Acid Spacers. *Tetrahedron Lett.* **2001,** *42,* 5533-5536.

94.	D'Costa, M.; Kumar, V. A.; Ganesh, K. N., Aminoethylprolyl Peptide Nucleic Acids (aepPNA): Chiral PNA Analogues That Form Highly Stable DNA:aepPNA$_2$ Triplexes. *Org. Lett.* **1999,** *1,* 1513-1516.

95.	Lagriffoule, P.; Eriksson, M.; Jensen, K. K.; Nielsen, P. E.; Wittung, P.; Nordén, B.; Buchardt, O., Peptide Nucleic Acids with a Conformationally Constrained Chiral Cyclohexyl-Derived Backbone. *Chem. Eur. J.* **1997,** *3,* 912-919.

96.	Kosynkina, L.; Wang, W.; Liang, T. C., a Convenient Synthesis of Chiral Peptide Nucleic Acid (PNA) Monomers. *Tetrahedron Lett.* **1994,** *35,* 5173-5176.

97.	Suparpprom, C.; Srisuwannaket, C.; Sangvanich, P.; Vilaivan, T., Synthesis and Oligodeoxynucleotide Binding Properties of Pyrrolidinyl Peptide Nucleic Acids Bearing Prolyl-2-Aminocyclopentanecarboxylic Acid (Acpc) Backbones. *Tetrahedron Lett.* **2005,** *46,* 2833-2837.

98.	Vilaivan, T.; Lowe, G., A Novel Pyrrolidinyl PNA Showing High Sequence Specificity and Preferential Binding to DNA over RNA. *J. Am. Chem. Soc.* **2002,** *124,* 9326-9327.

99.	Vilaivan, T., Pyrrolidinyl PNA with α/β-Dipeptide Backbone: From Development to Applications. *Acc. Chem. Res.* **2015,** *48,* 1645-1656.

100.	Manicardi, A.; Gyssels, E.; Corradini, R.; Madder, A., Furan-PNA: A Mildly Inducible Irreversible Interstrand Crosslinking System Targeting Single and Double Stranded DNA. *Chem. Commun.* **2016,** *52,* 6930-6933.

101.	Elskens, J.; Manicardi, A.; Costi, V.; Madder, A.; Corradini, R., Synthesis and Improved Cross-Linking Properties of C5-Modified Furan Bearing PNAs. *Molecules* **2017,** *22,* 2010-2020.

102.	Ditmangklo, B.; Muangkaew, P.; Supabowornsathit, K.; Vilaivan, T., Synthesis of

Pyrrolidinyl PNA and Its Site-Specific Labeling at Internal Positions by Click Chemistry. In *Peptide Nucleic Acids: Methods and Protocols*, Nielsen, P. E., Ed. Springer US: New York, 2020; pp 35-60.

103.   Ditmangklo, B.; Taechalertpaisarn, J.; Siriwong, K.; Vilaivan, T., Clickable Styryl Dyes for Fluorescence Labeling of Pyrrolidinyl PNA Probes for the Detection of Base Mutations in DNA. *Org. Biomol. Chem.* **2019,** *17,* 9712-9725.

104.   Lee, K.-H.; Ko, K.-Y., Catalytic Oxidation of Benzophenone Hydrazone with Alumina-Supported KMnO4 under Oxygen Atmosphere. *Cheminform* **2006,** *27,* 185-186.

105.   Miller, J., Notes- Preparation of Crystalline Diphenyldiazomethane. *J. Org. Chem.* **1959,** *24,* 560-561.

106.   Vilaivan, T.; Srisuwannaket, C., Hybridization of Pyrrolidinyl Peptide Nucleic Acids and DNA: Selectivity, Base-Pairing Specificity, and Direction of Binding. *Org. Lett.* **2006,** *8,* 1897-1900.

107.   Huang, Z.-P.; Du, J.-T.; Zhao, Y.-F.; Li, Y.-M., Synthesis of Site-Specifically Dimethylated and Trimethylated Peptides Derived from Histone H3 *N*-Terminal Tail. *Int. J. Pept. Res. Ther.* **2006,** *12,* 187-193.

108.   Stevens, K.; Claeys, D. D.; Catak, S.; Figaroli, S.; Hocek, M.; Tromp, J. M.; Schürch, S.; Van Speybroeck, V.; Madder, A., Furan-Oxidation-Triggered Inducible DNA Cross-Linking: Acyclic Versus Cyclic Furan-Containing Building Blocks-On the Benefit of Restoring the Cyclic Sugar Backbone. *Chem. Eur. J.* **2011,** *17,* 6940-6953.

109.   Patchornik, A.; Lawson, W. B.; Gross, E.; Witkop, B., The Use of N-Bromosuccinimide and N-Bromoacetamide for the Selective Cleavage of C-Tryptophyl Peptide Bonds in Model Peptides and Glucagon1. *J. Am. Chem. Soc.* **1960,** *82,* 5923-5927.

110.   Carrette, L. L.; Gyssels, E.; Madder, A., DNA Interstrand Cross-Link Formation Using Furan as a Masked Reactive Aldehyde. *Curr. Protoc. Nucleic Acid Chem.* **2013,** *54,* 5.12.1-5.12.16.

111.   Darwish, I.; Hussein, A.; Mahmoud, A.; Hassan, A., *N*-Bromosuccinimide/Fluorescein System for Spectrophotometric Determination of H2-Receptor Antagonists in Their Dosage Forms. *Int. J. Res. Pharm. Sci.* **2010,** *1,* 86-93.

112. Carrette, L. L.; Madder, A., A Synthetic Oligonucleotide Model for Evaluating the Oxidation and Crosslinking Propensities of Natural Furan-Modified DNA. *ChemBioChem* **2014,** *15*, 103-107.

113. Jiao, Y.; Stringfellow, S.; Yu, H., Distinguishing "Looped-out" and "Stacked-in" DNA Bulge Conformation Using Fluorescent 2-Aminopurine Replacing a Purine Base. *J. Biomol. Struct. Dyn.* **2002,** *19*, 929-934.

114. Kolbeck, P. J.; Vanderlinden, W.; Nicolaus, T.; Gebhardt, C.; Cordes, T.; Lipfert, J., Intercalative DNA Binding Governs Fluorescence Enhancement of SYBR Gold. *bioRxiv* **2020**.

115. Wittung, P.; Kim, S. K.; Buchardt, O.; Nielsen, P.; Norden, B., Interactions of DNA Binding Ligands with PNA-DNA Hybrids. *Nucleic Acids Res.* **1994,** *22*, 5371-5377.

116. Akisawa, T.; Ishizawa, Y.; Nagatsugi, F., Synthesis of Peptide Nucleic Acids Containing a Crosslinking Agent and Evaluation of their Reactivities. *Molecules* **2015,** *20*, 4708-4719.

117. Akisawa, T.; Yamada, K.; Nagatsugi, F., Synthesis of Peptide Nucleic Acids (PNA) with a Crosslinking Agent to RNA and Effective Inhibition of Dicer. *Bioorg. Med. Chem. Lett.* **2016,** *26*, 5902-5906.

118. Kim, Y.; Hong, I. S., PNA/DNA Interstrand Cross-links from a Modified PNA Base upon Photolysis or Oxidative Conditions. *Bioorg. Med. Chem. Lett.* **2008,** *18*, 5054-5057.

