國立臺灣大學醫學院醫學檢驗暨生物技術研究所

博士論文

Department of Clinical Laboratory Sciences and Medical Biotechnology

College of Medicine

National Taiwan University

Doctoral Dissertation

利用飛行時間質譜儀進行基因恆定性相關酵素活性之研究

Study of Genomic Maintenance Enzymes by MALDI-TOF Mass Spectrometry Analysis

張惠嵐

Hui-Lan Chang

指導教授：方偉宏 博士

Advisor: Woei-horng Fang, Ph.D.

中華民國 113 年 7 月

July 2024

doi:10.6342/NTU202401466

Acknowledgment

回顧我在台大這七年的研究生生活，是有笑有淚，有苦有樂，十分精采。首先我要感謝口試委員許濤老師、蔡芷季老師、蘇剛毅老師、郭靜穎老師，在學位論文、實驗設計、實驗方法上都給予我充分的支持與幫助。特別感謝我的指導老師方偉宏老師，在我最無助最徬徨的時候伸出援手，接納我並收我為徒，引領我踏上研究之路。方老師在科學研究上給予許多引導與啟發，在待人處事上也時常提醒，感謝老師對我的信任與鼓勵，讓我知道我還有很多潛能可以發揮，還有許多不足的地方，需要持續進步與學習。

在讀博士班的期間受到許多人的幫助，何其幸運!首先感謝賴婉君學姊幫助完成質譜儀上機，以及協助完成 JOVE 期刊的拍攝。感謝雅蘭學姊、肯肯學長、威霖學長幫忙許多大大小小的事情。感謝 521-3 實驗室的大家：感謝我的學姊顏榕宣與林貴卿，教導我實驗技術，讓我熟悉研究生活。感謝碩士班的夥伴周能安，一起讀分生考古題不孤單。感謝雅芬學姊在臨床生化實驗課的幫忙。感謝學弟妹倖林、禮醫、靜宜、哲宇、羿凱、屾玨、詠筑、曉沛，謝謝你們與我一同成長，是實驗的好幫手!感謝有一股衝動的博士班前輩們：偉嘉、文淵、佳儀、淯妃、筱芳，給予我很多寶貴的經驗分享。感謝博士班的同學呂呂總是替我加油打氣。感謝宜臻、詠涵、煒倫在實驗上給予幫助。感謝我的朋友們特別是蓓薇、德盈、庭瑋，謝謝你們時常提醒我不要忘記休息。

最後我想感謝我的家人，尤其是我的父母，他們陪著我一起經歷這些日子，給予我無條件的支持與愛，讓我可以無後顧之憂的學習。每當遭遇挫折灰心，甚至想要放棄的時候，是爸爸媽媽的愛給我動力讓我繼續前行。真心祝福一路上幫助過我的每一個人，謝謝你們出現在我的生命中，在未來的每一天我都會懷著感恩的心，謙虛並勇敢的面對挑戰。

doi:10.6342/NTU202401466

中文摘要

此篇研究主要為利用基質輔助雷射脫附游離飛行時間質譜技術 (MALDI-TOF MS)，搭配未標定的 DNA 寡核苷酸受質設計，研究三種維護基因體穩定性酵素的活性，提出新的檢測酵素活性的方法，期望未來可應用於生技產業或臨床檢驗。此篇研究分為三個部分，第一部分為第一型 DNA 聚合酶校正在單一核苷酸插入/缺失錯誤的活性之研究；第二部分為大腸桿菌中 DNA 修復蛋白 Uracil-DNA glycosylase (UDG)活性之研究；第三部分為大腸桿菌第五型內切酶活性檢測方法之建立與研究。

重複序列 DNA 因為易於滑動，在複製過程中常見插入/缺失錯誤的產生，此錯誤通常可以透過 DNA 聚合酶校正活性進行校正，但若錯誤位置距離引子 3’端較遠，則可能因為末端正確配對使聚合酶直接進行引子延伸。為了研究插入/缺失錯誤之校正活性，我們設計了距離 3’端倒數第 1 至第 9 個位置的單一核苷酸插入/缺失錯誤，利用 MALDI-TOF MS 的方法分析 DNA 聚合酶的校正活性。使用 3 種 dNTP 或 4 種 ddNTP 進行校正活性的測定，研究結果發現距離 3’端第 1 至 5 個核苷酸的插入/缺失錯誤能夠很好的被第一型 DNA 聚合酶進行校正，在距離 3’端第 6 個核苷酸的插入/缺失錯誤僅部分被校正，距離 3’端第 7 至 9 個核苷酸的插入/缺失錯誤無法被 DNA 聚合酶校正，而會直接進行引子延伸。根據以上結果我們推測此校正活性與傾向，和 DNA 聚合酶與 DNA 引子-模板交界的接觸有關。

doi:10.6342/NTU202401466

第二部分為大腸桿菌 DNA 修復蛋白 Uracil-DNA glycosylase (UDG)活性之探討。尿密啶 (Uracil)是 DNA 中常見的損傷，由胞密啶水解脫氨形成。此 DNA 損傷若未即時進行修復，則會在 DNA 複製後形成 G:C to A:T 的轉換突變。在大腸桿菌中，主要由 UDG 進行 U 的修復，此酵素於原核與真核細胞中在演化上高度保留。已知許多測定 UDG 活性的方法都有其限制性，故我們欲發展高專一性與非標定受質的檢驗方法，期望能應用於臨床上抑制劑篩選的方法。我們利用帶有 U 的寡核苷酸受質與 UDG 酵素進行反應，形成帶有 apyrimidinic site (AP site) 的產物與 U 受質分子量相差 94，經過 MALDI-TOF MS 的分析，能夠用質譜圖的信號強度計算酵素的活性。我們利用此方法針對 UDG 進行酵素活性的研究，結果顯示其 Km 為 50 nM, Vmax 為 0.98 nM/s ，Kcat 為 9.31 s^{-1}，此結果與傳統法所得結果相近。利用 uracil DNA glycosylase inhibitor (UGI) 作為抑制物進行抑制效果檢測，結果測得 IC50 為 7.6 pM。另外針對多種的 U 受質檢測 UDG 對不同受質的專一性，結果顯示 U 在正中間且與 G 配對的受質(T1/U+9)最適合用於檢測 UDG 的活性，且以質譜儀檢測 UDG 活性能夠很好的應用於臨床上的檢測。

第三部分則利用 MALDI-TOF MS 的平台建立用以檢測大腸桿菌內第五型核酸內切酶酵素(Endo V)的活性。DNA 中的腺嘌呤經脫氨作用會形成亞黃嘌呤(dI)的損傷，若未被及時修復再次複製後會形成 A:T to G:C 的轉換突變。在大腸桿菌中 Endo V 為參與 dI 修復的重要酵素，先前研究發現 dI 的修復由 Endo V、第一型 DNA 聚合酶(Pol I)與 DNA 連接酶完成。Endo V 在切割完 dI 受質後會停留在

doi:10.6342/NTU202401466

DNA 受質上，被認為可能與後續 Pol I 的作用有關。為了能夠更深入研究 Endo V 的活性，我們發展以 MALDI-TOF MS 方法搭配高週轉性的 dI 受質進行 Endo V 酵素活性的分析，結果發現高週轉性受質能夠提高 Endo V 週轉 5 次的切割活性，並能夠以動力學方法分析發現 Endo V 對不同 DNA 受質具有特異性(I-G> I-A ≒I-T> I-C)，且利用 MS 的方法相較於螢光法可提升 16 倍的靈敏度。我們也利用尿素變性聚丙烯醯胺膠體電泳搭配螢光標定的 DNA 受質，以研究 Endo V 與 Pol I 的交互作用，發現具有 3’端至 5’端外切酶活性 Pol I 可以促進 Endo V 週轉 1.5 倍的活性，而沒有外切酶活性的 Pol I 則促進週轉活性能力較差。因此可從這個結果判定 Pol I 與 Endo V 之間在修復過程有交互作用。

在此篇論文中我們展現了 MALDI-TOF MS 此平台檢驗的優勢，快速、高通量、高專一性，搭配非標定的 DNA 受質且彈性的受質設計，可以更全面的了解酵素的活性，更重要的是可以做抑制物的檢測，將可應用於生技製藥的研發。

關鍵詞:核酸修復、MALDI-TOF MS、Klenow fragment 校正活性、大腸桿菌 UDG、大腸桿菌第五型核酸內切酶。

doi:10.6342/NTU202401466

Abstract

Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) has been widely used in the detection of DNA modifications and SNPs. In this study, a non-labeled DNA substrate assay was established using MALDI-TOF MS to study the activities of DNA maintenance enzymes, including DNA polymerase I (Pol I), uracil DNA glycosylase (UDG), and endonuclease V (Endo V).

Insertion/deletion errors frequently occur during the replication of repetitive DNA sequences and could be corrected by DNA polymerase proofreading activity. A series of single nucleotide insertion/deletion (indel) error substrates, with the indel errors located 1 to 9 nucleotide (nt) from the 3' terminus, were designed for a proofreading assay and subjected to MALDI-TOF MS. The proofreading activity was performed with three dNTPs or four ddNTPs. The results revealed that indel errors located 1 to 5-nt from the 3' terminus could be effectively proofread by the Pol I, while partially proofread when indel errors at 6-nt from the 3’ terminus. Indel error at 7 to 9-nt from 3’terminus escaped proofreading, leading to primer extension. Based on these results, it suggests that proofreading activity is correlated with the interaction between DNA polymerase and primer-template junction.

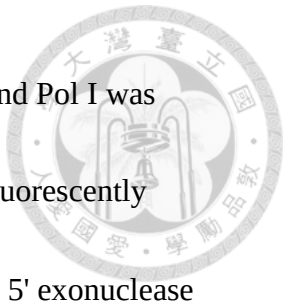
Uracil (U) is a DNA damage caused by the deamination of cytosine. It would cause G: C to A: T transition mutations if not repaired prior to DNA replication. In *E. coli*,

doi:10.6342/NTU202401466

uracil is mainly repaired by uracil UDG, an enzyme that is highly conserved in both prokaryotic and eukaryotic cells. The UDG detection assay was designed with site-specific uracil and performed by MALDI-TOF MS. UDG activity was determined using a non-labeled double-strand oligonucleotide carrying a uracil on its middle position. Kinetic parameters of UDG with Km at 50 nM, Vmax at 0.98 nM/s, and Kcat of 9.31 s^{-1} were obtained by mass spectrometric analysis and these parameters were comparable to those revealed by traditional methods. The inhibitor screen assay was performed with a uracil glycosylase inhibitor (UGI), yielding an IC50 at 7.6 pM. Additionally, the UDG substrate specificity for various uracil substrates was tested, concluding that the substrate with uracil located at the center and paired with guanine is the most suitable for detecting UDG activity.

Deoxyinosine (dI) is formed by the deamination of adenine and leads to A: T to G: C transition mutations if not repaired before DNA replication. In *E. coli*, dI repair involves Endo V, Pol I, and DNA ligase. Previous nicking assay showed tightly binding of Endo V to dI-containing DNA suggesting the possible role of dI-bound Endo V as a repair signal for Pol I. To study the interplay of Pol I and Endo V in the repair of dI, the activity of Endo V was determined using a MALDI-TOF MS-based assay and a PAGE assay. Analysis of Endo V activity using MALDI-TOF MS showed a 16-fold higher nicking activity compared to the fluorescence-based method, and substrate specificity in

doi:10.6342/NTU202401466



the order of I-G > I-A ≈ I-T > I-C. The interaction between Endo V and Pol I was studied by denaturing urea polyacrylamide gel electrophoresis with fluorescently labeled DNA substrates. The results indicated that Pol I KF with 3' to 5' exonuclease activity enhances Endo V turnover by 1.5 times, while Pol I without exonuclease activity showed much less turnover. These findings suggest a direct interaction between Pol I and Endo V in the repair process.

In conclusion, we demonstrate that MALDI-TOF MS is a powerful tool for enzyme activity analysis. It offers rapid, high-throughput, and highly specific results, and allows for flexible substrate design using unlabeled DNA. This approach provides a comprehensive understanding of enzyme activity and shows great potential for inhibitor screening in biomedical settings.

Keywords: MALDI-TOF MS, DNA polymerase I proofreading activity, uracil DNA glycosylase, Endonuclease V, turnover activity.

doi:10.6342/NTU202401466

Contents

Acknowledgment ····· i

中文摘要 ····· ii

Abstract ····· v

Contents ····· viii

List of Figures ····· xiv

List of Tables ····· xvi

Abbreviations ····· xvii

Chapter I: Introduction ····· 1

1.1 Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry (MALDI-TOF MS) for DNA Detection ····· 1

1.2 Indel Error and DNA Proofreading ····· 5

1.2.1 Insertion/ Deletion Error ····· 5

1.2.2 Correction of Insertion/Deletion Error ····· 7

1.2.3 DNA Polymerases in *E. coli* ····· 8

1.2.4 *E. coli* DNA Polymerase I ····· 10

1.2.5 The Proofreading Assays ····· 13

1.3 Base Excision Repair and Glycosylase ····· 15

1.3.1 Uracil in DNA ····· 15

doi:10.6342/NTU202401466

1.3.2 Base Excision Repair (BER) Pathway ····· 17

1.3.3 Uracil DNA Glycosylase (UDG) ····· 19

1.3.4 5-Fluorouracil and UDG Activity ····· 21

1.3.5 The Detection Method of UDG Activity ····· 22

1.4 Repair of Deoxyinosine and *E. coli* Endonuclease V ····· 26

1.4.1 Deoxyinosine in DNA ····· 26

1.4.2 Repair of Deoxyinosine in DNA ····· 27

1.4.3 *E. coli* Endonuclease V ····· 28

1.4.4 Homologs of Endonuclease V ····· 30

1.4.5 The Endonuclease V Repair Pathway ····· 31

1.5 Research Motivation and Strategy ····· 32

Chapter II Materials and Methods ····· 34

2.1 Materials ····· 34

2.1.1 Synthetic DNA Oligonucleotides ····· 34

2.1.2 Enzymes ····· 34

2.1.3 Reagents ····· 35

2.2 DNA Duplex Substrate Preparation ····· 35

2.3 DNA Polymerase Proofreading Assay for MALDI-TOF MS ····· 36

2.4 Uracil DNA Glycosylase Assay for MALDI-TOF MS ····· 37

doi:10.6342/NTU202401466

2.5 Endonuclease V Nicking Assay for MALDI-TOF MS…… 37

2.6 MALDI-TOF MS Analysis …… 38

2.7 MS Data Analysis…… 39

2.8 Fluorescence-Based Gel Assay…… 39

2.9 Detection of DNA Possessing Enzymes by Fluorescently Labeled DNA Duplex 40

Chapter III Result …… 42

3.1 Proofreading of Single Nucleotide Insertion/deletion Replication Errors Analyzed by MALDI-TOF Mass Spectrometry Assay …… 42

3.1.1 The Indel DNA Substrate Design and Reaction Condition…… 42

3.1.2 Single Nucleotide Deletion Error at 1 to 6-nt from 3' Terminus Can be Proofread by Pol I with 4ddNTPs…… 44

3.1.3 Single Nucleotide Deletion Error at 1 to 5-nt from 3' Terminus Can be Proofread by Pol I with 3ddNTPs…… 45

3.1.4 Single Nucleotide Insertion Error at 1 to 5-nt from 3' Terminus Can be Proofread by Pol I with 4ddNTPs…… 46

3.1.5 Single Nucleotide Insertion Error at 1 to 5-nt from 3' Terminus Can be Proofread by Pol I with 3dNTPs…… 47

3.1.6 The Ability of DNA Polymerase I to Correct Indel Errors…… 48

doi:10.6342/NTU202401466

3.2 Measurement of Uracil-DNA Glycosylase Activity by Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry Technique …… 49

3.2.1 The Scheme of Uracil DNA Glycosylase Activity Assay …… 49

3.2.2 UDG Kinetic Parameters Determined by MS-based Assay were Comparable to Traditional Method …… 51

3.2.3 Investigation of UDG Substrate Specificity and Glycosylase Activity with Varied Uracil Positions in DNA Oligonucleotides …… 54

3.2.4 The Inhibitor Screen Assay of Uracil DNA Glycosylase …… 55

3.3 Study of *E. coli* Endonuclease V Enzyme Activity by MS-based and PAGE-based Analysis …… 57

3.3.1 The Scheme of Endonuclease V Nicking Assay by MALDI-TOF MS Analysis 57

3.3.2 Enzymatic Analysis of Endonuclease V by MS-based Analysis …… 59

3.3.3 The Substrate Specificity of Endonuclease V Nicking Activity…… 61

3.3.4 The Fluoresce-based Analysis for Endonuclease V Nicking Assay · 62

3.3.5 Endonuclease V Nicking Activity Analyzed by The Fluorescence-based Assay …… 64

3.3.6 The Exonuclease Activity of Pol I Promotes Endonuclease V Turnover…… 65

doi:10.6342/NTU202401466

Chapter IV Discussion ····· 67

4.1 The MS-based Analysis of DNA Processive Proteins ····· 67

4.1.1 The Prospective of MALDI-TOF MS Analysis in Biomedical Research and Development ····· 67

4.1.2 Mass-to-Charge Ratio Consideration of DNA Substrates and Products Analysis ····· 68

4.1.3 Adducts Interference and Intervention in Measurement ····· 69

4.2 The Different Outcomes of 4 ddNTPs and 3 dNTPs in Proofreading Assay 70

4.3 The Contact of KF and DNA Junction Affects the Proofreading Efficiency ····· 73

4.4 Acid Termination Method Offers Safer Alternative to Phenol/Chloroform Extraction in Glycosylase Assay ····· 74

4.5 Comparison of UDG Kinetic Parameters Using MALDI-TOF MS and Traditional Radioisotope Methods ····· 75

4.6 The Comparison of MS-based and Fluorescence-labeled Assays for Endonuclease V Repair Pathway Study ····· 76

4.7 The Turnover Activity of Endonuclease V ····· 79

Chapter V Conclusion ····· 82

doi:10.6342/NTU202401466

Figures ····· **84**

Tables ····· **118**

References ····· **126**

Biography ····· **143**

Publications ····· **144**

doi:10.6342/NTU202401466

List of Figures

Figure 1. The graphic abstract of MALDI-TOF MS based analysis of DNA repair proteins 84

Figure 2. The scheme of the proofreading assay by MALDI-TOF MS analysis 85

Figure 3. The single nucleotide deletion error from position 1 to 5 could be proofread by Klenow fragment effectively with 4ddNTPs 86

Figure 4. The single nucleotide deletion error from position 1 to 4 could be proofread by Klenow fragment effectively with 3dNTPs and ddCTP 88

Figure 5. The single nucleotide insertion error from position 1 to 5 could be proofread by Klenow fragment effectively with 4ddNTPs 90

Figure 6. The single nucleotide insertion error from position 1 to 5 could be proofread by Klenow fragment effectively with 3dNTPs 92

Figure 7. The scheme of DNA Polymerase I proofreading capacity for indel errors 94

Figure 8. Model system for DNA uracil glycosylase assay..... 95

Figure 9. The UDG enzymatic activity measured by MS-based analysis showed time and dose dependency 97

doi:10.6342/NTU202401466

Figure 10. The UDG kinetic parameter determined by MALDI-TOF MS was comparable to traditional method. 99

Figure 11. UDG assay by MALDI-TOF MS analysis was comparable to the manufacturer-assigned unit..... 101

Figure 12. The inhibition curve of UGI 102

Figure 13. Signal intensity and molecular weight..... 104

Figure 14. Example of high turnover substrate for Endo V nicking activity analyzed by MALDI-TOF MS 106

Figure 15. Endo V enzymatic activity analysis by MALDI-TOF MS..... 108

Figure 16. Endo V nicking specificity analysis using 21bp dI-substrate..... 110

Figure 17. The detection limit of fluorescence-based gel assay 111

Figure 18. Endo V nicking assay using fluorescence-labeled substrate and sensitivity assay 113

Figure 19. Endo V substrate specificity analysis using fluorescence-based gel assay 115

Figure 20. Pol I variants promote Endo V turnover using fluorescence-labeled gel assay 116

doi:10.6342/NTU202401466

List of Tables

Table 1 The comparison of uracil DNA glycosylase detection assay 118

Table 2 sequences of template-primers containing a single nucleotide deletion or insertion error 119

Table 3 U-containing substrate compositions and UDG specificities 120

Table 4 The sequence for MS normalization 121

Table 5 dI-containing substrate for MS-based Endo V analysis..... 122

Table 6 dI-containing substrate for fluorescence-based Endo V analysis..... 123

Table 7 Example of absorbance monitoring at λ = 260 nm for substrate duplex preparation...... 124

Table 8 Comparison of MS-based and PAGE-based analysis for Endo V 125

doi:10.6342/NTU202401466

Abbreviations

2-AP	2-Aminopurine
5’-dRP	5’-Deoxyribose Phosphate
A	Adenine
AER	Alternative Excision Repair
AID	Activation-Induced Cytidine Deaminase
AlkA	3-methyladenine DNA glycosylase II
a-NHEJ	Alternative Non-Homologous End Joining
AP	Apyrimidine/ Apurine
AP site	Abasic site
APE1	AP endonuclease 1
BER	Base Excision Repair
bp	Base Pair
C	Cytosine
C_H	Heavy Chain of Immunoglobulin
c-NHEJ	Classic Non-Homologous End Joining
CSR	Class Switch Recombination
dATP	Deoxyadenosine 5'-Triphosphate
dCTP	Deoxycytidine 5'-Triphosphate
ddATP	2', 3'-Dideoxyadenosine 5'-Triphosphate
ddCTP	2', 3'-Dideoxycytidine 5'-Triphosphate
ddGTP	2', 3'-Dideoxyguanosine 5'-Triphosphate
ddNTP	2', 3'-Dideoxynucleotide 5'-Triphosphate
ddTTP	2', 3'-Dideoxythymidine 5'-Triphosphate

doi:10.6342/NTU202401466

DEA	Diethanolamine
dGTP	Deoxyguanosine 5'-Triphosphate
dI	Deoxyinosine
dIMP	Deoxyinosine-5'-Monophosphate
dITP	Deoxyinosine Triphosphate
DNA	Deoxyribonucleic Acid
dNTP	Deoxynucleotide 5'-Triphosphate
dRP lyase	Deoxyribose-Phosphodiesterase
ds	Double Stranded
DSB	Double Strand Break
DTT	1,4-Dithiothreitol
dTTP	Deoxythymidine 5'-Triphosphate
dUTP	Deoxyuridine Triphosphate
dXTP	Deoxyxanthosine Triphosphate
E. coli	*Escherichia Coli*
EDTA	Ethylenediaminetetracetic Acid
Endo V	Endonuclease V
exo-	Exonuclease-deficient
exo site	3'→ 5' Exonuclease Activity Site
FAM	Fluorescein amidite
FEN1	Flap Endonuclease 1
Fpg	Formamidopyrimidine DNA Glycosylase
FRET	Fluorescent Resonance Energy Transfer
G	Guanine
GC	Glassy Carbon

doi:10.6342/NTU202401466

IMP	Inosine-5'-Monophosphate
Indel	Insertion/Deletion
ITPase	Triphosphate Pyrophosphatase
KF	Klenow Fragment
LC-MS/MS	Liquid Chromatograph/Tandem Mass Spectrometer
m/z	Mass-To-Charge Ratio
MALDI-TOF MS	Matrix Assisted Laser Desorption Ionization-Time Of Flight Mass Spectrometry
MBD4	Methyl-CPG Binding DNA Glycosylase
MBP	Maltose Binding Protein
MMR	Mismatch Repair
MS	Mass
nt	Nucleotide
PAGE	Polyacrylamide Gel Electrophoresis
PCR	Polymerase Chain Reaction
PE	Primer Extension
PMF	Peptide Mass Fingerprint
Pol I	DNA polymerase I
Pol II	DNA polymerase II
Pol III	DNA polymerase III
Pol IV	DNA polymerase IV
Pol V	DNA polymerase V
Pol β	DNA Polymerase β
PR	Proofreading Product
qPCR	Quantitative polymerase chain reaction

doi:10.6342/NTU202401466

RNA	Ribonucleic Acid
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
SHR	Somatic Hypermutation
SMUG1	Single-Stranded Selective Monofunctional Uracil DNA Glycosylase
SNP	Single-Nucleotide Polymorphism
ss	Single Stranded
T	Thymine
TBE	Tris/Borate/EDTA buffer
TE	Tris- EDTA
ThT	Thioflavin T
TOF	Time-Of-Flight
UDG	Uracil DNA Glycosylase
UNG	Uracil N-Glycosylase

doi:10.6342/NTU202401466

Chapter I: Introduction

1.1 Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry (MALDI-TOF MS) for DNA Detection

Matrix-assisted laser desorption ionization mass (MALDI MS) was initially employed in 1985 for analyzing and detecting amino acids. The strong laser pulse would cause large molecular damage and fractionation, making this method unsuitable for analyzing the intact substance [1]. This technique was developed in 1987 by Japanese engineer Koichi Tanaka, refining the matrix component for detecting high-mass molecular compounds, leading him to earn the Nobel Prize in Chemistry in 2002.

MALDI MS analysis begins with the preparation of the sample, involving mixing or coating it with a solution of an energy-absorbent organic compound, such as glycerin and cobalt metal powder, known as the matrix. The analyte is co-crystalized with the matrix and ionized with a laser beam, which induces analyte desorption and ionization leading to the formation of singly protonated ions [2]. These protonated ions can be separated by different mass-to-charge ratios (m/z) under the fixed potential fields [3]. These results can be analyzed and detected by specialized analyzers, such as quadrupole mass analyzers, ion trap analyzers, and time-of-flight (TOF) analyzers [4].

doi:10.6342/NTU202401466

The TOF analyzer determines the mass-to-charge ratio of detected ions by measuring the time to reach the detector. The traveling time of the ions only depends on the molecular weight when all the analytes carry the same charge. Some TOF analyzers include an ion mirror at the rear end of the flight tube, reflecting ions through the tube to a detector [5]. This ion mirror not only extends the length of the flight tube but also corrects minor differences in energy among ions. Therefore, the lighter ion reaches the detector first then the heavier ion follows [3].

TOF mass analyzers are mainly used in microbiology [3,6]. Utilizing the TOF information, a characteristic spectrum known as the peptide mass fingerprint (PMF) is generated for analytes in the sample. Identification of unknown microbes using MALDI-TOF MS involves either comparing the PMF to the database or matching the masses of biomarkers from the unknown organism with the proteome database. In the last decade, MALDI-TOF MS has been widely used in clinical diagnostic assays for microbiology and virology [6,7]. High sensitivity, high resolution, and fast speed scan are the advantages compared to other analyzers.

Alterations in DNA methylation and DNA adducts are strongly associated with human cancers and can serve as indicators for disease prognosis and diagnosis. MALDI-TOF MS has been used for analysis of DNA methylation [8,9], DNA adducts

doi:10.6342/NTU202401466

[10,11], and SNP detections [12] in many pathological conditions. The MALDI-MS technique for SNP detection dubbed Pinpoint assay was first described by Haff and Smirnov in 1997 [13]. The protocol of Pinpoint assay involved three steps: SNP primer design, primer extension, and MALDI-TOF MS analysis. The SNP primer was designed for detection of the polymorphic site at the 3' terminus adjacent to the primer. During the primer extension reaction, DNA polymerase polymerized one 2'3-dideoxyribonucleoside 5'-triphosphates (ddNTPs) to match the template strand. The extension products were separated based on their molecular weight by MALDI-TOF MS, which could detect a 1 Da difference in mass [14]. We took advantage of a high-throughput and high-resolution commercially developed MALDI-TOF MS method for clinical diagnostic application of SNP detection and streamline the procedure to develop several assays for DNA maintenance proteins [15,16].