APPENDIX
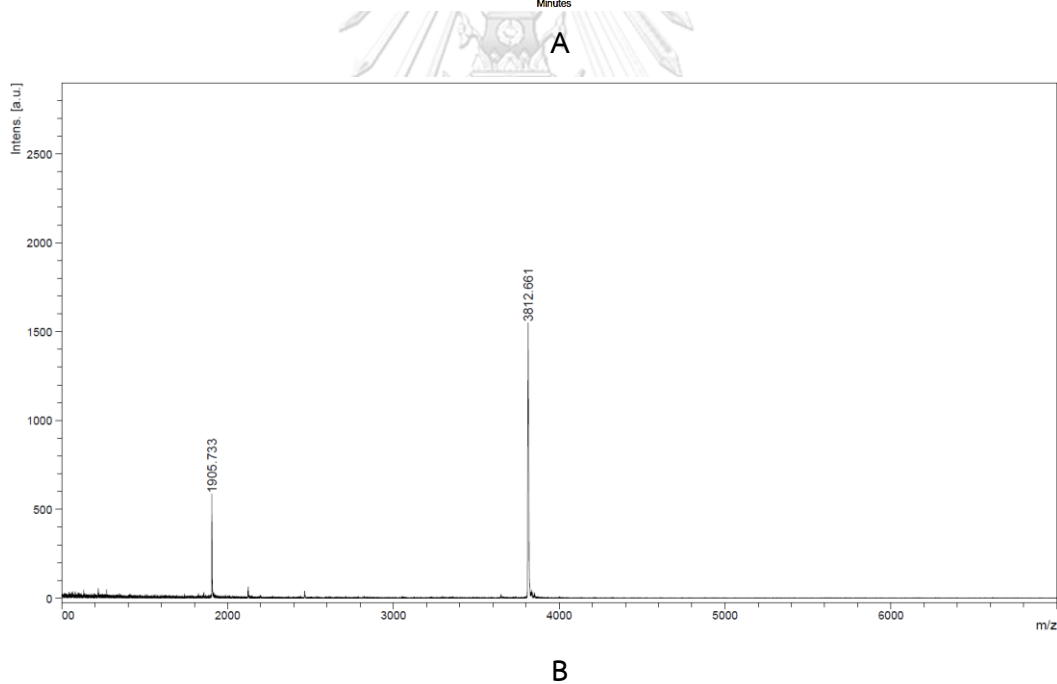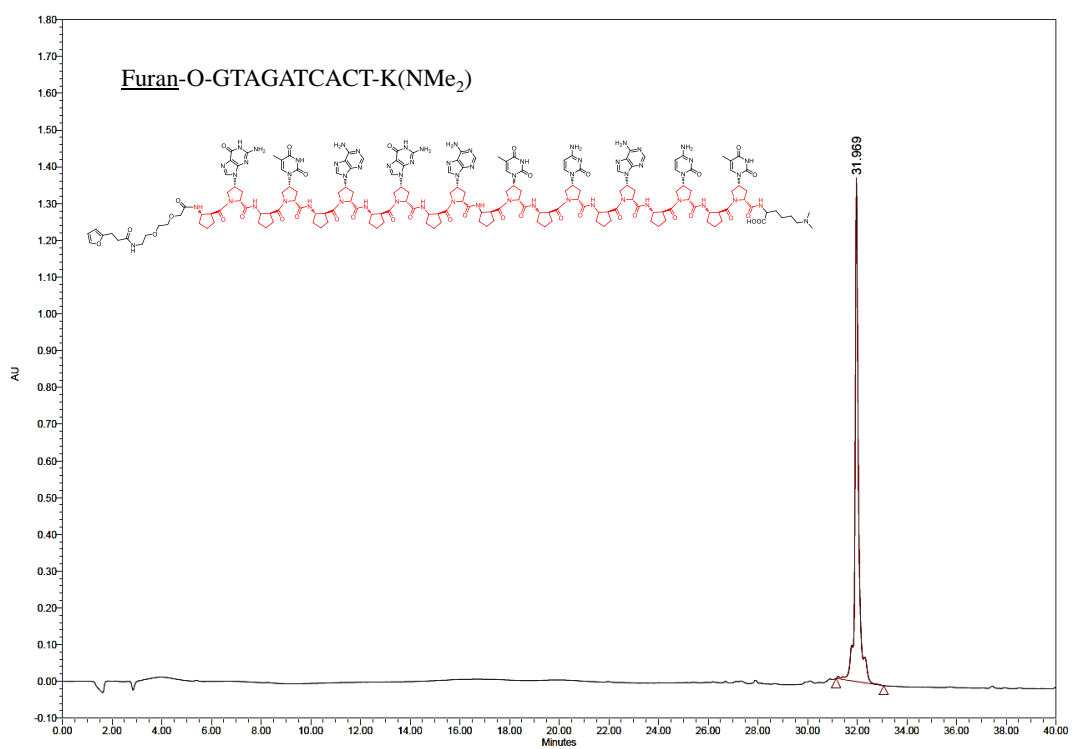
Figure A1 (A) Analytical HPLC chromatogram and (B) MALDI-TOF mass spectrum of PNA(T)
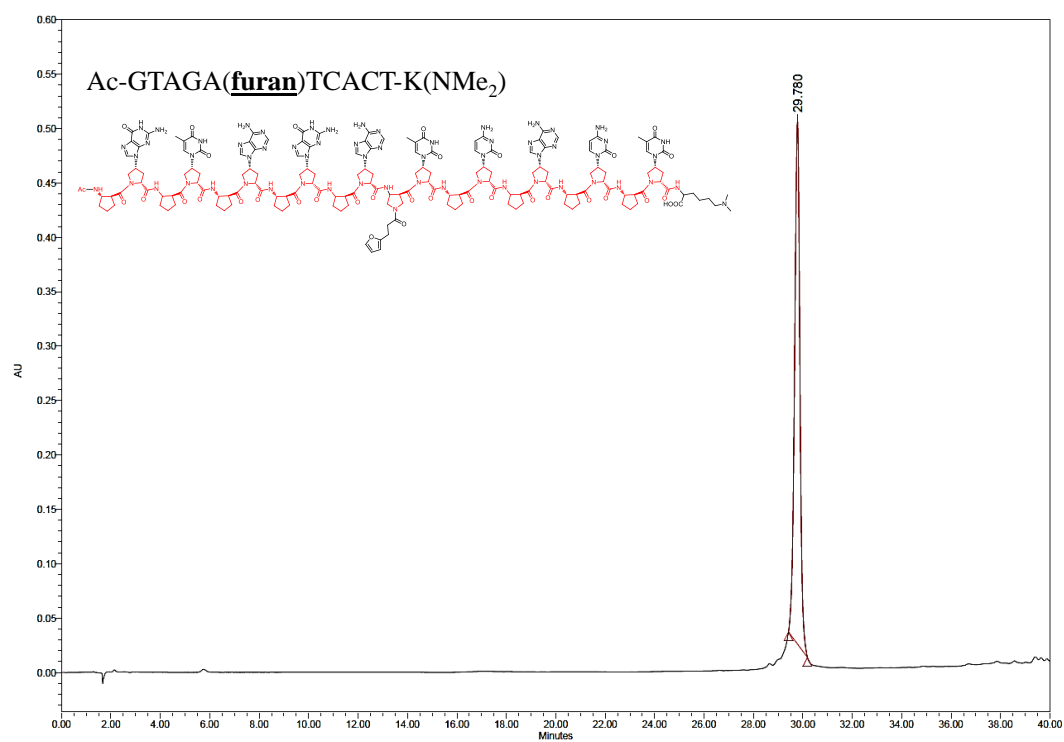
A
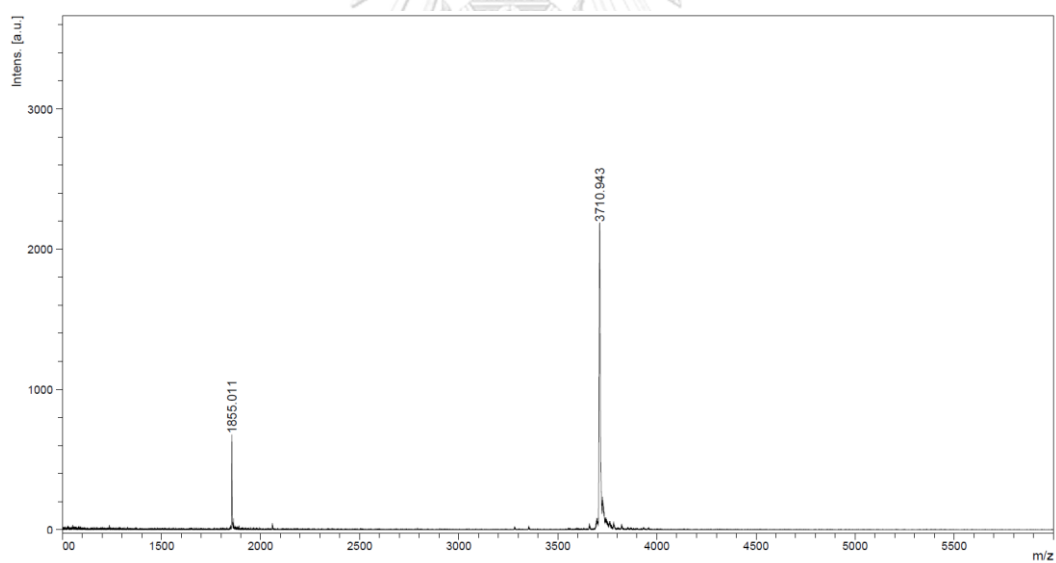
B