We have developed a DNA analysis assay by using MALDI-TOF MS. In 2018, we issued a MALDI-TOF MS assay to understanding *Escherichia coli* (*E. coli)* DNA polymerase I (Pol I) proofreading and DNA repair mechanism [17,18]. The labeled free substrates were used in assays. We took advantage of the high resolution of MALDI-TOF MS to understand the Pol I proofread patch and process. A versatile design of the DNA substrate was prepared, which contained a transition mutation at different

doi:10.6342/NTU202401466

positions from the 3’ terminus at the primer-template junction. The proofreading products and excision products can be easily separated on the mass spectrum. The results revealed that internal mismatches located 2-4 nucleotides from the 3’ terminus could be proofread by Pol I, while mismatches located 6-9 nucleotides from the 3’ ends escaped the proofreading activity. Furthermore, we subjected the deoxyinosine (dI) lesion at 3’ penultimate of the primer for mimicking *E. coli* endonuclease V (Endo V) nicking products. The findings indicated that T-I proved to be a more effective substrate compared to G-I and A-I, but C-I was resistance to repair.

We then used this MS-based *in vitro* proofreading assay to explore the involvement of Pol I in the deoxyinosine (dI) repair pathway. In *E. coli*, Endo V served as the core enzyme to repair the dI lesion in DNA. Yet, the detailed mechanism of the dI repair pathway is unclear. By employing MALDI-TOF MS for the analysis of single-nucleotide extension products, we demonstrated that the Pol I proofreading exonuclease removed only 2-nt nucleotides upstream of the Endo V incision site. This result further refined our understanding of the role of Pol I in the Endo V-dependent dI-repair pathway. Our conclusion highlights the crucial role of the proficient exonuclease activity of DNA polymerase I in the deoxyinosine repair pathway [19].

doi:10.6342/NTU202401466

These noteworthy discoveries offer a systematic approach for studying related DNA repair proteins and mechanisms using MALDI-TOF MS as a high-resolution tool. In this study, we utilized the MALDI-TOF MS coupled with unlabeled DNA oligonucleotide substrates to study the activity of Pol I proofreading, *E. coli* uracil DNA glycosylase (UDG), and *E. coli* Endo V.

In the first part of the study, we focus on the DNA Pol I proofread activity to the insertion/deletion error. In the second part of the study, we established a new detection method to analyze *E. coli* UDG activity. In the last part of the study, we attempted to identify the possible interaction between Endo V and Pol I in the repair pathway of deoxyinosine.

1.2 Indel Error and DNA Proofreading

1.2.1 Insertion/ Deletion Error

During DNA replication, the insertion or deletion of nucleotides can cause frameshift mutations. In frameshift mutagenesis, misalignment of DNA intermediates leads to the addition of one or more nucleotides to either the template or the primer strands. The exact placement of these extra nucleotides can be affected by the sequence

doi:10.6342/NTU202401466

context at the template-primer junctions, potentially leading to errors in the polymerase catalytic cycle [20-22].

Two well-known theories explain the insertion/deletion mutations resulting from DNA misalignment. The first theory, known as the slipped-strand mispairing theory, was introduced by Streisinger et al. in 1966 [23]. It suggests that repetitive DNA sequences are prone to slipping during replication. This slippage leads to misalignment and the potential insertion of an incorrect nucleotide at the end of the primer-template junction. Such a mismatch could impede DNA elongation due to the resulting protruding structure, necessitating intervention by the proofreading process. However, if the inserted base aligns with the template upon DNA realignment, DNA polymerase may resume synthesis, causing the insertion/deletion error after replication [24].

The second theory involves error-prone DNA synthesis, particularly associated with some members of the Y-family DNA polymerases. These polymerases are known to foster insertion/deletion errors. A study of their crystal structures revealed a mechanism by which error-prone and lesion bypass synthesis occur. This mechanism is involved with a loose interaction between the polymerase and the replicating base pair, allowing for the concurrent incorporation of two template bases into the active site [25]. Such

doi:10.6342/NTU202401466

errors can lead to genetic rearrangements and microsatellite instability, with potential implications for disease development [20,26].

1.2.2 Correction of Insertion/Deletion Error

In *E. coli*, the fidelity of DNA replication is safeguarded by several mechanisms, including (1) base selection through the geometric specificity for correct base pairing, (2) 3→5' exonuclease proofreading activity for the correction of replication errors, and (3) post-replication mismatch repair. Both base selection and 3→5' exonuclease proofreading activity are intrinsic functions of DNA polymerase [27]. DNA polymerases with a functional proofreading exonuclease domain demonstrate lower error rates in frameshift mutations. This suggests that misaligned DNA intermediates containing insertion-deletion (indel) errors can be corrected through the proofreading process [28].

However, DNA intermediates that are misaligned but stabilized by a matching 3' primer terminus can be extended by a DNA polymerase and escape from proofreading. Extrahelical nucleotides located a significant distance from the 3' terminus of the primer are less likely to be corrected by the 3→5' exonuclease activity, thus escaping proofreading [28].

doi:10.6342/NTU202401466

To determine the mechanism of extrahelical structure fidelity, it is important to investigate the capacity of DNA polymerase to process misaligned primer-template structures. The *E. coli* DNA polymerase I (Pol I) is the prototype enzyme of family A DNA polymerase. The Klenow fragment (KF) of DNA polymerase I is an outstanding model system for studying structure and function due to the availability of its crystal structure and its complex with a DNA substrate [29]. The features of DNA polymerase I and its proofreading mechanism are further discussed in section 1.2.3.

1.2.3 DNA Polymerases in *E. coli*

DNA polymerases play important roles in DNA replication, synthesis, and proofreading processes. These enzymes remove mismatched DNA in the mismatch repair pathway (MMR), as well as the damaged bases in both the base excision repair pathway (BER) and nucleotide excision repair pathway (NER) [30].

In *E. coli*, five distinct DNA polymerases have been identified, designated as I through V. DNA Polymerase I (Pol I), the first to be discovered in the 1950s [31], is encoded by the *polA* gene [32]. It is the most abundant polymerase in *E. coli*, with an estimated 400 molecules per cell, making it the most thoroughly researched and characterized among the polymerases isolated from *E. coli* [33].

doi:10.6342/NTU202401466

DNA polymerase II (Pol II), encoded by the *polB* (*dinA*) gene, was discovered in 1970 [34]. It is a high-fidelity enzyme that possesses both polymerization and a 3'→5' exonuclease active site. Pol II is involved in various DNA repair pathways such as the repair of oxidized DNA lesions, UV-induced DNA damage, and restarting the replication after DNA lesion [35]. Its concentration increases sevenfold under the SOS response stimulation, from a baseline level of approximately 30 molecules per cell [36-38]. During the replication process, Pol II interacts with the β clamp, and serves as a backup to DNA polymerase III (Pol III) through polymerase switching, thereby ensuring the continuity and accuracy of DNA replication [39].

The DNA Polymerase III holoenzyme (Pol III) was discovered by Thomas Kornberg and Malcolm Gefter in the 1970s [40]. It is a high-fidelity replicative enzyme in *E. coli* and features a rapid synthesis rate of approximately 1000 bases per second [41,42]. Pol III is a multi-subunit enzyme composed of ten subunits. Its structure includes a central core, including the α, ε, and θ subunits, which possessed both the 5'→3' polymerase activity and the 3'→5' exonuclease proofreading function. Additionally, the complex includes the β2 sliding clamp and the clamp loader complex, essential for processivity and substrate specificity [43].

doi:10.6342/NTU202401466

DNA polymerase IV and V in E. coli are translesion synthesis (TLS) enzymes characterized by their lack of proofreading activity and relatively low fidelity [44,45]. Expression of Pol IV can increase up to tenfold above its basal intracellular level in response to SOS induction [46]. Pol IV is often recruited to the replication fork to address misaligned or bulged primer-template DNA structures, enabling it to bypass such lesions during DNA replication [47]. Polymerase V exhibits higher error rates and is a key player in SOS mutagenesis [48], primarily responsible for UV-induced mutagenesis in *E. coli.* This demonstrates its proficiency in circumventing lesions caused by UV radiation [49]. Together, all five DNA polymerases are crucial for DNA replication and genome stability maintenance.

1.2.4 *E. coli* DNA Polymerase I

The *E. coli* DNA Polymerase I (Pol I) polypeptide comprises two functional domains: the large domain, known as the Klenow fragment (KF), exhibits both 5'→3' polymerase activity and 3'→5' proofreading exonuclease activity; the small domain possesses 5'→3' exonuclease activity [50]. The three-dimensional structure of the Klenow fragment was determined by X-ray crystallography in 1985 [51]. Comprehensive enzymatic studies of the KF, including mutational analyses [52-54] and

doi:10.6342/NTU202401466

kinetic evaluations [55,56], have been conducted. These studies demonstrated that the dNTP polymerization region and the exonucleolytic proofreading domain are separated by about 30-35 Å [57,58].

The Klenow fragment contains a thumb, finger, and palm domain. The finger domain contacts with single-strand DNA and the incoming nucleotides, facilitating the identification and incorporation of the correct nucleotide during DNA synthesis. The thumb domain primarily contacts with 5-8 bp of double-stranded DNA, stabilizing the polymerase-DNA complex [59-61]. The palm domain, crucial for catalytic activity, binds two metal ions that are essential for nucleotide addition and DNA polymerization [58,62,63].

A tight pocket formed by the finger and palm domains only adopted the Watson-Crick base pairing geometry [64], ensuring the correct base was selected during replication [65]. The suitable structure was detected by interacting with specific hydrogen bond acceptors found at the pyrimidine O-2 and purine N-3 atoms [66]. Therefore, DNA polymerase ensures the correct base synthesis during the replication process and discriminates the errors through the proofreading process.

One of the major factors for the low error rate is attributed to the proofreading activity of DNA polymerase. If DNA polymerase encounters mismatched DNA during

doi:10.6342/NTU202401466

replication, the replicative DNA would be transferred from the polymerase active site to the 3'→5' exonuclease active site [67]. A previous study employing T7 DNA polymerase has demonstrated that the polymerase synthesizes correctly paired DNA with a rate constant of 300 s^{-1} and transfers the DNA to the 3'→5' exonuclease active site with a rate constant of 0.2 s^{-1}. The rate constant for the synthesis of dNTP into mismatched DNA by DNA polymerase is 0.01 s^{-1}, and DNA moves to the 3'→5' exonuclease active site with a rate constant of 2.3 s^{-1} [68]. Notably, when DNA polymerase encounters mismatched DNA, it tends to rectify the situation. This is manifested by a reduction in synthesis efficiency and an increase in the efficiency of transferring the DNA from the polymerase active site to the 3'→5' exonuclease active site.

The DNA proofreading process *in vitro* can be outlined in three steps. First, DNA binds to the polymerization site of DNA polymerase. Second, the 3' end of the DNA primer moves to the 3'→5' exonuclease site. Third, the exonuclease removes the incorrect nucleotide [69]. During the proofreading process, about 4 base pairs from the 3’ terminal in the DNA polymerase melted and then translocated to the 3'→5' exonuclease active site. Finally, the terminal nucleotide is excised before transferring the primer end to the polymerase active site [29,70,71]

doi:10.6342/NTU202401466

1.2.5 The Proofreading Assays

The proofreading efficiency of extrahelical structures by DNA polymerase has been studied by using different methodologies over the years. The DNA polymerase proofreading activity can be estimated by the thermodynamic-based method. The thermodynamic stability (ΔΔG) is influenced by the presence of a single base bulge, making a DNA substrate with an extrahelical structure less stable than perfectly paired DNA. The melting temperature (Tm) of the DNA duplex is determined by measuring absorbance at 10-second intervals during temperature elevation. It is noteworthy that the ranking of bulge-induced destabilization is in the following order: G > C ≫ T > A [22].

A real-time deletion mutation assay was established to examine polymerase β and μ by a fluorescent substrate. A fluorescent analog of adenine, 2-aminopurine (2-AP), base paired with thymine. The fluorescence signal alteration served as an indicator to understand the DNA stacking and structure. A decrease in fluorescence when dTTP is directly incorporated into the 2-AP template or strand slippage takes place. Conversely, the signal increased when nucleotide incorporation stabilized DNA stacking and formed the 2-AP loop out. The absorption wavelength for 2-AP is 310 nm, while its emission occurs at 370 nm. The limitations of this method include the design of the special

doi:10.6342/NTU202401466

sequence and the constraint imposed by the 2-AP-labeled DNA, which limits misalignment structures [72].

The kinetic study of proofreading was investigated using a quenched-flow assay [73]. To understand the proofreading activity on DNA substrates with various mismatched positions from the 3' terminal end, the kinetic of DNA transfer between the polymerization site and the exonuclease site of the polymerase was monitored. Changes in fluorescence signal intensity were indicative of DNA partitioning to the exonuclease site, thereby suggesting the occurrence of proofreading activity. The results revealed that a mismatch buried by 1 to 3 correct nucleotides was bound to the exonuclease site, while a mismatch buried by 4 to 10 nucleotides was bound to the exonuclease site less than 10% of the time [73]. However, the observation of DNA partitioning provided indirect evidence of proofreading activity for substrates with buried mismatches. The most straightforward method to assess proofreading activity is by quantifying the resultant proofreading products. Over the past decade, we have developed both *in vivo* and *in vitro* assays to thoroughly investigate Pol I proofreading activity.

The system for evaluating proofreading activity was first developed in 2013 [74]. This system not only illustrates the significance of proofreading activity in the *E. coli* repair mechanism of deoxyinosine but also allows us to calculate the initial rate of Pol I

doi:10.6342/NTU202401466

proofreading among all twelve mismatches at the 3’-penultimate. A circular DNA duplex was constructed with 12 mismatches, and the three mutagenic deoxyinosine lesions (A-I, G-I, and T-I). Those DNA substrates were designed to disrupt a restriction endonuclease recognition site and scored by restriction enzyme assay. The *in vitro* repair assays were performed by using either a cell-free extract or a reconstituted system with purified proteins. The results indicated that the proficient exonuclease polymerase is essential in the repair of deoxyinosine.

1.3 Base Excision Repair and Glycosylase

1.3.1 Uracil in DNA

Uracil (U) is a normal nucleotide in RNA, it replaces thymine (T) and pairs with adenine (A) during DNA transcription. However, uracil (U) is considered a lesion in DNA, leading to two types of pairings: U-A or U-G [75]. Uracil in DNA may interrupt the transcription process, including gene fidelity and transcription factor binding. It is believed that accumulation of DNA base damage is highly related to tumorgenicity, aging, and genetic disease [75,76].

The U-A pairing can occur due to the misincorporation of dUTP during semiconservative replication. Since deoxyuridine triphosphate (dUTP) is an

doi:10.6342/NTU202401466

intermediate in the cellular synthesis of deoxythymidine triphosphate (dTTP), DNA polymerase can erroneously substitute dUTP for dTTP to form a U-A mismatch [77]. The misincorporation of dUTP in the DNA replication process depends on the ratio of dTTP and dUTP in the dNTP pool. In *E. coli*, the dCTP is converted into dUTP through dCTP deaminase. During the *E. coli* replication process, the incorporation of dUTP occurred every 1500-3000 bases, and about 4000 uracil in DNA after replication [78]. In mammalian cells, the production of dUTP is less than *E. coli* due to the lack of dCTP deaminase. Therefore, the dUTP incorporation frequency in mammalian cells is a hundred times lower than *E. coli*, with only one dUTP polymerized by polymerase every 10^5 bases, and ~10^4 dUTP per genome per cell division cycle [79]. U-A pairing might not cause mutations, but it can lead to cell damage through multiple repair cycles and DNA breakage [80].

The U-G base pair in DNA arises from cytosine deamination within a C-G base pair. Cytosine deamination can be induced by heat, nitrous anhydride, sodium bisulfite, alkaline, and ultraviolet (UV) irradiation, or occur spontaneously [81]. The rate of spontaneous cytosine deamination is estimated to be between 100 and 500 events per cell per day [82] and occurs about 100 times faster on single-stranded DNA (ssDNA) than on double-stranded DNA (dsDNA) [83]. If the uracil is not repaired before DNA

doi:10.6342/NTU202401466

replication, it may lead to a G-C to A-T transition mutation [84]. To prevent mutation and maintain gene integrity, uracil can be repaired through base excision repair (BER).

1.3.2 Base Excision Repair (BER) Pathway

Base excision repair (BER) is a repair pathway that corrects DNA damages, including alkylation, deamination, and oxidative damage [85,86]. This pathway is highly conserved in both prokaryotic and eukaryotic cells. In mammalian cells, BER is initiated by DNA glycosylase, which identifies and excises specific modified nucleobases. DNA glycosylase can be classified into two types: monofunctional and bifunctional. Monofunctional DNA glycosylases hydrolyze the glycosidic bond resulting in an abasic site (AP site). The AP site is then processed by AP endonuclease 1 (APE1) which removes the AP site by generating a nick at the 5' phosphodiester bond, forming a 3' hydroxyl and 5'-deoxyribose phosphate (5'-dRP) groups [87]. In contrast, bifunctional DNA glycosylases cleave the glycosidic bond via β-elimination of the DNA backbone 3’ to the AP site and form a Schiff base, then create a single strand break with 3’-α, β-unsaturated aldehyde, and 5’-phosphate [89,90]

After the reaction of APE1, BER may proceed via one of two pathways: short-patch and long-patch pathway. In short patch BER, the 5’ dRP group is removed by the DNA

doi:10.6342/NTU202401466

polymerase β (Pol β) which possesses deoxyribose-phosphodiesterase (dRP lyase) activity, and synthesis of a complement nucleotide. Then the BER pathway is accomplished by ligase IIIα/XRCCI complex to seal the nick. For long patch BER, the Pol β or Pol δ/ε undergoes strand displacement synthesis, and flap endonuclease 1 (FEN1) removes the replaced DNA fragment, then seals the nick by ligase I/PCNA complex.

In *E. coli*, the BER pathway was searched thoroughly by Lindahl [88]. *E. coli* DNA glycosylase recognized and excised the lesion base, cleaving an N-glycosidic bond and leaving an AP site. The 5' sugar-phosphate chain at the AP site is cleavage by *E. coli* APE exonuclease III (XthA) and endonuclease IV (Nfo). Exonuclease III removes the 3' phosphoryl group then generates a 3' hydroxyl group, which is a suitable substrate for DNA Pol I gap filling [89]. Endonuclease IV generates a 3' OH group by attacking the phosphodiester bond at the 5' to the AP sites [90]. In *E. coli*, short patch BER is involved with dRP lyase, such as formamidopyrimidine DNA glycosylase (Fpg) and RecJ, to remove dRP and synthesize one nucleotide by DNA polymerase I. In long patch BER, DNA polymerase I uses its 3'→5' exonuclease activity to remove the dRP group and synthesize two or more nucleotides [91]. The displaced strand is then removed, and the resulting nick in the DNA is sealed by DNA ligase.

doi:10.6342/NTU202401466

1.3.3 Uracil DNA Glycosylase (UDG)

DNA glycosylases are a group of damage-specific enzymes, which recognize various base lesions on single or double-stranded DNA. For repairing uracil in DNA, uracil-DNA glycosylase (UDG) initiates and remove uracil from DNA, then fills the correct nucleotide with a series of downstream proteins. UDG has been widely observed across various organisms, including viruses, pathogenic bacteria, and eukaryotic and mammalian cells since its discovery in *E. coli*. It has been identified as a highly conserved superfamily through evolution, as shown in various studies [92,93]. Based on their substrate specificity, UDGs are categorized into six families [94].

In *E. coli*, Uracil N-glycosylase (UNG) is encoded by the *ung* gene and primarily removes uracil from DNA, belonging to family I UDG [95]. The U-G mismatch-specific DNA glycosylase (Mug) which recognizes U-G mismatch, belongs to family II UDG [96]. In human cells, there are at least 11 unique glycosylases and four of them mainly remove uracil: Uracil N-glycosylases (UNG, family I), single-strand-selective monofunctional uracil-DNA glycosylase (SMUG1, family III), thymine-DNA glycosylase (TDG, family II) and methyl-CpG binding DNA glycosylase (MBD4, HhH-GPD superfamily) [75,97]. Uracil on DNA is primarily removed by the *UNG* gene

doi:10.6342/NTU202401466

product, which is further divided into UNG2 in the nucleus and UNG1 in the mitochondria based on its mRNA N-terminal sequence [98]. UNG1 primarily repairs U-G mispair in mitochondria. UNG2 is cell cycle-regulated glycosylase, highly expressed in nuclear at early to mid-S phase cells, and rapidly degraded at the G2 phase. SMUG1, which has a high binding affinity toward the AP site and shows a lower catalytic turnover rate compared to UNG1/2, is suggested to play an important role in the repair of uracil in non-proliferating cells. TDG and MBD4 repair pyrimidine derivates in mismatch and uracil in the CpG sequence, respectively [86,99].

A comparison of the sequences of *E. coli* Ung and human UNG proteins shows that they share 58% of identical amino acids, revealing a remarkably high degree of sequence conservation across the two species [100]. UDG is highly specific to uracil, remove uracil from both double and single-stranded DNA with the preference: ssU>U-G>U-A.

DNA glycosylase also plays an important role in immunological diversity and takes part in somatic hypermutation (SHR) and class switch recombination (CSR) in the B cell maturation process [101]. Irregular expressions of UDG have been linked to the development of several diseases, such as lymphoma, Bloom syndrome, human

doi:10.6342/NTU202401466

immunodeficiency, and cancer [84,102]. Therefore, the detection of UDG activity is critical for clinical diagnosis.

1.3.4 5-Fluorouracil and UDG Activity

5-Fluorouracil (5-FU) was the first synthetic fluoropyrimidine analog introduced in 1957, marking a significant advancement in cancer chemotherapy [103]. It is predominantly utilized in the treatment of breast and colorectal cancers with its high efficacy [104]. The cytotoxic mechanism of 5-FU involves multiple pathways. Primarily, its incorporation into RNA and DNA as fluoronucleotides disrupt normal cellular processes. Additionally, 5-FU inhibits thymidylate synthase (TS), an essential enzyme in the nucleotide synthesis pathway, leading to a reduction in thymidine triphosphate (dTTP) pools and subsequently causing DNA damage and cell death [104].

When the misincorporation of 5-FU occurred and paired with adenine, it could be recognized and excised by DNA glycosylase [105]. Conversely, when 5-FU pairs incorrectly with guanine, it could be detected by the mismatch repair (MMR) protein MutS [106]. Previous study has demonstrated that both BER and MMR pathways are integral to the cytotoxicity and sensitivity of cells to 5-FU [106]. These pathways not only repair 5-FU-induced DNA lesions but also play roles in modulating the cellular

doi:10.6342/NTU202401466

response to 5-FU, influencing its efficacy as a chemotherapeutic agent. In vitro assays have shown that 5-FU can be removed by four major DNA glycosylases in mammalian cells, highlighting the broad capacity of the BER pathway to address 5-FU-induced damage [107]. Among these glycosylases, uracil-DNA glycosylase (UNG) and single-strand selective monofunctional uracil-DNA glycosylase 1 (SMUG1) have been the most extensively studied, underscoring their importance in the repair process [108]. These studies suggest that understanding of DNA repair pathways could lead to improvement of therapeutic strategies. Additionally, exploring the interplay between these repair mechanisms and other cellular pathways may offer new avenues for overcoming resistance to 5-FU and increasing its clinical effectiveness [109]. Therefore, developing a new strategy that combines 5-FU and glycosylase inhibitors may have great importance for the success of the therapy. Undoubtedly, establishing a fast, high-throughput, and highly sensitive methodology is crucial for inhibitor screening in DNA glycosylase.

1.3.5 The Detection Method of UDG Activity

Several assays assessing the *in vitro* base excision repair (BER) activity have been previously described and listed in Table 1. The UDG assay using isotope-labeled

doi:10.6342/NTU202401466

substrate was first described in 1981 [110]. The uracil excised from radioactively labeled DNA substrates and free uracil can be separated from the DNA by different solubility in ethanol then followed by chromatography analysis. However, the use of radioactive materials in this method poses significant hazards. Additionally, the preparation process is both time-consuming and labor-intensive.

Fluorescence-labeled oligonucleotide assays combined with gel electrophoresis have been devised as an alternative to radioisotope labeling [111]. Fluoresce labeled oligonucleotide was used for measuring glycosylase activity, containing a base lesion in the restriction enzyme sequence. The base lesion was removed by glycosylase activity and repaired by a following repair protein. The glycosylase activity was analyzed by restriction enzyme digestion and resolved by gel electrophoresis. In this assay, the reaction buffer should be taken into consideration when analyzing different glycosylase activity.

The fluorescently labeled substrate is also widely used in fluorescent resonance energy transfer (FRET) based assay. The technique employs two complementary oligonucleotides containing multiple U-A pairings. One is labeled with a 3' donor fluorophore and the other with a 3' acceptor fluorophore. The stable complex dissociates after the uracil is excised by a glycosylase reaction. Monitoring the reaction by

doi:10.6342/NTU202401466

fluorometer, energy transfer from the donor fluorophore to the acceptor fluorophore resulting in fluorescence emission could be detected [112]. However, the specific labeling requirements limit substrate design and high background levels remain problematic.

There are more UDG detection methods mentioned with special substrate designs. The colorimetric-based method with a G-quadruplex DNA probe was developed in 2015, and utilized on UDG activity in 2021 [113,114]. The substrate contained multiple U-G pairings with 3' poly dG and it formed a G-quadruplex structure when the uracil was removed by glycosylase reaction. The G-quadruplex structure would then bind with a fluorescent substance such as thioflavin T (ThT) or heme. Lastly, the fluorescent complex would then covalently bind to redox indicators and could be detected at the specific wavelength for measuring UDG activity. However, the precipitation of redox substances may cause a high background.

The electrochemical-based method was reported in 2016 [115]. Multiple A-U pairing DNA duplex dissociated after UDG excised the uracil. Single-stranded DNA can be absorbed onto the graphene-modified glassy carbon (GC) electrodes, causing the current response as the signal indicator for evaluation of UDG activity. Nonetheless, the preparation of graphene-modified GC electrode is time-consuming and hard to perform.

doi:10.6342/NTU202401466

Recently, a PCR-based assay was developed to measure uracil DNA-glycosylase activity, providing a rapid, label-free, and quantitative *in vitro* assay [116]. A U-containing DNA substrate was prepared by randomly incorporating dUTP into the DNA. The cell or tissue extract was mixed with this U-containing DNA, followed by qPCR to measure glycosylase activity. However, a disadvantage of using this protocol is low specificity.

Recently, the use of enzyme-free electrochemical biosensor-based UDG detection assay was reported [117]. The UDG detection assay involves three steps. First, a metal nanoparticle was prepared using nickel (Ni), cobalt (Co), and single-stranded U-containing DNA to form NiCoP@PtCu@sDNAs nanosheets. Second, a metal-linked DNA substrate was subjected to a glycosylase reaction. The released products replaced the protector on the capture probe anchored to the electrode surface. Finally, changes in the electrical signal proportional to UDG activity could be measured. The authors claimed that this method featured high sensitivity. However, the application of this protocol seems to be limited by its tedious substrate preparation.

Taking into account the limitations of the methods described above, there is a need to develop a simple, accurate, highly sensitive, and cost-effective assay for UDG activity.

doi:10.6342/NTU202401466

1.4 Repair of Deoxyinosine and *E. coli* Endonuclease V

1.4.1 Deoxyinosine in DNA

Deoxyinosine (dI) is an abnormal nucleoside with hypoxanthine serving as its base component. Most of the deoxyinosine present in DNA arises from the spontaneous deamination of deoxyadenosine residues, or catalysis by adenosine deaminase or AMP deaminase, leading to the formation of a dI-dT base pair. Deamination of dATP results in dITP, as it has a similar structure to deoxyguanosine, causing dI-dC pairing. Meanwhile, dITP pairing with dATP or dGTP during DNA synthesis is a rare event [118]. In *E. coli*, the incorporation of dITP is non-mutagenic since DNA polymerase recognizes it as guanine and stably pairs with cytosine through the near hydrogen bond. The deamination events of adenine occurred only 2% compared to that of cytosine [119]. LC/MS-MS quantification of damaged DNA products indicates that deoxyinosine could be identified at baseline levels ranging from 1 to 10 per 10^6 nucleotides in tissues or cells [120]. It would cause an A-T to G-C transition mutation if deoxyinosine is unrepaired before replication.