Figure A2 (A) Analytical HPLC chromatogram and (B) MALDI-TOF mass spectrum of PNA(I)
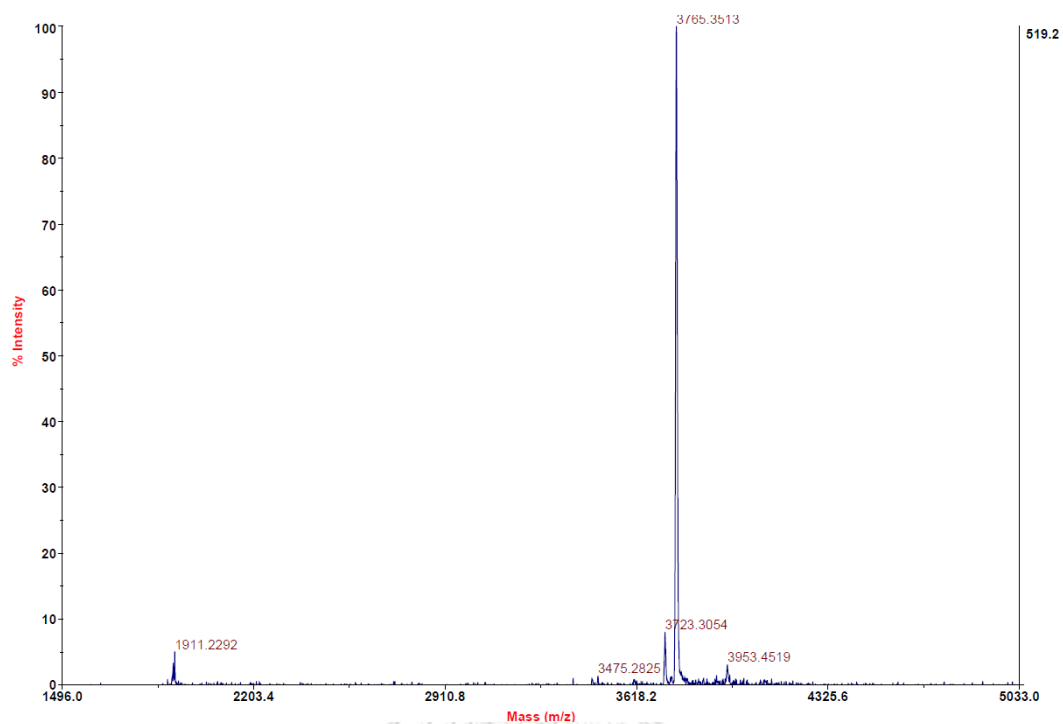
Figure A3 MALDI-TOF mass spectrum of Rf aegPNA
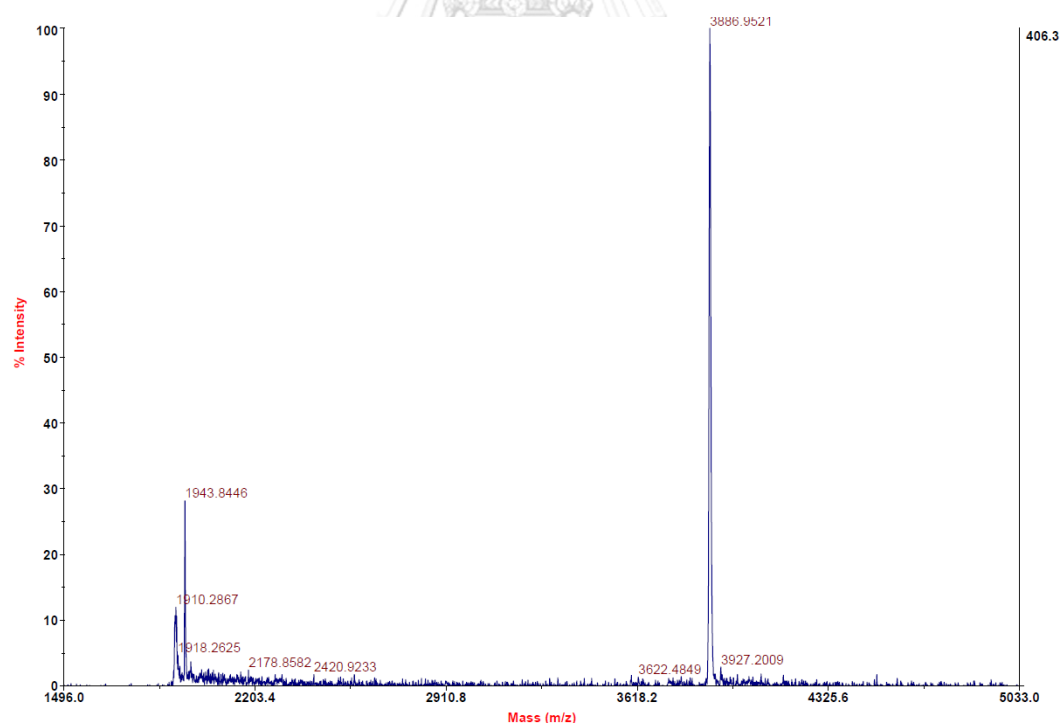
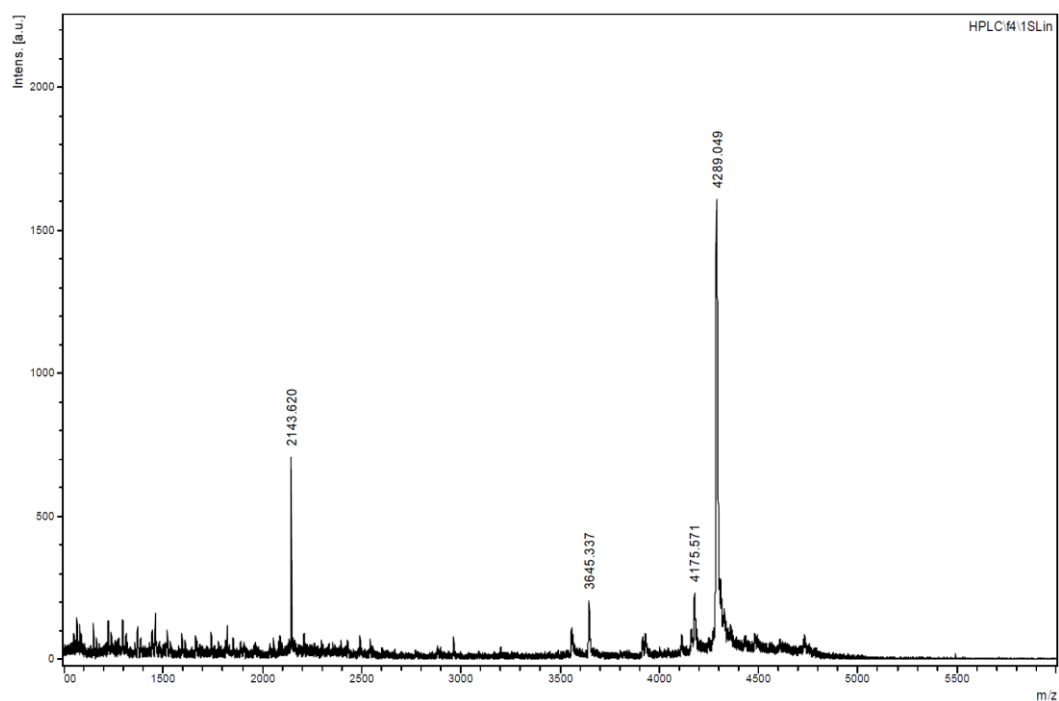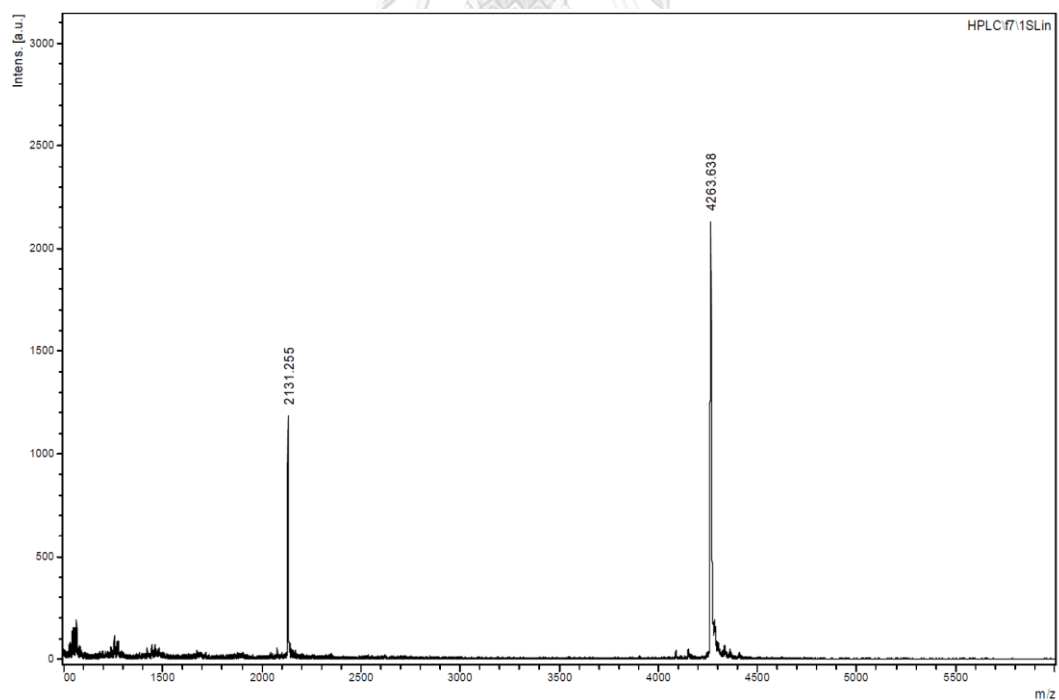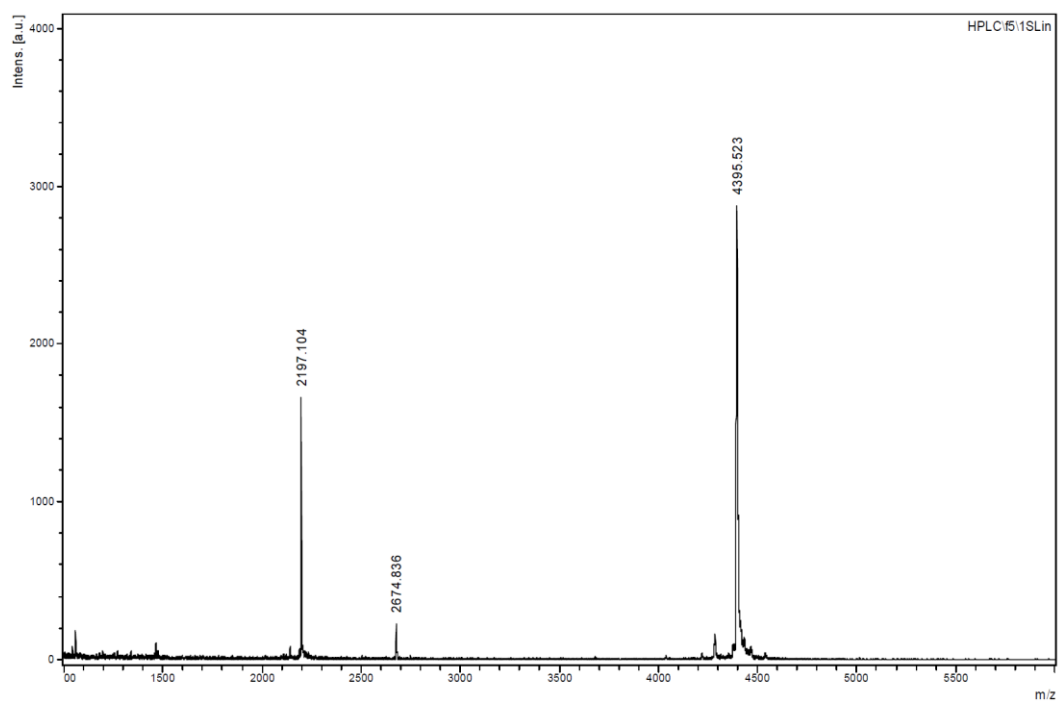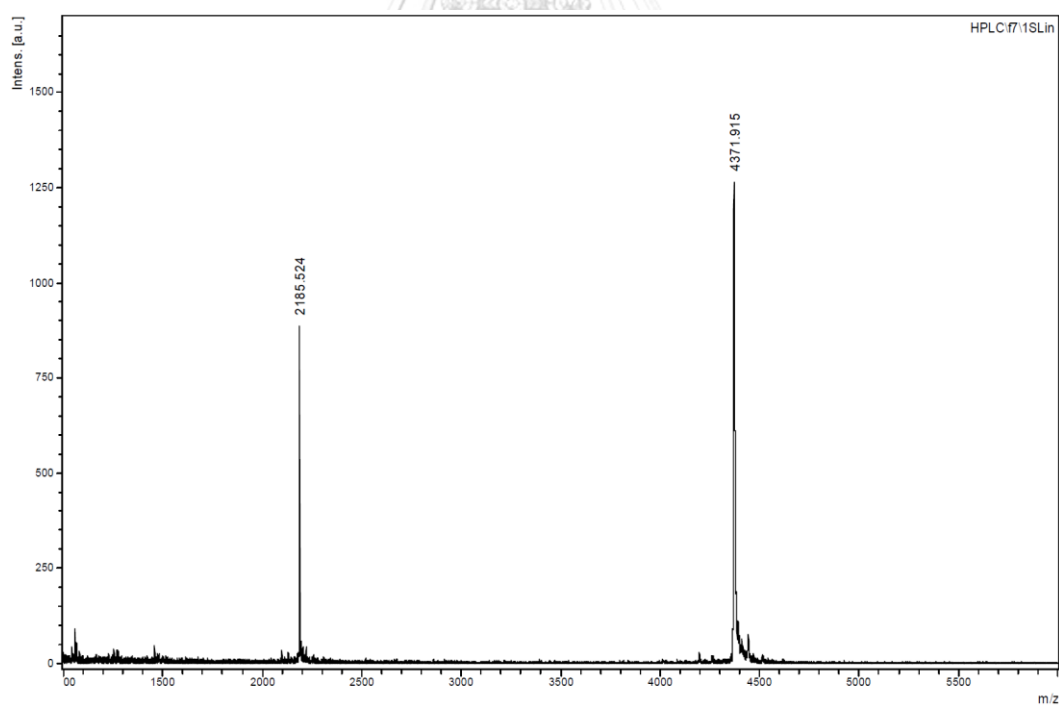**Figure