In the *de novo* purine nucleotide metabolism pathway, inosine monophosphate (IMP) is a normal intermediate metabolite in which hypoxanthine is the base moiety.

doi:10.6342/NTU202401466

Previous studies have found that inosine triphosphate pyrophosphatase (ITPase) encoded by the *ITPA* gene hydrolyzes inosine triphosphate (ITP) and dITP into IMP and dIMP with similar efficacy in mammalian cells. This process reduces the vital dysfunction caused by the accumulation of inosine derivatives [121,122]. Hence, it is crucial to repair the presence of deoxyinosine in DNA to ensure the preservation of genome integrity.

1.4.2 Repair of Deoxyinosine in DNA

Studies have indicated that deoxyinosine (dI) produced in DNA can be removed by various DNA repair systems in both prokaryote and eukaryote cells. In *Escherichia coli*, 3-methyladenine DNA glycosylase II (AlkA) encoded by the *AlkA* gene featured with board substrate recognition includes hypoxanthine, oxidized guanine, 8-oxoguanine, 5-hydroxymethyluracil, 5-formyluracil,3-methyladenine, 7-methylguanine, 1,N^6-ethenoadenine (εA), and other alkylated purine bases [123]. The AlkA protein in *E. coli* excises the hypoxanthine base from dI in DNA, creating an apurinic/apyrimidinic (AP) site and initiating a base excision repair (BER) pathway [124]. The homology of the AlkA protein is highly conserved over the organism, MAG in *S. cerevisiae* and AAG in *H. sapiens.* The AlkA protein acts on all base pairs, including A-I, C-I, G-I, and T-I,

doi:10.6342/NTU202401466

without exhibiting clear specificity [125,126]. However, *E. coli* AlkA protein excises hypoxanthine much slower than other alkylated DNA, while mammalian AlkA protein removes hypoxanthine efficiently [127].

In *E. coli*, the alternative excision repair (AER) pathway triggered by endonuclease V (Endo V), which is encoded by the *nfi* gene [128], has been demonstrated as the primary pathway for processing deoxyinosine (dI) both in vivo and in vitro [129,130]. In AER, Endo V identifies and nicks the DNA at the second phosphodiester bond 3' to the dI without releasing the damaged base [131]. *E. coli* Endo V is thought to require another enzyme to accomplish dI repair.

These comprehensive views highlight the complexity and efficiency of DNA repair mechanisms in dealing with deoxyinosine. While AlkA and Endo V provide critical pathways for repairing dI, the differences in efficiency and mechanisms between prokaryotic and eukaryotic systems underscore the evolutionary adaptations of these organisms to maintain genomic integrity. Further research into these pathways could provide deeper insights into the prevention and treatment of genetic diseases related to DNA repair deficiencies.

1.4.3 *E. coli* Endonuclease V

doi:10.6342/NTU202401466

Endonuclease V (*nfi* gene, M.W.:24.9 kDa) was first discovered as the fifth endonuclease from *E. coli* K12 strain deficient in endonuclease I by Gates and Linn in 1977 [132]. It has been described as exhibiting an optimal pH at an alkaline level of 9.5, necessitates Mg^{2+}, and generates 3'-hydroxy and 5'-phosphate termini. *E. coli* Endo V also nicked the duplex DNA exposed to OsO4, UV light, or acid. Uracil contains DNA and single-stranded DNA served as a good substrate for nicking, suggesting that the Endo V could trigger the base excision repair [133].

The deoxyinosine 3' endonuclease activity was isolated and purified by Min Yao [131,134,135]. In 2012, Weiss and Kow cloned and sequenced *nfi* gene, confirmed that the deoxyinosine 3' endonuclease is identical to Endo V [133]. Endo V features a broad substrate specificity including hypoxanthine, cleave AP, uracil sites, base mismatches, flap, pseudo Y structures, and small insertions/deletions [134,135]. The Endo V binds the major groove of the dI-containing DNA duplex and then nicks the DNA at the 3’ of deoxyinosine without removing the lesion base. The binding and nicking activity is Mg^{2+} dependent [136]. Additionally, *E. coli* Endo V formed a stable complex with dI-containing DNA due to the 6-keto group of hypoxanthine from deoxyinosine [137]. The substrate specificity was determined by examining the nicking efficiency of oligonucleotide substrate. For double-stranded dI-containing DNA, Endo V had a four-

doi:10.6342/NTU202401466

fold preference over single-stranded DNA, and showed a similar preference for dI paired with A, T, C, or G. Furthermore, Endo V shows a similar affinity toward the substrate and nicked products, proposing that the DNA-protein complex is a target to other repair proteins and initiates an unconventional repair pathway [134].

1.4.4 Homologs of Endonuclease V

Endonuclease V is evolutionary conserve in prokaryotic and eukaryotic cells. The structure of Endo V has been investigated intensively by using the thermophilic bacterium *Thermotoga maritime* Endo V [138]. The 3’ exonuclease activity and non-specific 5’ exonuclease activity of Tma Endo V were measured by fluorescently labeled substrate [139]. The *Salmonella typhimurium* Endo V exhibits a strong affinity for deaminated bases and is the only variant of Endo V that has demonstrated noticeable binding to DNA containing deoxyxanthosine [140].

The substrate preference of homologs of Endo V was altered from DNA in bacteria to RNA in eukaryotes. Human Endo V (M.W.: 50.2 kDa) shares 30% sequence identity with bacterial Endo V and has been identified as a nucleolar protein that recognizes and binds to branched DNA structures [141]. As a deoxyinosine endonuclease, human Endo V cleaves dI on single- and double-stranded DNA, exhibiting a preference order of ssI >

doi:10.6342/NTU202401466

G-I > T-I > A-I > C-I [142]. Human Endo V also cleaves the inosine on RNA suggesting that Endo V plays an important role in RNA metabolism [143]. It has been reported that the activity or expression level of human Endo V is correlated to atherosclerosis [144], hepatocellular carcinoma [145], and re-localization of cytoplasmic stress granules [146], all of them show the biological significance of Endo V.

1.4.5 The Endonuclease V Repair Pathway

The repair pathway of deoxyinosine by Endo V remained unclear until the 2010s. Lee. et al. conducted an *in vitro* assay by dI-containing DNA duplex and demonstrated that the dI repair pathway could be facilitated by a maltose binding protein fusion Endo V, DNA polymerase I, and DNA ligase. By using bacterial cell-free extract, the removal of the dI lesion is highly dependent on the presence of the *nfi* gene [129]. The removal of dI lesion in DNA relies on the 3'-5' exonuclease activity of DNA polymerase, and dI at 3' penultimate can be corrected efficiently through the proofreading process by DNA polymerase I, with the preference I-T ≅ I-G > I-A> I-C. The repair of I-C was refractory to DNA polymerase I [74]. Considering the available evidence, a working model for the multistep Endo V repair pathway has been proposed. The dI repair is

doi:10.6342/NTU202401466

initiated by Endo V, which creates a strand break at the second phosphodiester bond 3’ to the lesion. The 3’-5’ exonuclease activity of Pol I removes the 2 nucleotides including the 3’ terminal nucleotide and the 3’ penultimate dI lesion, resulting in a small 2-nt gap. Subsequently, repair synthesis by Pol I fills in this 2-nt gap, and the nick is sealed by DNA ligase to complete the repair process [19].

The strong binding affinity to a DNA duplex containing dI seems to fulfill a significant biological function. The formation of a protein-DNA complex might be important for efficiently targeting the subsequent protein, which is Pol I, to remove the lesion. However, due to the high binding affinity of Endo V after nicking the substrate, determining the kinetic parameters of Endo V is challenging. Moreover, it is still not clear whether the Endo V-DNA complex acts as a signal to Pol I.

Establishing a more effective approach is necessary for comprehending the detailed mechanism and interaction between Endo V and Pol I in the Endo V repair pathway.

1.5 Research Motivation and Strategy

Several DNA polymerase proofreading analyses have been conducted using MALDI-TOF MS [17-19]. Based on previous findings, it is hypothesized that the KF can directly contact the four bases at the end of the DNA primer, and therefore has a

doi:10.6342/NTU202401466

better ability to correct mismatches at the first to fourth bases from the primer-template junction.

This method offers several notable advantages. First, label-free substrates help reduce costs. Second, the approach allows for rapid and high-throughput testing. A single MALDI-TOF MS chip has the capacity to detect up to 384 objects, aligning well with the clinical requirements. Third, the method demonstrates high resolution, capable of analyzing differences in 1 Dalton samples. Considering these three aspects collectively, we believe that MALDI-TOF MS is a highly effective approach for detecting the activity of DNA repair proteins. Our goal is to establish a method that achieves rapid, large-scale output while requiring less labor. In this study, we used MALDI-TOF MS to analyze three DNA maintenance proteins. The first part is the evaluation of the proofreading capacity for the insertion/deletion error at the 3’ primer-template junction. The second part is to establish a new detection method for uracil DNA glycosylase enzymatic by using MALDI-TOF MS. The last but not least, is to analyze the Endo V and polymerase I interaction through MALDI-TOF MS analysis. The graphic abstract is shown in Figure 1.

doi:10.6342/NTU202401466

Chapter II Materials and Methods

2.1 Materials

2.1.1 Synthetic DNA Oligonucleotides

Non-labeled oligonucleotides with standard desalting procedures were purchased from Integrated DNA Technologies (Singapore) and Genomics Inc. (Taiwan). The fluorescently labeled oligonucleotides purified by HPLC were purchased from Synbio Tech Inc. (Taiwan) and Genomics Inc. (Taiwan). ESI-MS checked the quality of synthetic oligonucleotides or MALDI-TOF MS. The oligonucleotide pellet was dissolved in TE solution (1mM EDTA, 10mM Tris-Cl, pH=7.6) to prepare 100μM stock by following the data sheets provided by the manufacturer. The accuracy of the concentration was subsequently verified by measuring the absorbance of the oligonucleotide solution at $\lambda = 260$ nm. The sequence of nucleic acid for proofreading assay is listed in Table 2, for uracil glycosylase activity assay is listed in Table 3, and for endonuclease V activity assay is listed in Table 4,5 and 6.

2.1.2 Enzymes

DNA Polymerase I, Large (Klenow) Fragment (KF, cat#M0210L, 5000U/mL), Uracil-DNA Glycosylase (UDG, cat#M0280L, 5000U/mL), Uracil Glycosylase

doi:10.6342/NTU202401466

Inhibitor (UGI, cat#M0281S, 2000U/mL), Endonuclease V (Endo V, cat#M0305S, 10000U/mL) were purchased from New England BioLabs.

2.1.3 Reagents

The dideoxy ribonucleotides triphosphate sets (ddNTPs, including ddATP, ddTTP, ddCTP, ddGTP) were purchased from Trilink Biotechnologies (San Diego, CA, US). The deoxyribonucleotide triphosphate sets (dNTPs, including dATP, dTTP, dCTP, and dGTP) were purchased from Clubio (Guishan Dist, Taoyuan City, Taiwan).

Tris/Borate/EDTA(TBE) buffer was used in the urea-PAGE assay. Sample loading dye contained 90% formamide, 0.5% EDTA, 0.1% xylene cyanol, and 0.1% bromophenol blue. The fixation solution contained 10% methanol, 20% ethanol, and 0.5X TBE buffer.

2.2 DNA Duplex Substrate Preparation

In a previous study of MALDI-TOF MS analysis on DNA oligonucleotide, we observed that 20-50 pmol of DNA oligonucleotide resulted in the generation of remarkably strong and reliable signals [17]. These signals were indicative of successful analysis and provided valuable information about the molecular characteristics of the

doi:10.6342/NTU202401466

oligonucleotides under investigation. Equal molar amounts of primers and templates were mixed in a 1.5 mL microcentrifuge tube, incubated at 65 ºC for 30 minutes, then at 37 ºC for 30 minutes, and finally put on ice for proper annealing.

To confirm the formation of a canonical heteroduplex resulting from the annealing of the substrate and template strands, we employed A260 nm measurements both before and after hybridization process (Table 7). These measurements offer numerical evidence of the hybridization process, confirming the proper formation of the targeted double-stranded complex.

2.3 DNA Polymerase Proofreading Assay for MALDI-TOF MS

Equal molar amounts of primers and templates were mixed in a 1.5 mL microcentrifuge tube, incubated at 65 ºC for 30 minutes, then at 37 ºC for 30 minutes, and finally put on ice for proper annealing. The 20μL reaction mixture contained 50 pmol (2.5μM) insertion/deletion-DNA substrate, 0.1mM of ddNTPs or dNTPs mixture, 50mM NaCl, 10mM $MgCl_2$, 1mM dithiothreitol, 10mM Tris-HCl (pH 7.9 at 25 ºC). Reactions were started by adding 2U KF and the reaction was performed at 37 ºC, then quenched by adding 2μL 0.25M HCl and put on ice for 6 min, to drop down the pH

doi:10.6342/NTU202401466

ratio of the mixture. The reaction mixture was neutralized by 2μL 0.23M Tris base, then a 6μL TE solution was added. The reaction mixture tubes were stored at -20 ºC before MS analysis.

2.4 Uracil DNA Glycosylase Assay for MALDI-TOF MS

The DNA substrate preparation is described in section 2.1. For *E. coli* uracil DNA glycosylase assay, a 10μL reaction mixture in a 1.5 mL microcentrifuge tube containing 50 pmol (5μM) uracil-containing substrate, 20 mM Tris-HCl, 1mM dithiothreitol, 1mM EDTA (pH 8.0 at 25 ºC). Reactions were started by adding a designated unit of UDG and performed at 37 ºC. Reactions were quenched by adding 1μL 0.25M HCl and put on ice for 6 min, to drop down the pH ratio of the mixture. Neutralized the mixture by 1μL 0.23M Tris base, then added 15μL TE solution. The reaction mixture tubes were stored at -20 ºC before MS analysis.

2.5 Endonuclease V Nicking Assay for MALDI-TOF MS

The DNA substrate preparation is described in section 2.1. A 10μL Endo V nicking assay was performed in 1.5 mL microcentrifuge tube containing 20 pmol (2 μM) inosine-containing substrate, 50 mM Potassium Acetate, 20 mM Tris-acetate, 10 mM

doi:10.6342/NTU202401466

Magnesium Acetate, 1mM DTT (pH 7.0 at 25 ºC). Reactions were started by adding a designated unit of Endo V and performed at 37 ºC. The following procedure is the same as the UDG glycosylase assay.