A4** MALDI-TOF mass spectrum of RF aegPNA
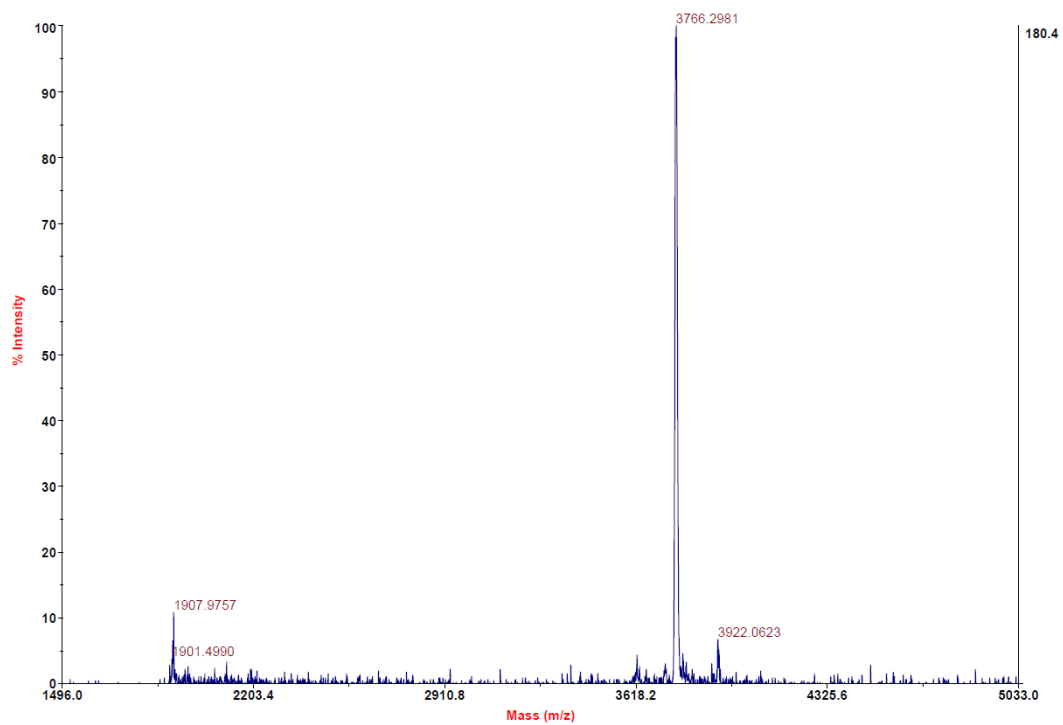
Figure A5 MALDI-TOF mass spectrum of RT aegPNA
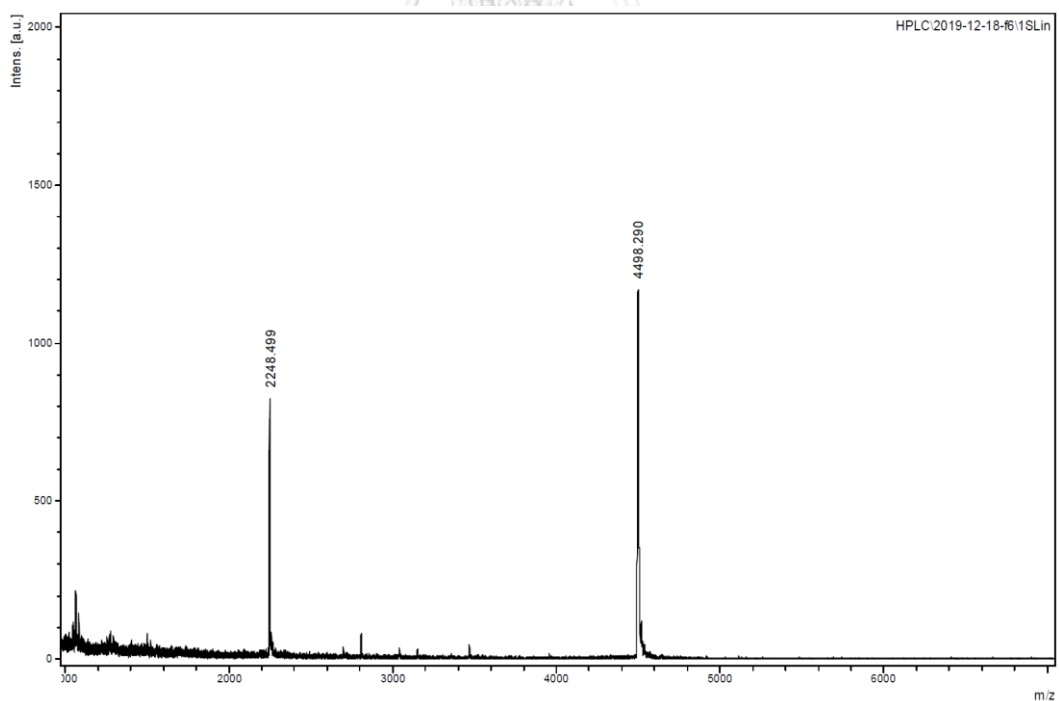
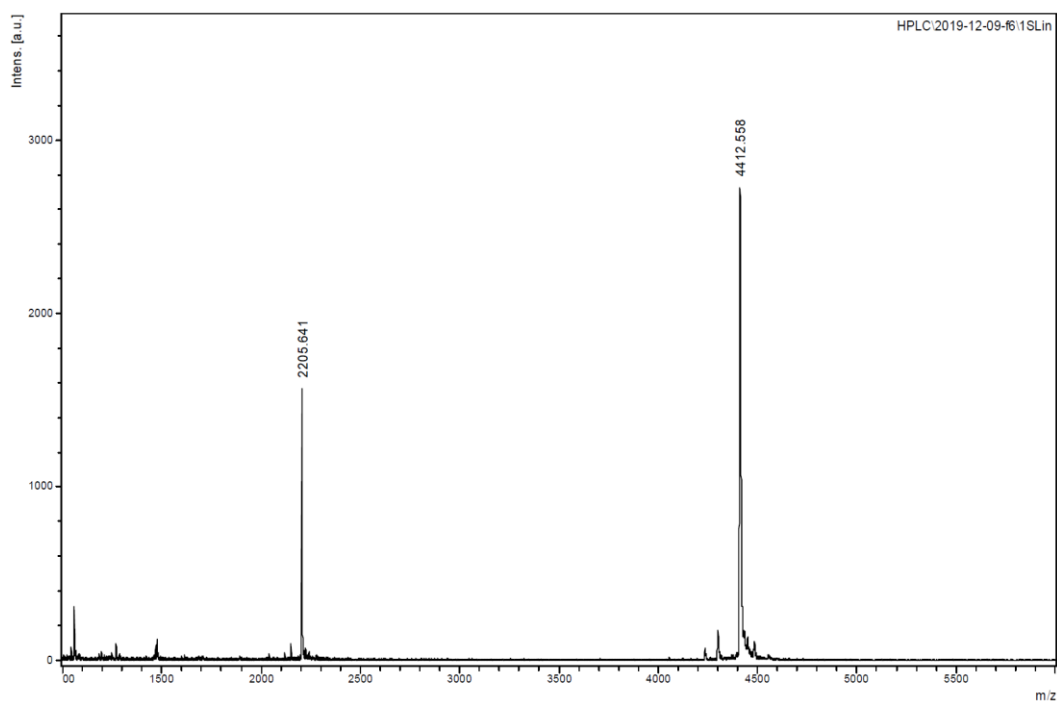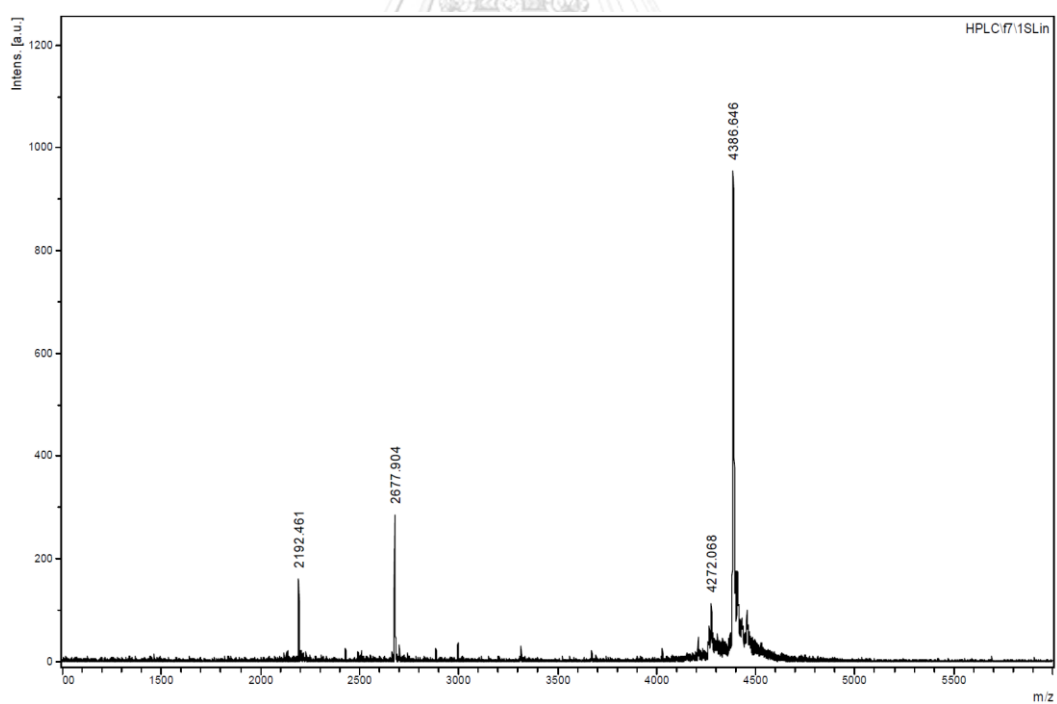**Figure A6** MALDI-TOF mass spectrum of Pf acpcPNA