2.6 MALDI-TOF MS Analysis

The MALDI-TOF MS analysis procedure included sample spotting, MALDI-TOF MS operation, and data analysis. Before starting the MS assay, a Microsoft Excel file containing the m/z value of DNA substrates and expected products were imported to the AssayDesigner module of Type4 application program (Agena Bioscience, CA). The 30μL of the reaction mixture was transferred to a 384-well plate, then desalted mixture by resin and rotated at room temperature for 15 min. Then palleted down the resin by centrifuge, 3200g, for 5 minutes. The 384-well plate was then loaded onto the nanoliter dispenser MassARRAY system (Agena Bioscience, CA) and ready for sample spotting. Each sample would be dispensed from a 384-well plate to a matrix chip in the 5 to 10 nL range. Checked the monitor to ensure every sample was spotted on the chip with proper volume. The chip was then loaded into the mass spectrometer (Agena Bioscience, CA).

For the MALDI-TOF MS operation, the chip ID and operation program were set up and linked to the mass spectrometer. Press the run button on the MS machine to initiate

doi:10.6342/NTU202401466

the mass spectrometer, after which the spectra are automatically gained and saved. Each sample receives 20 laser shots to acquire mass spectra, and the spot on the analysis software is marked in green when the spectral signal matches the expected m/z ratio.

2.7 MS Data Analysis

The Type 4 software (Agena Bioscience, CA) was used for raw data analysis. The mass spectra graph could be customized by the analysis program, with the X-axis representing the range of m/z ratio and the Y-axis showing the upper and lower limits of signal peak intensity. Signal intensity and peak height could be measured on the screen by moving the cursor over the specific peak. The data were exported in JPEG format with dimensions of 1600x1200.

2.8 Fluorescence-Based Gel Assay

The assay was performed as described previously [147] with minor modifications DNA samples were prepared by serial dilution of the FAM-40I oligonucleotide. The 1 μL of diluted DNA and 2 μL of sample loading dye were mixed in a 1.5 mL microcentrifuge tube and then heated at 95°C for 5 minutes. The sample tubes were then placed on ice and quickly spun down. The DNA was resolved on a 20%

doi:10.6342/NTU202401466

polyacrylamide-8M urea gel and electrophoresed for 20 minutes using 1X TBE running buffer at 180 V. After electrophoresis, the gel was transferred to a fixation solution to fix the DNA and then destained. The gel was visualized using the ChemiDoc XRS+ System (Bio-Rad), and the relative signal intensity was quantified using Image Lab Software (Bio-Rad).

2.9 Detection of DNA Possessing Enzymes by Fluorescently Labeled DNA Duplex

Equal molar amounts of FAM-labeled oligonucleotides and non-labeled template oligonucleotides were mixed in a 1.5 mL microcentrifuge tube and incubated at 65°C for 30 minutes, then at 37°C for another 30 minutes. For proper annealing, we placed the tube on ice for 5 minutes after incubation. For the 40bp-IG, IA, IC substrates, FAM-40I was annealed with 40G, 40A, and 40C, respectively. For the 21bp-IT substrate, FAM-21I was annealed with 4SIT. The nicking reaction mixture, containing 10 pmol of dI substrate in a 10 μL volume, also included 50 mM potassium acetate, 10 mM magnesium acetate, 20 mM Tris-acetate, 1 mM dithiothreitol (pH 7.9 at 25°C), and various units of endonuclease V. The nicking assay was performed at 37°C and heat-inactivated at 65°C for 20 minutes, followed by the addition of 2 μL sample loading dye

doi:10.6342/NTU202401466

and heating at 95°C for 10 minutes. The DNA was resolved by Urea-PAGE as described in section 2.8.

doi:10.6342/NTU202401466

Chapter III Result

3.1 Proofreading of Single Nucleotide Insertion/deletion Replication Errors Analyzed by MALDI-TOF Mass Spectrometry Assay

*** The results in part I were published in DNA repair, 2020. doi: 10.1016/j.dnarep.2020.102810.**

3.1.1 The Indel DNA Substrate Design and Reaction Condition

In previous study, the mismatch at the position of 5-nt from the 3’ end could be proofread proficiently by KF [18]. It suggested that the thermostability of the DNA duplex affected the proofreading efficiency to mismatch substrates. To further investigate the proofreading efficiency for internal errors, we designed a series of indel error substrates. In this study, the indel error proofreading activity was measured by MS-based analysis. The indel substrate design was based on previous proofreading assay for internal mismatches [18]. The same 28-nt template (T28) and a 21-mer primer (P21A) were used for the template design. Table 2 listed all the oligonucleotide indel substrates utilized in this study. The 29/21-mer primer-template formed the insertion substrate, and the 28/22-mer primer-template formed the deletion substrate. The proofreading assay methodology was illustrated in Figure 2. Previous research observed an inverse relationship between peak intensity and the mass-to-charge (m/z) ratio of the

analyte, indicating MS signal intensity should be normalized [19]. Therefore, the same DNA sequences were used in this study to maintain consistent parameters in proofreading signal calculations.

The substrate design aimed to investigate the impact of unpaired extrahelical nucleotides on polymerase proofreading activity (Table 2). Single-nucleotide bulges were introduced on either the template or primer strands to evaluate proofreading activity. DNA substrates, designated as Pdel N (N=1 to 9)/P21A, were created by incorporating an unpaired pyrimidine loop on the template strand of a normally perfectly paired 28/21-mer template-primer. These structures could result in deletion mutations if not corrected before further replication. Conversely, DNA substrates designated as T28/Pins N (N=1 to 9) were produced by adding an extra unpaired pyrimidine to the primer strand of an otherwise perfectly paired 28/21-mer template-primer, representing potential intermediates of insertion mutations. Identifying proofreading products for single-nucleotide deletion errors is straightforward, as shown in Figure 2. The polymerase KF exonuclease removed misaligned nucleotides from the 3’ terminus. Once the unpaired nucleotide was removed, the polymerase underwent DNA synthesis and added a matched ddNTP in protocol A, or added matched dNTP in protocol B, resulting in proofreading products. To terminate the reaction, an acid

doi:10.6342/NTU202401466

quenching protocol was employed instead of hazardous phenol/chloroform reagents. The reaction mixture was stored at -20°C before the MALDI-TOF analysis.

3.1.2 Single Nucleotide Deletion Error at 1 to 6-nt from 3' Terminus Can be Proofread by Pol I with 4ddNTPs

As the spectrum shows in Figure 3, Pdel 1 contained a single nucleotide bulge at the 3' terminus, with the last nucleotide of the P21 primer removed. Subsequently, ddGTP was added, yielding a 21-nt proofreading product (PR). Similarly, for Pdel 2, the last 2 nucleotides were removed, and ddGTP was added, resulting in a 20-nt PR, and this pattern continued accordingly. Utilizing high-resolution MALDI-TOF MS, with precision as low as 1 Da, the signals of primers, proofreading products, and excision intermediates were distinctly identified by their corresponding m/z values on the mass spectrum. The sum of all signals on the spectrum was 100%, and changes in each signal could be calculated and normalized. As shown in Figure 3, a clear pattern in proofreading capacity emerged: deletion errors within 1 to 5 nucleotides from the 3' terminus of the primers were efficiently corrected by KF, with correction percentages ranging from 98% to 29% (Fig. 3, PR signals for Pdel 1 through 5), without primer extension products. Pdel 6 displayed both proofreading and primer extension, albeit at

doi:10.6342/NTU202401466

low levels. In contrast, deletions positioned at 7, 8, and 9 nucleotides from the 3' terminus showed no proofreading activity; instead, 22% to 84% of these primers underwent extension (Fig. 3, primer extension (PE) signals for Pdel 7, 8, and 9).

The excision products (Ex-1 and Ex-2 signals) were observed in the spectrum for Pdel 3, 4, 5, 6, and 7. These were induced by ddNTPs in the reaction, which acted as inhibitors in the polymerase synthesis [17]. To substantiate this model in the following section, we designed a reaction involving 3 dNTPs and ddCTP. This setup facilitated primer extension over several nucleotides at a regular replication rate before reaching the chain terminator ddCTP.

3.1.3 Single Nucleotide Deletion Error at 1 to 5-nt from 3' Terminus Can be Proofread by Pol I with 3ddNTPs

Kinetic studies showed that the excision rate of DNA polymerase for terminal mismatches was 100 times higher than for matched base pairs. This kinetic barrier prevented the excision process on correctly paired DNA, avoiding idle turnover and blocking the elongation of mismatched DNA [148]. Additionally, this barrier caused DNA polymerase to stall on the DNA, increasing the idle-turnover that a mis-paired primer would be transferred to the excision active site. Furthermore, the instability of

doi:10.6342/NTU202401466

the heterologous DNA duplex facilitated the entry of the partially melted 3’ single-stranded primer end into the exonuclease site [149].

We used three canonical dNMPs and one ddNTP to overcome the non-specific excision intermediated in the proofreading reaction. According to the template sequence, the dATP, dTTP, dGTP, and ddCTP were used in the proofreading assay as shown in Figure 2 protocol B. As the spectrum in Figure 3 and the sequences listed in Table 2, if the deletion substrates were able to be proofread, the loop-out nucleotides would be removed and incorporated 4-nt of dNMPs to form a 25-nt proofreading product (PR). The 24-nt primer extension product (PE) was formed by adding 2 dNMPs and one ddC.

The proofreading efficiency improved when reacting with 3 dNTPs. The correction rate for single nucleotide deletion errors within 4 nucleotides from the 3’ end was 100%. However, the proofreading efficiency decreased to 56% for Pdel 5 and 6% for Pdel 6. No proofreading (PR) products were detected in Pdel 7, 8, and 9, with over 90% of primer extension (PE) products were observed.

3.1.4 Single Nucleotide Insertion Error at 1 to 5-nt from 3’ Terminus Can be Proofread by Pol I with 4ddNTPs

doi:10.6342/NTU202401466

As shown in Figure 5, we found a noticeable trend in proofreading capacity: single nucleotide insertion errors from Pins 1 to 5 were effectively proofread by KF, with correction efficiencies ranging from 100% to 51% (Table 2 and the PR signal in Figure 5), with no apparent primer extension. However, Pins 6 showed 23% extension products and only 18% proofreading. This reaction also resulted in a significant number of non-specific 1-nt excision products, alongside unprocessed substrate. In the cases of Pins 7, 8, and 9, 86%–92% of the input primers were extended (Fig. 5, PE signals for Pins 7, 8, and 9), and no proofreading product was detected.

3.1.5 Single Nucleotide Insertion Error at 1 to 5-nt from 3' Terminus Can be Proofread by Pol I with 3dNTPs

The dATP, dTTP, and dGTP were used in the proofreading assay for analyzing the insertion substrate, which extended several nucleotides after the removal of the insertion error. As shown in Figure 6, the 24-nt proofreading product (PR) was formed by removing 2 nucleotides from the 22-nt primer, followed by the synthesis of 3-nt dNMP. The 24nt-primer extension product was formed by 2-nt dNMP, without the proofreading process. 100% primer correction was observed in Pins 1 to Pins 5, and 100% primer

doi:10.6342/NTU202401466

extension was observed in Pins 7 to Pins 9. A partial correction was observed at Pin 6, with 53% PR and 47% PE, and no unreacted primer.

3.1.6 The Ability of DNA Polymerase I to Correct Indel Errors

Using our newly developed MALDI-TOF MS proofreading assay, we examined indel error substrates located up to 9 nucleotides from the 3’ terminus of the primer to determine the full range of indel proofreading by KF polymerase. Two sets of experiments were conducted under different conditions to evaluate proofreading effectiveness: one using 4 ddNTPs to induce chain termination and another with 3 dNTPs to induced chain termination. Interestingly, the proofreading ranges determined from both methods were consistent. As illustrated in the schematic in Figure 7, when an indel error results from polymerase misincorporation or template-primer misalignment, DNA synthesis stalls, enabling the primer to be transferred to the 3'-5' exonuclease site for proofreading. Indel errors occurring within 1-4 nucleotides from the primer’s 3’ end likely initiate proofreading, possibly because this region remains in contact with the polymerase [59]. However, indel errors located 5 and 6 nucleotides from the 3’ end of the primer are only partially proofread, likely due to their limited contact with the polymerase. Indel errors more than 6 nucleotides from the primer’s 3’ terminus escape

doi:10.6342/NTU202401466

proofreading, probably because these indel errors lose contact with the polymerase, allowing the primer terminus to remain in the polymerase site for continued DNA synthesis and potentially caused insertion or deletion mutation.

3.2 Measurement of Uracil-DNA Glycosylase Activity by Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry Technique

***The results in part II were published in DNA repair, 2021. doi: 10.1016/j.dnarep.2020.103028.**

3.2.1 The Scheme of Uracil DNA Glycosylase Activity Assay

The UDG assay was established as illustrated in Figure 8A, with the resulting spectrum shown in Figure 8B. Synthetic oligonucleotides of the lesion strand contained one uracil and annealed with a template strand, forming a duplex containing a single G-U mismatch. UDG excised the uracil from the U-containing DNA, creating the abasic (AP) products. The reaction was terminated by phenol/chloroform, or acid quenching/tris neutralization. Then the reaction mixture would be subject to MS analysis (Fig. 8A).

A typical mass spectrum of the UDG reaction can be seen in Figure 8B. The enzyme blank control was conducted only with a U-containing DNA duplex (Figure 8B,

doi:10.6342/NTU202401466

upper), and the excess UDG (0.5U) with incubation of 30 minutes could make a complete enzyme digestion (Figure 8B, lower). The MS brought at high-resolution spectra which separate the AP product signal (m/z = 5447.6) from the substrate signal (m/z = 5541.6) based on the mass difference resulting from the hydrolysis of uracil from DNA (Δm/z = -94).

The open aldehyde form of the deoxyribose sugars at the AP site is labile, as the 3' phosphodiester bonds can be hydrolyzed through a β-elimination reaction induced by alkaline pH and elevated temperature [150]. However, the stability of deoxyribose residues during the MALDI process was not reported. The termination methods employed, including phenol/chloroform extraction and acid quenching followed by Tris neutralization, were selected to ensure the reaction stopped promptly and the products remained stable for MS analysis. This careful approach allowed for the precise measurement of UDG activity, as illustrated in the mass spectra shown in Figure 8B. No detectable signals of the fractured AP products were found at corresponding m/z values in the spectrum. Using the signal from the 19-nt template DNA as a reference, the relative intensity of the AP-product signal after complete digestion was comparable to the signal from the unprocessed U-substrate.

doi:10.6342/NTU202401466

The relative quantification of products within each test was used in our MS method. This approach has been employed to calculate the mutation allele frequency in the previous studies [12,151]. Each glycosylase reaction contained 50 pmol of U-containing substrate as confirmed in Section 2.1.1. The glycosylase activity can be calculated by following formula:

$$\text{Glycosylase activity} = \frac{AP\ product\ signal}{U\ substrate\ signal + AP\ product\ signal} \times U\ substrate$$

3.2.2 UDG Kinetic Parameters Determined by MS-based Assay were Comparable to Traditional Method

A time course analysis was conducted using various UDG concentrations, ranging from 0.01 Unit (3.2 pM) to 0.1 Unit (32 pM). As shown in Figure 9, the reaction demonstrated both dose and time dependency. The activity rates were linear across the range of UDG concentrations (0.01 Unit to 0.1 Unit) within a 3-minute incubation period. It was determined that 3.2 pM of UDG (0.1 Unit per 100 μL reaction) was appropriate for determining the Km and Kcat kinetic parameters.

The UDG activity was measured using a G-U substrate prepared with an 18-nt U-containing strand and a 19-nt template. The G-U substrate was mixed with UDG and incubated at 37℃ for 30 and 60 seconds, followed by phenol/chloroform extraction.

doi:10.6342/NTU202401466

The reaction products from five substrate concentrations, ranging from 10 to 200 nM, were analyzed using MALDI-TOF MS (Figure 10A).

The relative quantities of product and substrate were determined by MS signal intensity and then converted into concentrations (nM). The Michaelis-Menten curve was created with initial velocity (v) represented on the vertical axis and substrate concentration (nM) on the horizontal axis. The resulting graph yielded the Lineweaver-Burk plot by plotting 1/[S] on the y-axis and 1/v on the x-axis. In this curve, the y-intercept corresponds to 1/Vmax, while the x-intercept represents -1/Km. The kinetic parameters Km = 50 nM, Vmax = 0.98 nM s^{-1}, and Kcat = 9.31 s^{-1}, respectively, were tested at least three independent experiments (Figure 10B). Our findings are comparable to the earlier studies using the conventional ^{3}H-uracil releasing assay which reported values were Km = 40 nM and Kcat = 13.3 s^{-1} [152].

The acid quenching protocol was found to be more suitable for the automation process in MALDI-TOF MS, as opposed to phenol/chloroform extraction, which involves phase separation. To adjust the reaction pH to 2, HCl was introduced, followed by a 6-minute incubation on ice and subsequent neutralization with tris-base until reaching pH 6.8. Notably, there were no indications of base loss or modification, and

doi:10.6342/NTU202401466

the reaction levels exhibited no significant differences between these two methods of terminating the reaction.

Based on the unit definition provided by the manufacturer, one unit of UDG is capable of catalyzing 60 pmol of uracil per minute from an uracil-contained double-stranded DNA substrate. The measurement of enzyme activity involved the excised of tritium-labeled uracil in a 50 μL reaction, with 0.2 μg of PBS1 DNA (equivalent to 230 pmol of dUMP), for 30 minutes at 37 ◦C. Our MS-based method utilized a 10 μL reaction containing specifically a 50 pmol (approximately 0.56 μg DNA) 18/19-nt G-U DNA duplex. The UDG was diluted with buffer (composed of, 20 mM Tris-HCl (pH 7.5), 30 mM NaCl, 0.5 mM EDTA, 1 mM dithiothreitol, and 50% glycerol) on ice, and enzyme activity was assessed through a 30-minute incubation at 37 ◦C. As shown in Figure 11, a plot was generated comparing the MALDI-TOF MS measured unit to the manufacturer-defined unit within the range of 0.001 U to 0.02 U. The results showed that the MS-determined unit exhibited a proportional relationship with the UDG concentration, represented by the correlation equation $y = 0.933x + 0.0003$, and a high correlation coefficient ($R^2 = 0.9974$). We believe that this method is comparable to the conventional approach using ^{3}H-uracil labeling. The detection limit of UDG in our method is 0.001 U, with a coefficient of variation (CV%) of 9.2% based on 5

doi:10.6342/NTU202401466

independent tests. Given that the typical application of UDG is at the unit level [153], we believe that this assay offers adequate sensitivity.

3.2.3 Investigation of UDG Substrate Specificity and Glycosylase Activity with Varied Uracil Positions in DNA Oligonucleotides

In a prior publication, it was demonstrated that UDG catalyzes the liberation of uracil from single-stranded or double-stranded DNA substrates [154]. Various uracil-containing substrates were tested for analyzed substrate specificity of UDG and listed in Table 3. When uracil is positioned in the middle of the oligonucleotide substrates, the glycosylase activity toward the single-stranded substrate (ssU+9) is 3.7 times higher than that toward the double-stranded G-U substrate (G-U) (Table 3, ssU and G-U entries). In comparison to G-U, the reaction rate for A-U was significantly decreased, exhibiting an almost threefold difference (Table 3, G-U and A-U entries).

Typically, the glycosylase substrate was designed with a single uracil near the center of oligonucleotides. This study investigated how glycosylase activity is influenced by varying the position of uracil within DNA substrates. Uracil was tested at the ultimate, penultimate, and antepenultimate positions at the 5’ end (named G-U5’+1,

doi:10.6342/NTU202401466

G-U5'+2, and G-U5'+3) as well as the 3' end (named G-U3'+1, G-U3'+2, and G-U3'+3). We found out the G-U5'+1, G-U3'-1, and G-U3'-2 substrate were not cleaved by UDG (Table 3, G-U5'+1, G-U3'-1, G-U3'-2 entries), these results were compatible with the previous finding [155,156]. Notably, G-U5'+3 and G-U5'+2 exhibited a higher reaction rate compared to the uracil in the center (G-U), showing a 2-fold increase for the G-U5'+2 and a 1.5-fold increase for the t G-U5'+3 (Table 3, G-U5'+2 and G-U5'+3 versus G-U). G-U3'-3 substrate could be removed by UDG, albeit the efficiency was approximately 8% less than G-U substrate.

Based on the results and the kinetic data in Table 3, we suggest that the G-U duplex is a suitable standard substrate for analyzing glycosylase activity. The G-U duplex appears to be more biologically relevant than the A-U duplex due to the natural occurrence of deamination in the C-G base pairing. While higher glycosylase activity has been found in ssU and U at the 5' terminal substrates, however these substrates are uncommon in *E. coli*. Taken together, the G-U substrate can be used in both general glycosylase assays and inhibitor screening assays.

3.2.4 The Inhibitor Screen Assay of Uracil DNA Glycosylase

doi:10.6342/NTU202401466

To demonstrate the potential of the MALDI-TOF MS based analysis in pharmaceutical usage, we used UDG inhibitor (UGI) in our inhibition assay. the UGI is a small protein (9.5 kDa) made from *Bacillus subtilis* bacteriophage PBS1 which effectively inhibits E. coli uracil-DNA glycosylase (UDG) and UDG from other species. This inhibition occurs through reversible protein binding, maintaining a 1:1 UGD:UGI stoichiometry, and leading to the dissociation of UDG-DNA complexes [157,158]. The impact of UGI as shown in Figure 12. In our assay conditions, UDG (0.05 U, 8 pM) was effectively inhibited by UGI at a concentration of 0.05 Unit (100 pM), with no detectable signal of AP product observed in the corresponding spectrum. The inhibition curve was plotted in Figure.12, and the IC50 was measured to be 7.6 pM.

All the experiments detailed above were conducted by using microcentrifuge tubes and manual pipetting. In total, we conducted less than 30 tests for UDG reaction in each batch. Our samples were analyzed concurrently with routine genetic tests from clinical trial samples at the NCFPB Integrated Core Facility for Functional Genomics and NRPB Pharmacogenomics Lab. However, using established automatic processes it could be possible to analyze 300 samples by 384-well microplate within 2 hours. This would be suitable for pharmaceutical usage.

doi:10.6342/NTU202401466

3.3 Study of *E. coli* Endonuclease V Enzyme Activity by MS-based and PAGE-based Analysis

3.3.1 The Scheme of Endonuclease V Nicking Assay by MALDI-TOF MS Analysis

Based on previous works, we have demonstrated the application of MALDI-TOF MS-based assay on DNA processing proteins by analyzing DNA intermediates. Then MALDI-TOF MS was further used to study Endo V with label-free substrates. DNA substrates ranging from 15 to 32 nucleotides were designed to provide the best resolution for MS analysis [17,19]. The shorter substrate would melt after Endo V nicking, potentially leading to the turnover of Endo V to another dI substrate. Although the turnover activity of Endo V may not be biologically relevant, it has been suggested that the high affinity of Endo V for the dI lesion could serve as a signal for polymerase I to remove the dI [136]. However, using convenient artificial substrates for analyzing enzyme activity is common practice in biochemistry and clinical chemistry. Therefore, seeking a highly efficient turnover rate and sensitive assay for Endo V could be beneficial for both research purposes and pharmaceutical industry applications.

The mass signal should be normalized due to the inverse correlation between peak intensity and analyte mass. Therefore, the DNA primer signal intensity was tested under

doi:10.6342/NTU202401466

the same concentration. The spectrum was shown in Figure 13A and all the MS signals in this study were normalized according to the average relative intensity as shown in Figure 13B. The normalization factors were obtained through the calculation of the intensity of all DNA signals relative to the signal intensity of the 15-nt DNA ratio. For instance, the intensity of 17-nt DNA was 69% of 15-nt DNA intensity. Therefore, the signal intensity of 17-nt DNA should be multiplied by a normalization factor of 1.44.

The substrate design was shown in Figure 14A. The 21-IT duplex contained one dI lesion at in the 6th nucleotide from the 3' end. Endo V recognized and generated a nick and formed a 17-nt nicking product (NP). The reaction mixture was then subjected to MS analysis. To calculate the nicking efficiency, the signal intensity on the spectrum should be normalized by the relative ratio according to oligonucleotide lengths. As explained above, the signal intensity of 17-nt nicking products and 21-nt dI lesion strands should be normalized by factors of 1.44 and 2.56, respectively. The nicking efficiency of Endo V was formulated as the following equation:

$$Endo\ V\ nicking\ activity = \frac{\text{nicking products signal intensity}/1.44}{\text{dI subtrate signal intensity}\ /2.56 + \text{nicking products signal intensity}/1.44} x\ input\ DNA(pmol)$$

Two conditions for the Endo V nicking assay were tested, and the spectrum is shown in Figure 14B. An input of 20 pmol 21-IT duplex and 2U Endo V (Figure 13B,

doi:10.6342/NTU202401466

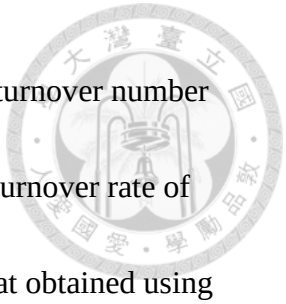
upper panel) yielded a clearer signal compared to 5 pmol 21-IT duplex and 0.5U Endo V (Figure 13B, lower panel). Therefore, 20 pmol of 21-IT DNA substrate was used for the standard nicking reaction.

3.3.2 Enzymatic Analysis of Endonuclease V by MS-based Analysis

A time course analysis was conducted using 0.5U (0.3875 pmol) or 1U (0.775 pmol) Endo V reacted with 20 pmol 21-IT duplex. As shown in the spectrum in Figure 15A and the line plot in Figure 15B, the reaction exhibited dose dependency but no time dependency. This phenomenon can be explained by the previously reported high affinity of the dI binding activity [136]. According to the manufacturer's unit definition of Endo V, one unit of Endo V is capable of nicking 1 pmol of 34bp dI containing DNA duplex in 1 hour incubation time [159]. The measurement of enzyme activity involved the process of 10 pmol of FAM-labeled DNA oligonucleotide in a 10 μL reaction, for 60 minutes at 37 °C. The MS-based method utilized a 10 μL reaction with 20 pmol 21/28-nt IT DNA duplex.

It was observed that Endo V exhibited slow turnover activity toward the 21-IT substrate. Under the manufacturer's assay conditions, the turnover number (Kcat) was

doi:10.6342/NTU202401466



0.02 min^{-1}. However, when measured using the MS-based assay, the turnover number (Kcat) of Endo V was found to be 0.52 min^{-1}. This indicates that the turnover rate of Endo V analyzed by the MS-based assay was 25 times higher than that obtained using the manufacturer's assay conditions.

A sensitivity analysis was conducted to compare the MS-based assay with the manufacturer's detection assay. Endo V was diluted with a buffer composed of 10 mM Tris-HCl (pH 8), 25 mM NaCl, 0.1 mM EDTA, 1 mM dithiothreitol, 200 μg/ml BSA, and 50% glycerol on ice. Enzyme activity was then measured after a 1 hour incubation at 37°C [159]. According to the manufacturer's definition, one unit of Endo V produces 1 pmol of nicking product from 10 pmol of a 34-bp dI-containing substrate within 1-hour incubation in a 10 μL reaction volume [159]. Figure 15C shows a positive correlation between the MS-determined units and the concentration of Endo V, and the relationship is represented by the equation $y = 16.09x - 0.1318$. The high correlation coefficient ($R^2 = 0.9869$) indicates a strong relationship within the concentrations ranging from 0.2 U to 1 U of Endo V. The nicking activity measured by the MS-based analysis was 16 times higher than the manufacturer's FAM-labeled assay at the same Endo V concentration, indicating that MS-based analysis was much more sensitive. It will be discussed in the later section.

doi:10.6342/NTU202401466

3.3.3 The Substrate Specificity of Endonuclease V Nicking Activity

In a previous study, Endo V showed similar affinity toward deoxyinosine pairing with four different nucleotides [136]. However, the substrate specificity cannot be interpreted well due to the non-turnover condition of the Endo V nicking assay. To clarify this, a 21-nt dI-containing oligonucleotide was annealed with a 28-nt template strand to prepare four different DNA duplexes: 21-IT, 21-IA, 21-IG, and 21-IC. The reaction was terminated at various time points, with representative spectra shown in Figure 16A and the time course results presented as a line plot in Figure 16B. The Endo V nicking preference was: I-G> I-A≒I-T> I-C. The I-G pairing is the best substrate for Endo V. Endo V demonstrated a nicking preference in the following order: I-G > I-A ≒ I-T > I-C. The I-G pairing was identified as the optimal substrate for Endo V, likely due to the instability of the I-G base pair, which makes it more susceptible to melting during the nicking assay. Additionally, the nicking activity on single-stranded DNA was eight times higher than on double-stranded DNA, similar to the previous finding that Endo V have a preference on single stranded DNA substrate [160].

doi:10.6342/NTU202401466

3.3.4 The Fluoresce-based Analysis for Endonuclease V Nicking Assay

In the previous section, a label-free Endo V nicking assay utilizing MALDI-TOF MS was presented, demonstrating its efficacy as a robust tool for enzymatic analysis with high turnover and sensitivity. To investigate the protein-protein interaction between DNA polymerase I and Endo V, a non-turnover model was established. The longer substrate would not easily melt after nicking and would form a stable complex with Endo V, which seems to be a more natural condition. In our previous work, we employed the 40-bp DNA duplex to analyze the Pol I and Endo V turnover reaction. The 40-bp substrate can be presented as a non-turnover model, preventing Endo V dissociation from the DNA duplex after nicking. However, oligonucleotides longer than 32-nt were prone to encounter poor resolution and fewer sensitivity issues during MS analysis [124]. Thus, it is essential to shift towards a better technique for understanding the Endo V turnover issue.