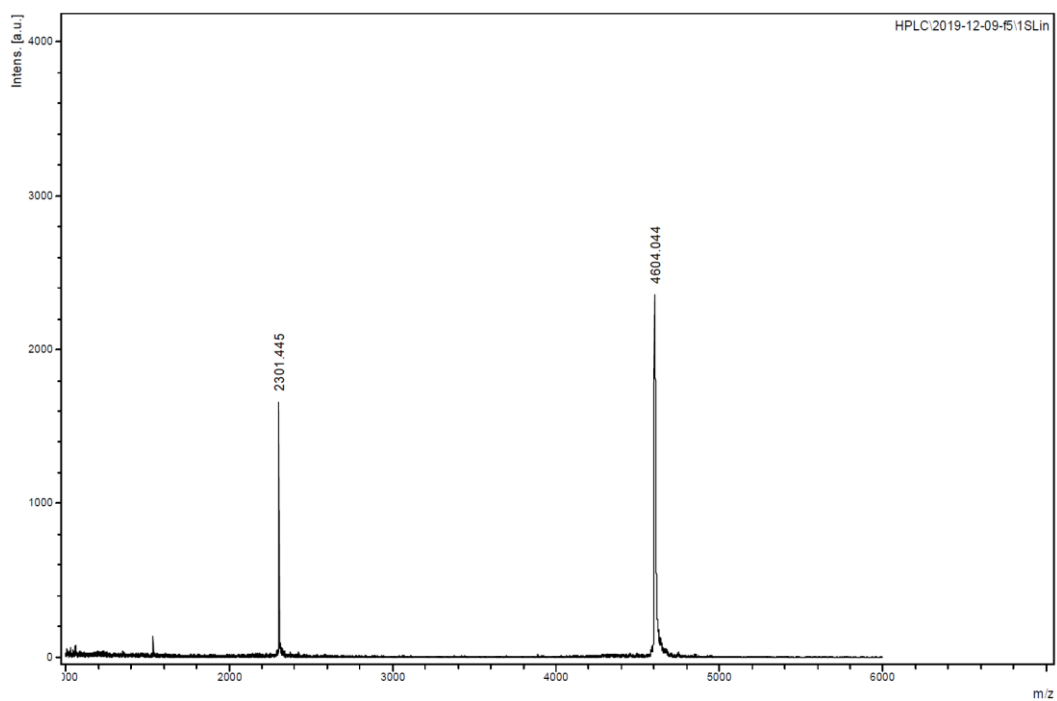
Figure A7 MALDI-TOF mass spectrum of PF acpcPNA
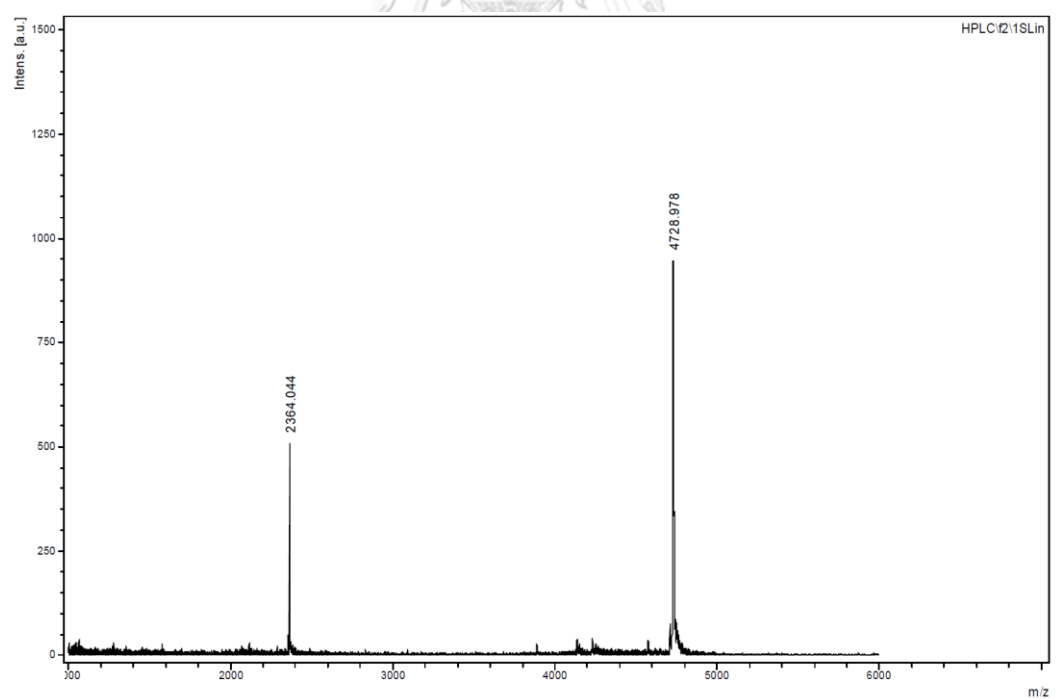
Figure A8 MALDI-TOF mass spectrum of PT acpcPNA
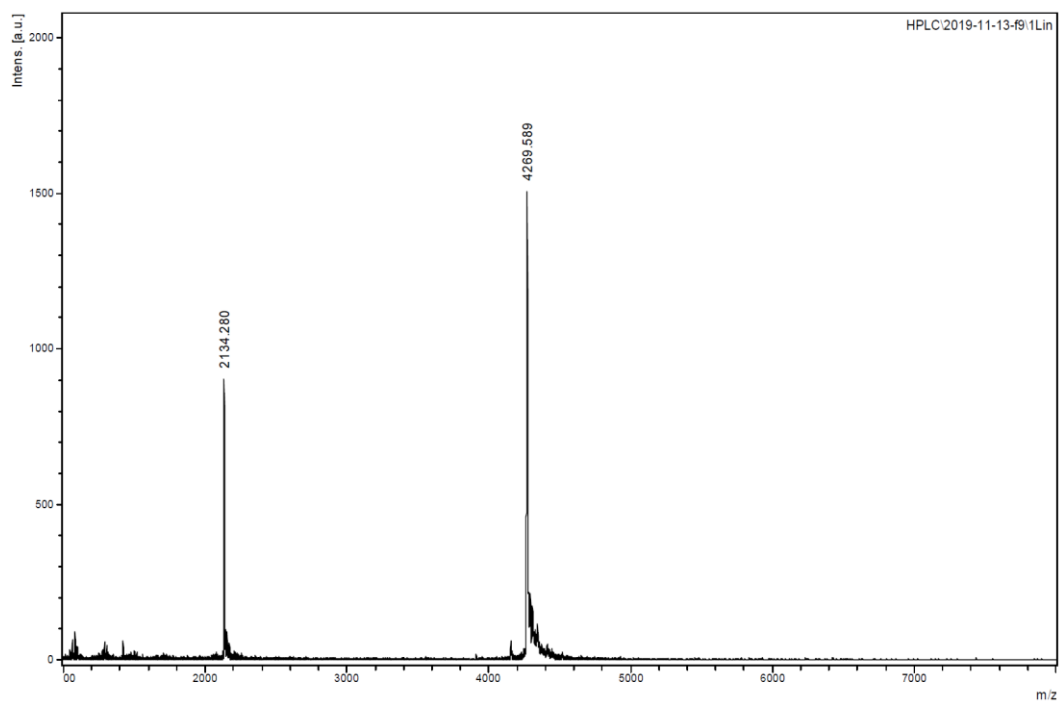
Figure A9 MALDI-TOF mass spectrum of AF acpcPNA
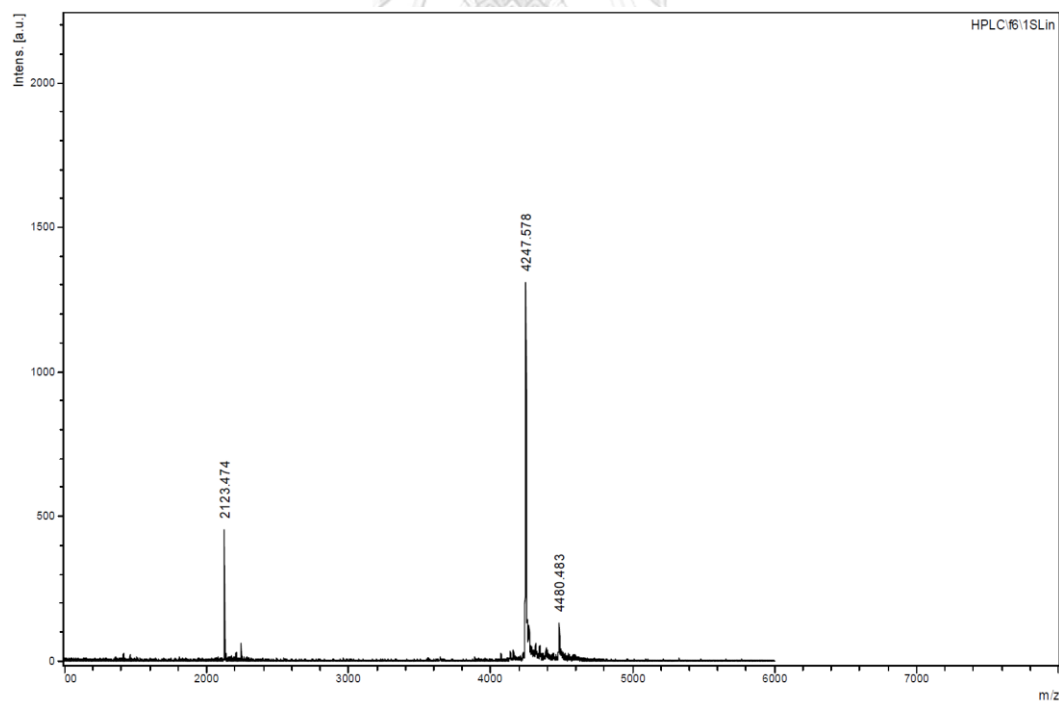
Figure A10 MALDI-TOF mass spectrum of CF acpcPNA

Figure A11 MALDI-TOF mass spectrum of GF acpcPNA

Figure A12 MALDI-TOF mass spectrum of TF acpcPNA

Figure A13 MALDI-TOF mass spectrum of AT acpcPNA

Figure A14 MALDI-TOF mass spectrum of CT acpcPNA

Figure A15 MALDI-TOF mass spectrum of GT acpcPNA

Figure A16 MALDI-TOF mass spectrum of TT acpcPNA
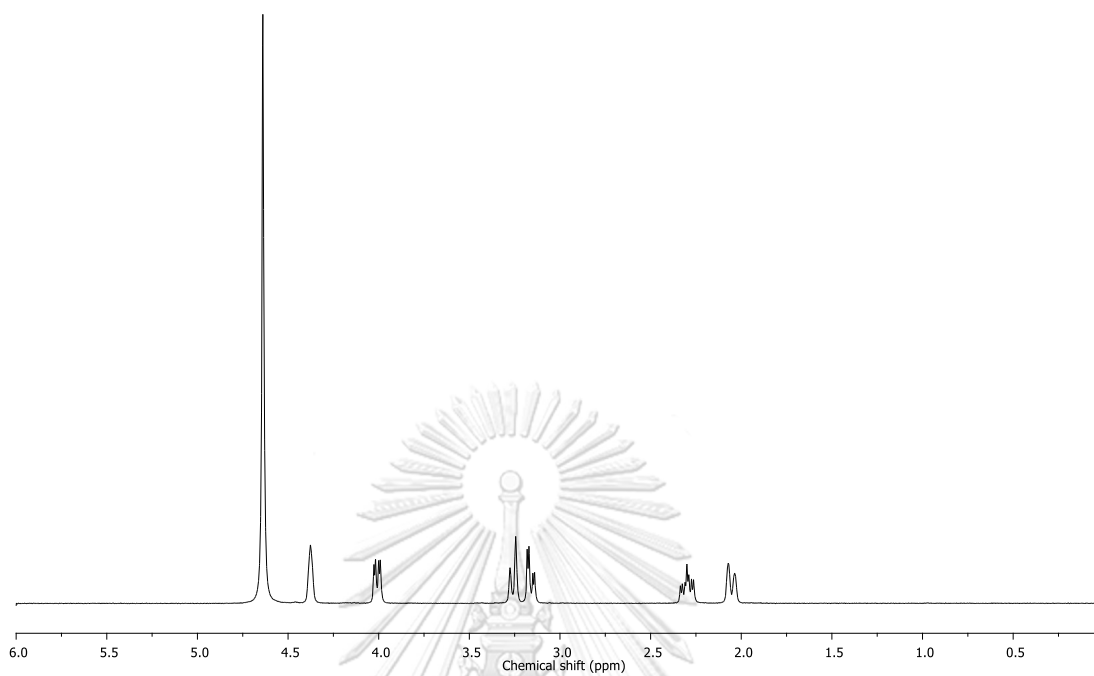
Figure A17 ^{1}H NMR spectrum of *cis*-4-Hydroxy-D-proline
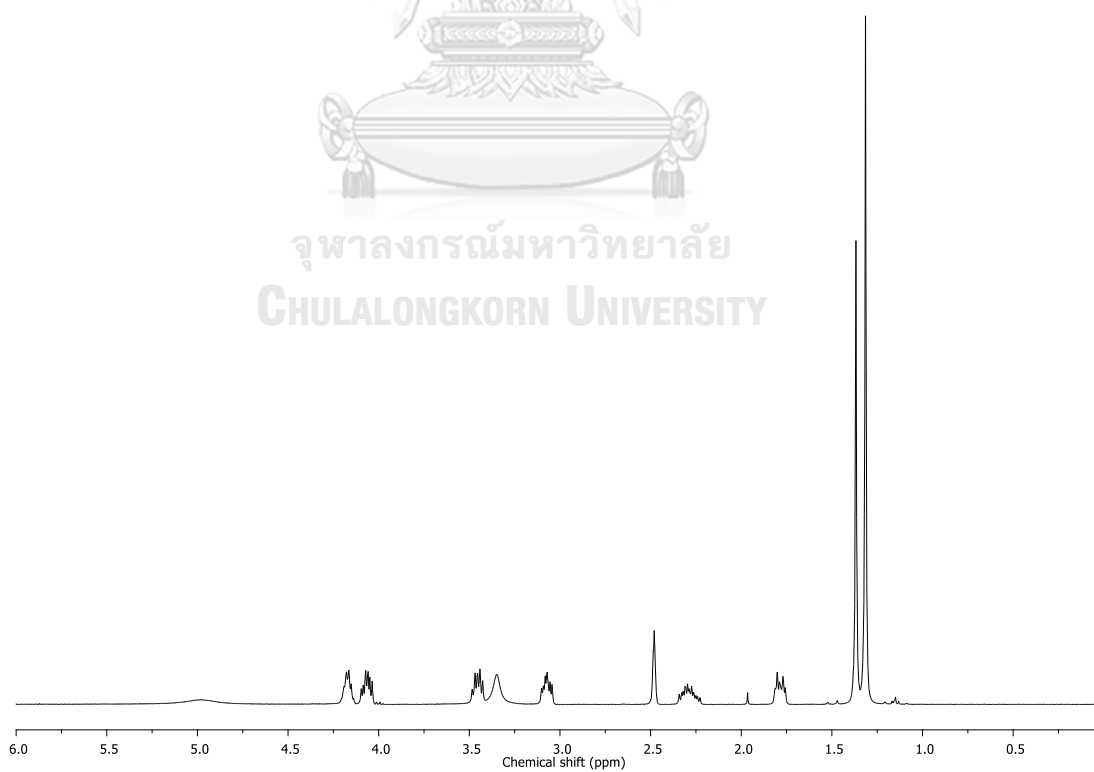
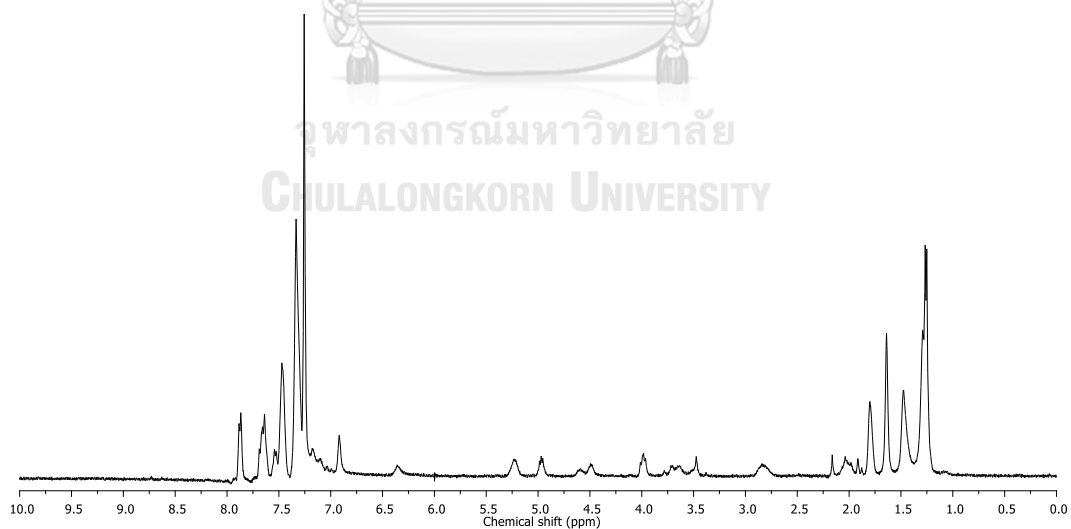**Figure A18** ^{1}H NMR spectrum of *N*-Boc-*cis*-4-hydroxy-D-proline
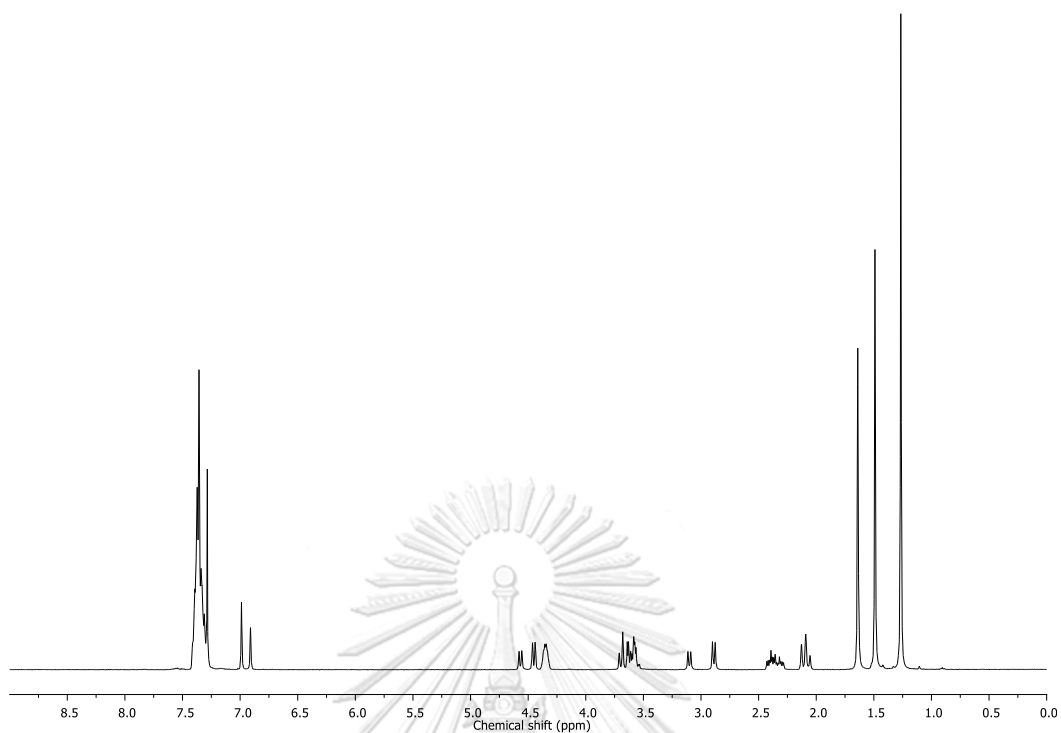
Figure A19 ¹H NMR spectrum of *N*-Boc-*cis*-4-hydroxy-D-prolinediphenylmethyl ester

Figure A20 ¹H NMR spectrum of *N*-Boc-*cis*-4-(*N*³-benzoylthymin-1-yl)-D-prolinediphenylmethyl ester
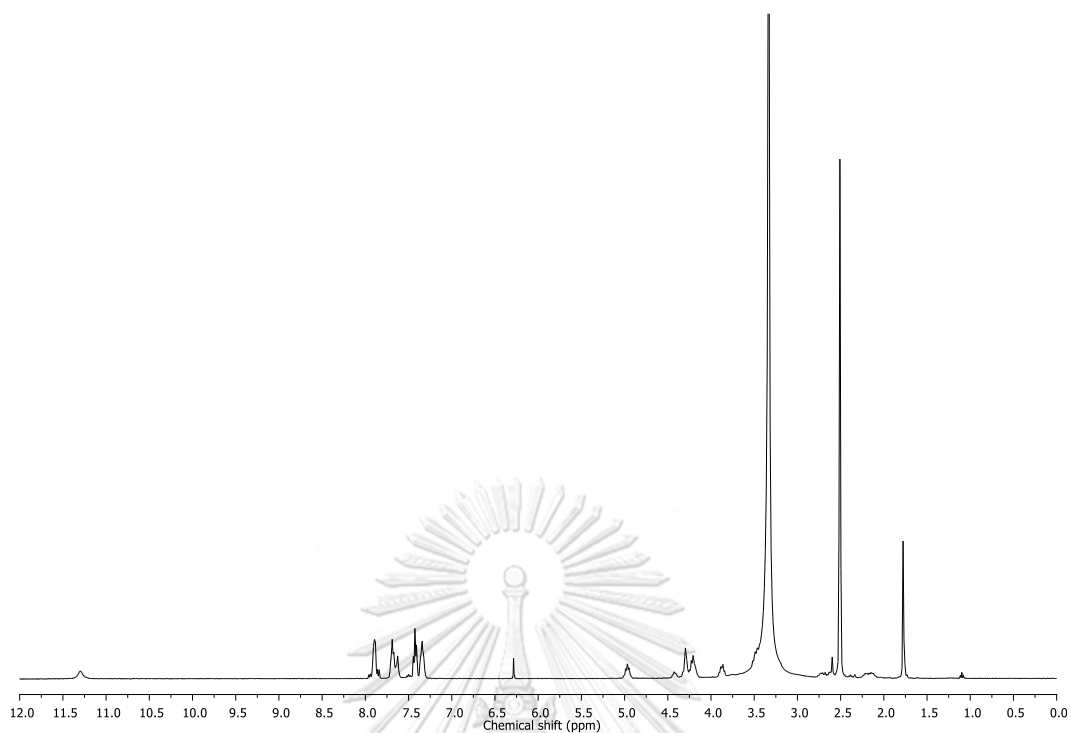
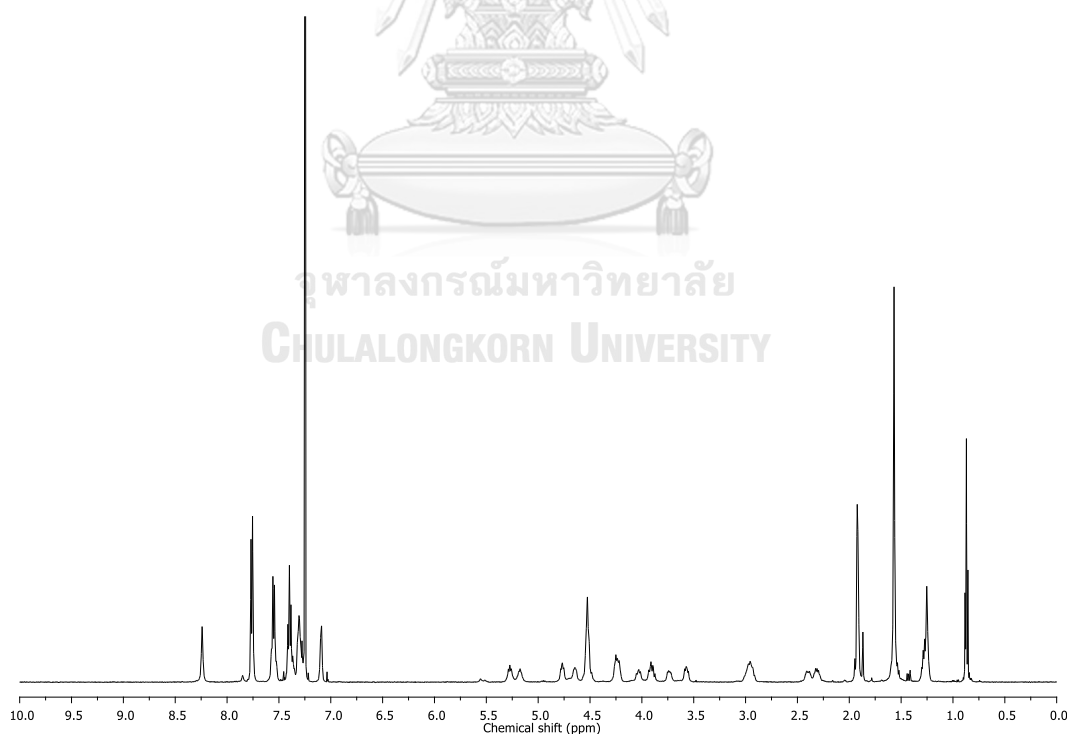**Figure A21** ^{1}H NMR spectrum of *N*-Fmoc-*cis*-4-(*N*^{3}-benzoylthymin-1-yl)-D-proline

Figure A22 ^{1}H NMR spectrum of *N*-Fmoc-*cis*-4-(*N*^{3}-benzoylthymin-1-yl)-D-proline pentafluorophenyl ester
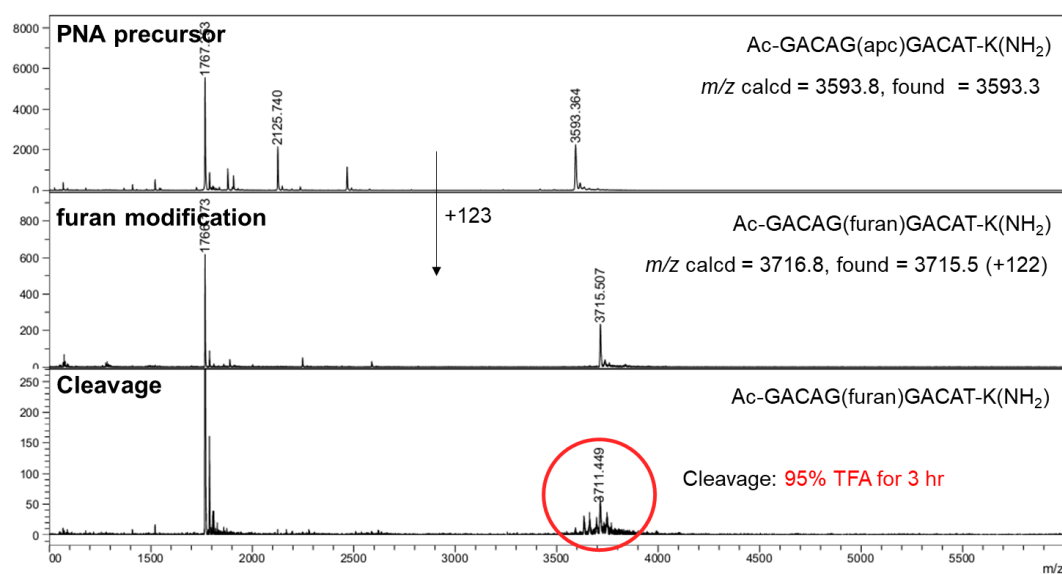
Figure A23 MALDI-TOF mass spectra showing the degradation of a furan-modified PNA (Ac-GACAG(furan)GACAT-K-NH$_2$ during the TFA cleavage step
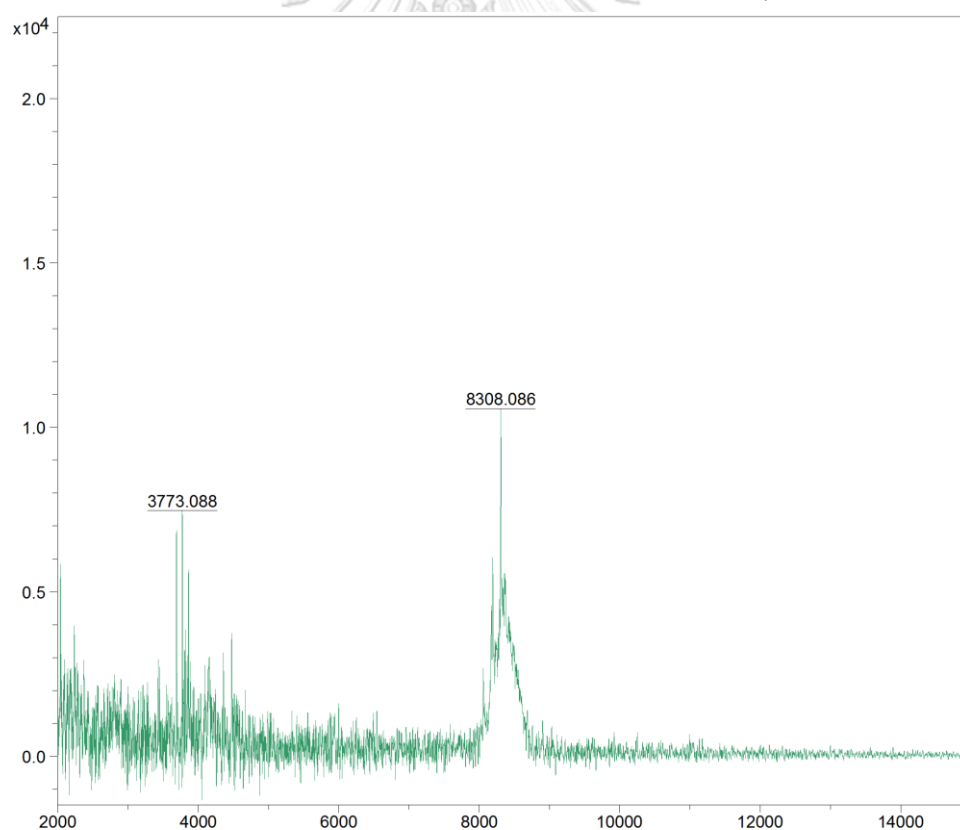

Figure A24 MALDI-TOF mass spectrum of the cross-linked product of PNA(T) and DNA-C5; (C$_{322}$H$_{405}$N$_{112}$O$_{124}$P$_{14}$)
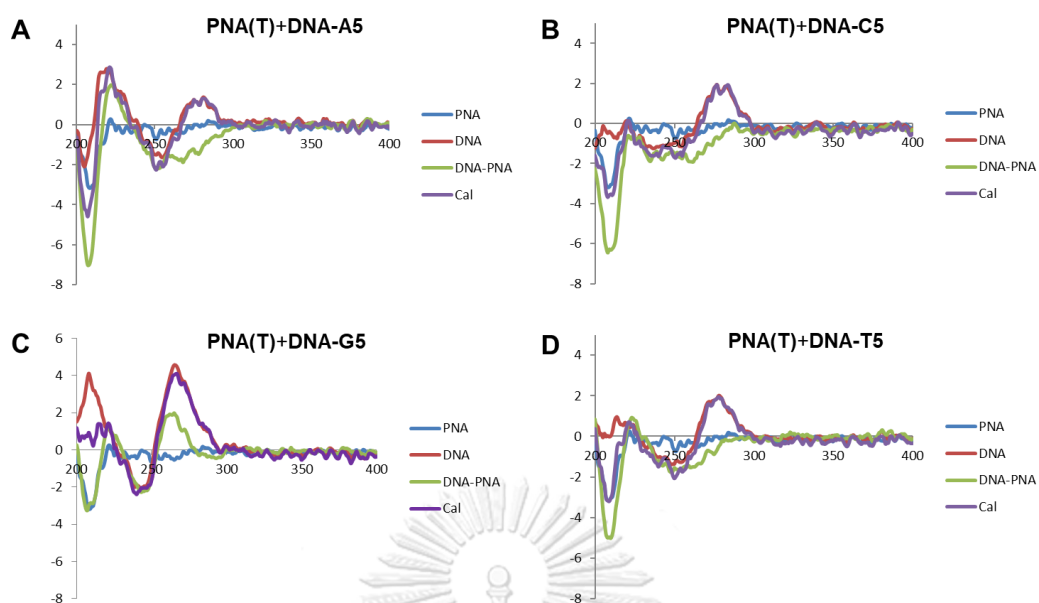
Figure A25 CD spectra of PNA(T) with DNA targets carrying different base overhangs [DNA-A5, DNA-C5, DNA-G5, and DNA-T5]. Conditions: [DNA] = [PNA] = 1 µM, 100 mM NaCl, 10 mM sodium phosphate buffer, pH 7.0. Color code: PNA(T) alone (blue), DNA alone (red), mixture of PNA and DNA (green), sum of PNA and DNA spectra (purple). The non-superimposable green and purple curves indicate duplex formation. The red curve in the case of DNA-G5 indicated the formation of a parallel G-quadruplex structure which diminished after the hybridization with PNA(T).

MALDI-TOF MS optimization

Matrix preparations

- 2,5-DHB (2,5-dihydroxy benzoic acid): 100 mg in 1 mL of H_2O: MeCN: TFA (1:2:0.1%)

- SA (Sinapic acid): 10 mg in 1 mL of H_2O: MeCN: TFA (1: 1: 0.1%)

- HCCA (α-cyano-4-hydroxy cinamic acid): saturated HCCA in 1 mL of H_2O: MeCN: TFA (7: 3:0.1%)

| matrix	method	sample	result
2,5-DHB	1. mix sample and matrix		✘
	2. sample first then matrix		✓
	3. matrix first then sample		✘
	4. matrix first then sample then matrix		✓
SA	1. mix sample and matrix	RFC1	✘
	2. sample first then matrix		degradation products
	3. matrix first then sample		✘
	4. matrix first then sample then matrix		degradation products
HCCA	1. mix sample and matrix		✘
	2. sample first then matrix		degradation products
	3. matrix first then sample		✘
	4. matrix first then sample then matrix		degradation products
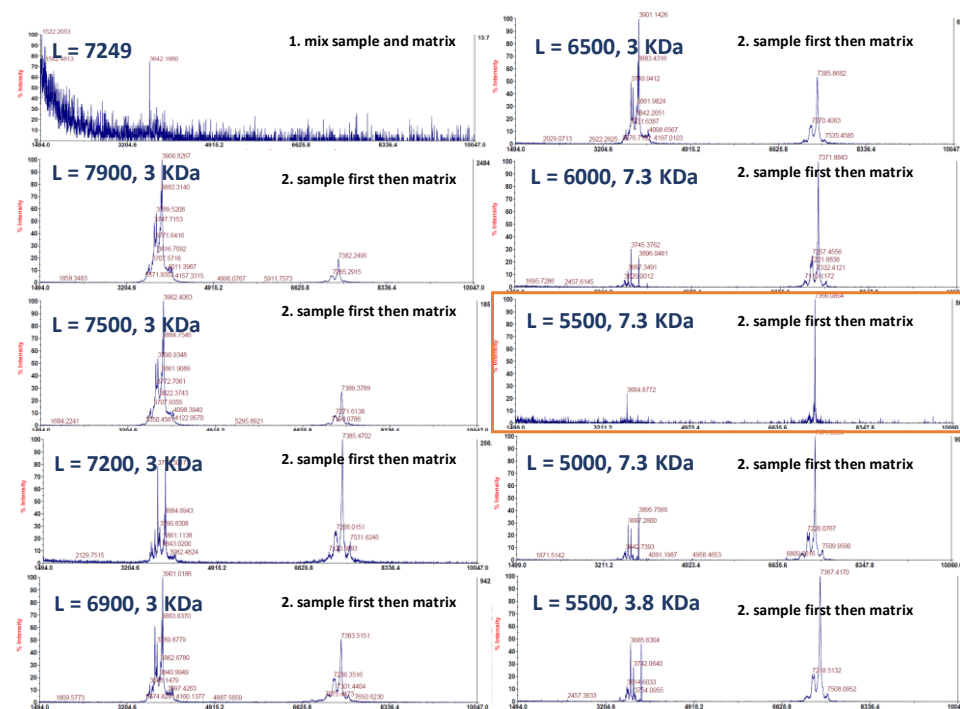
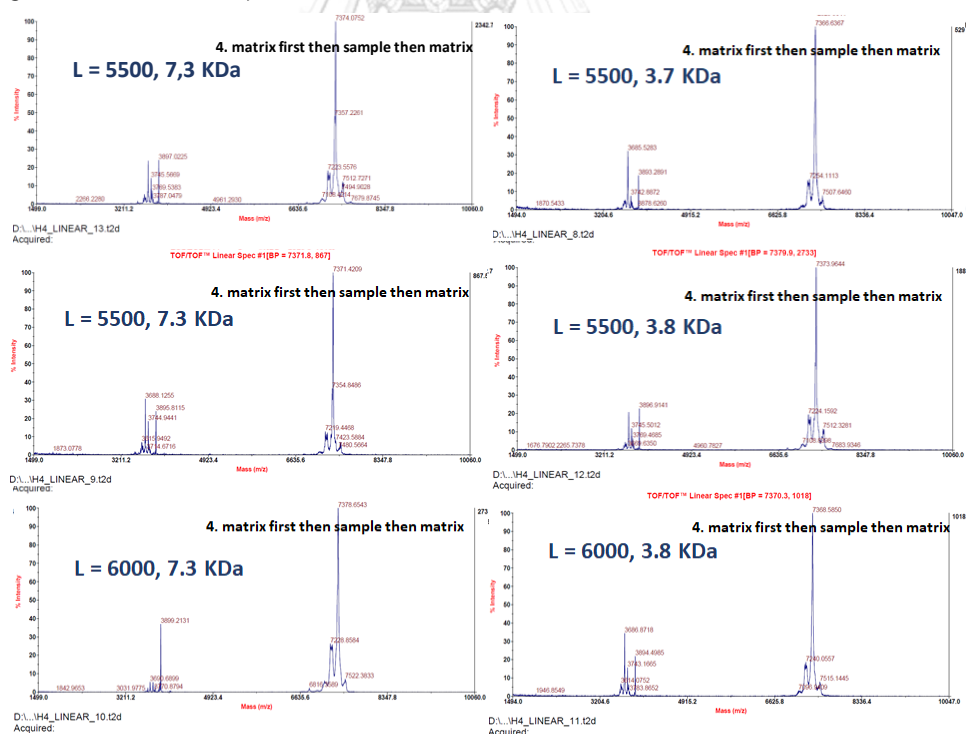**Figure A26** MALDI-TOF MS of the optimization of method 2 using 2,5-DHB as a matrix varying the laser intensity

Figure A27 MALDI-TOF MS of the optimization of method 4 using 2,5-DHB as a matrix varying the laser intensity

VITA

NAME	Penthip Muangkaew
DATE OF BIRTH	19 June 1991
PLACE OF BIRTH	NAN, Thailand
INSTITUTIONS ATTENDED	Bachelor's Degree of Science (Chemistry) in 2013

PUBLICATION

1. Ditmangklo, B.; Muangkaew, P.; Supabowornsathit, K.; Vilaivan, T., Synthesis of Pyrrolidinyl PNA and Its Site-Specific Labeling at Internal Positions by Click Chemistry. In Peptide Nucleic Acids: Methods and Protocols, Nielsen, P. E., Ed. Springer US: New York, 2020; pp 35-60.

2. Muangkaew, P.; Vilaivan, T., Modulation of DNA and RNA by PNA. Bioorg. Med. Chem. Lett. 2020, 127064.

3. Muangkaew, P.; Vilaivan, T., Pyrrolidinyl Peptide Nucleic Acid Probes Capable of Crosslinking with DNA: Effects of Terminal and Internal Modifications on Crosslink Efficiency. ChemBioChem. doi.org/10.1002/cbic.202000589