For longer substrates, oligonucleotides in the DNA duplex can be labeled with fluorescein, and the resulting compounds resolved using PAGE analysis [100]. To elucidate the biological relevance of Endo V activity, a nicking assay with fluorescence-based analysis was designed.

doi:10.6342/NTU202401466

First, the detection limit of fluorescein-labeled oligonucleotides was tested. As demonstrated in Figures 17A and 17B, the fluorescence intensity of various amounts of FAM-labeled oligonucleotides was measured, and the relative intensity of each band was calculated. The results indicated that fluorescence intensity saturated at 20 pmol, making input of 10 pmol suitable for relative quantification. The relative intensity showed a high correlation with the amount of DNA input, with an R^2 value of 0.9813 and a regression equation of $Y = 0.1036X - 0.005606$, where Y represents the relative intensity and X represents the DNA input. This approach is deemed suitable for relative quantification in Endo V nicking analysis.

Next, the nicking activity of Endo V on the 40 bp DNA duplex was tested. As shown in Figure 18A, a nick was first generated on the 40IT DNA duplex using Endo V, and the 19-nt nicking products were then resolved using polyacrylamide gel electrophoresis. The signal intensity of FAM-40I was calculated relative to the intensity of the nicking products for quantification. Only 1.4 pmol of nicking products were generated by 1U Endo V within 60 minutes of incubation. The PAGE system was then compared to the manufacturer's assay condition, which contained FAM-labeled 34 bp DNA in a 10 μL reaction volume. Based on the results in Figure 18B, the PAGE

doi:10.6342/NTU202401466

analysis system was found to be comparable to the manufacturer's system, with a regression equation of $Y = 0.8216X + 0.3774$ and an R^2 value of 0.9541.

3.3.5 Endonuclease V Nicking Activity Analyzed by The Fluorescence-based Assay

In the time course analysis, Endo V generated 1 pmol of nicking products per unit (Endo V 1U =0.775 pmol) within 60 minutes of incubation and showed dose dependency (Figure 19A). The linear plot demonstrated that Endo V quickly bound and nicked the dI substrate within the first 10 minutes, after which the nicking rate slowed down, showing a slow-turnover model.

The Endo V showed no nicking preference toward the 4 kinds of dI substrates, pairing with either A, T, C, or G within 10 minutes (Figure 19B). The nicking activity with the single-stranded dI substrate was approximately 2.5 times higher than that with the double-stranded I-containing substrate after 60 minutes of incubation. Based on these results, we suggested that the stability of the dI DNA duplex affects the nicking efficiency. The ssDNA structure is more favorable for nicking activity.

doi:10.6342/NTU202401466

3.3.6 The Exonuclease Activity of Pol I Promotes Endonuclease V Turnover

We analyzed the Endo V and Pol I turnover activity by MS-based assay [161]. Only the nicking products could be detected on the spectrum due to the detection range limitation. Thus, we used an internal calibrator for relative quantification in the long substrate MS-based assay. The result shown that KF promotes Endo V turnover 2.6 times, indicating a potential interaction between two proteins [161]. However, the low resolution and sensitivity of detecting long oligonucleotides by MS makes it challenging to interpret the results. Therefore, we attempted to utilize PAGE-based analysis to improve the measurement of the interaction between the two proteins by using a longer DNA duplex.

Two experiments were designed to test the interaction between Endo V and Pol I, with the results shown in Figure 20A and B. In Figure 20A, the turnover assay was performed under the reaction conditions of 20 pmol of a 40-bp dI DNA duplex (40IT) and 1U Endo V, along with 1U of either KF or KF (exo-). The KF (exo-) lacks 3' to 5'-exonuclease activity while maintaining normal polymerase activity [57], KF(exo-) may recognize and possess nicked DNA. The results indicated that the turnover activity was promoted by approximately 1.5 times in the presence of KF. Interestingly, we found

doi:10.6342/NTU202401466

that KF (exo-) cannot promote Endo V turnover. Even using 10 times the amount of KF (exo-), no difference was observed.

In Figure 20B, the Pol I variants were spiked after 15 minutes of Endo V incubation and then prolonged incubation time to 30 minutes. Surprisingly, the processed dI products increased 1.2-fold when spiked with KF. The KF (exo-) promoted Endo V turnover 1.1 times as well.

These findings revealed that Pol I KF, possessing 3' to 5' exonuclease activity, increases Endo V turnover by 1.5 times, whereas Pol I lack exonuclease activity exhibited lower turnover. These results imply a direct interaction between Pol I and Endo V during the repair process.

doi:10.6342/NTU202401466

Chapter IV Discussion

4.1 The MS-based Analysis of DNA Processive Proteins

In this study, we demonstrate three MS-based analyses of DNA processive proteins. We utilize non-labeled oligonucleotides with a flexible substrate design to understand the proofreading capacity. Furthermore, we demonstrate the potential pharmaceutical usage of MS assay for inhibitor screening by employing UDG and UGI. Additionally, we define the enzymatic characteristics of Endo V using a high-turnover substrate.

The high throughput and automation process make the MS-based analysis more suitable for biomedical usage than other methods. We believe that our results have demonstrated the potential application in biomedical usage.

4.1.1 The Prospective of MALDI-TOF MS Analysis in Biomedical Research and Development

The MS-based method is particularly suitable for inhibitor screening in biomedical research due to its efficiency, automatically, and scalability. This approach is easily scalable; while the current study manually performed around 30 tests per day using micropipettes and microcentrifuge tubes, the capacity can be significantly increased. By adapting an automated pipetting system for a 384-well microplate format, the daily

doi:10.6342/NTU202401466

output could be substantially boosted. Therefore, with automation, the streamlined process maintains the same turnaround time of 2 hours for 300 UDG MALDI-TOF MS assays. Thus, MS-based approach demonstrates a great potential in the process of screening and evaluating new therapeutic compounds to combat chemoresistance of base analogues.

4.1.2 Mass-to-Charge Ratio Consideration of DNA Substrates and Products Analysis

Several precautions should be taken into consideration when applying MS-based analysis, specifically when designing a set of experiments. Due to the feature of MALDI-TOF MS, the signal intensity of the analyte is inversely related to the molecular weight of the analyte. Our assay was first adapted from the Pinpoint assay which is used for SNP detection [16]. When detecting SNPs, using well-designed primers improves outcomes during MALDI-TOF analysis. Primer lengths ranging from 15 to 28 nucleotides are suitable for both primer extension and MALDI-TOF analysis [16]. Therefore, the best resolution for DNA oligonucleotides is in the range of 4000 to 9000 m/z, corresponding to oligonucleotides of 15-nt to 28-nt. Oligonucleotides shorter than 15-nt or longer than 28-nt might encounter difficulties in accurately quantifying the

doi:10.6342/NTU202401466

signal observed on the spectrum. To gain a better result of the spectrum, we typically use 20 to 50 pmol of oligonucleotides in the MS-based assays, which is relatively abundant compared to traditional isotope-labeled assays. We adapted an MS protocol from a clinical application for the detection of point mutations, and the clinical samples were amplified through the PCR process before being subjected to MS analysis. Since our experimental design did not include DNA amplification, thus it is required to use a higher amount of DNA substrate to gain higher MS signals. Consequently, more enzymes may be required in the reaction.

4.1.3 Adducts Interference and Intervention in Measurement

We noticed the monovalent alkaline metal (Na^+ and K^+) adducts signal followed behind the major signal on the spectrum in the proofreading assay and Endo V nicking assay, which came from the reaction buffer containing sodium and potassium, respectively. The *E. coli* UDG glycosylase activity is non-ion dependent; no cation adducts could be found in the analysis results. Adducts have a negative impact on the sensitivity of MS-based analyses, causing the ion signal to shift to higher m/z values, thereby interfering with the accurate determination of molecular weights. The soft ionization process in MALDI analysis causes less sample fragmentation and a larger

doi:10.6342/NTU202401466

cation adduct peak [162,163]. Also, the negative charge of the phosphate backbone tends to form salt adducts with cation. To avoid cation adducts, the samples were desalted using resin before being subjected to MS analysis. However, the 50 mM of salt in the reaction cannot be fully removed through the resin cleaning protocol. The clean-up kit is a useful tool for removing sodium or potassium ions from the reaction, resulting in a clear signal on the spectrum [164].

4.2 The Different Outcomes of 4 ddNTPs and 3 dNTPs in Proofreading Assay

In part I study, we aimed to investigate how extrahelical bases in DNA influence DNA polymerase proofreading, based on Streisinger's strand-slippage hypothesis. We designed fixed extrahelical bulges to explore the full range of the KF polymerase's ability to edit extrahelical nucleotides within the primer-template, instead of the slippage strand caused by repetitive sequence.

We first observed that the proofreading range defined by using 3 dNTPs and 4 ddNTPs generally displayed consistency, though there were some subtle differences (Table 2). Regarding deletion errors with 4 ddNTPs, we noted that single nucleotide deletions up to 6 nucleotides from the 3’ terminus could be proofread, with the

doi:10.6342/NTU202401466

efficiency gradually decreasing from Pdel 1 to Pdel 6. However, from Pdel 7 to 9, KF was unable to proofread and instead extended a significant portion of the substrates. When using 3 dNTPs, KF was able to proofread Pdel 1 to 6, but unlike the 4 ddNTPs reaction, a considerable amount of substrate extension was also observed for Pdel 5 and 6.

Likewise, for single nucleotide insertion errors, the proofreading efficiency of KF paralleled that of deletion errors. Using 4 ddNTPs, single nucleotide insertions up to 6 nucleotides from the primer terminus could be proofread, with efficiency decreasing from Pins 1 to Pins 6. However, for Pins 7, 8, and 9, KF failed to proofread and instead extended a high proportion of the substrates. When using 3 dNTPs, KF exhibited higher correction levels for Pins 1 to 5 but yielded different product outcomes compared to the ddNTP reactions, especially for Pins 6.

This phenomenon may be explained by the kinetic barrier inhibiting erroneous elongation and causing the polymerase to stall on the DNA. In the context of T4 polymerase proofreading, as described by the presteady-state kinetic model [148,165], the polymerase encounters a mismatch at primer 3’ end, leading to a temporary halt. Subsequently, the polymerase induces partial melting of the primer-terminus, facilitating the transfer of the primer from the active site to the exonuclease site. This

doi:10.6342/NTU202401466

transfer constitutes a rate-limiting step with a rate constant of 4 s^{-1}. As the primer moves towards the exonuclease site, further strand separation occurs, compelling the primer-terminus to bind at the exonuclease active site with a faster rate constant of 20 s^{-1}. Excision takes place through the hydrolysis of the phosphodiester bond at a remarkable rate constant of 100 s^{-1}, followed by the fast repositioning of the primer back into the polymerase's active site. The kinetic analysis reveals a rate constant of 400 s^{-1} for extending a correct base pair [68]. This kinetic barrier not only prevents erroneous primer elongation but also facilitates idle turnover—a repetitive cycle involving the extension and excision of a nucleotide. This activity effectively stalls the DNA polymerase on DNA, contributing to the fidelity of DNA synthesis.

The difference observed between reactions utilizing 4 ddNTPs and those utilizing 3 dNTPs for KF processing Pdel 5, 6, and Pins 6 can be attributed to a kinetic obstacle formed by the addition of ddNTPs. ddNTPs demonstrate lower efficiency as substrates for Pol I compared to dNTPs. While dNTPs synthesis with a rate constant of approximately 80 s^{-1}, ddNTPs exhibit a significantly slower extension rate constant of 0.015 s^{-1} [166]. Consequently, the comprehensive proofreading rate observed when employing ddNTPs was notably inferior compared to using a singular dNTP [17].

doi:10.6342/NTU202401466

The study provides insights into the complex interplay between DNA polymerase, extrahelical nucleotides, and proofreading mechanisms, shedding light on the factors influencing DNA replication fidelity.

4.3 The Contact of KF and DNA Junction Affects the Proofreading Efficiency

While the crystal structure of DNA bound to the Klenow fragment (KF) editing site has been resolved and extensively discussed in several studies [71,166,167], the crystal structure of the KF-DNA complex at the polymerization site remains elusive. Consequently, certain widely accepted structural characteristics of DNA-bound KF are derived from extrapolations from T7 polymerase structures, which belong to the same family as Pol I [168].

Structure studies have revealed that T7 DNA polymerase interacts with the DNA backbone approximately 7 base pairs upstream of the primer terminus [168]. These interactions with the DNA duplex are believed to enable the polymerase to detect geometric distortions of the helix caused by mispairs and/or extra nucleotides [169]. As a result, single nucleotide insertion and deletion (indel) errors occurring within the primer-template region in close proximity to the polymerase are likely to be corrected.

doi:10.6342/NTU202401466

The T7 DNA polymerase studies have laid a foundation for these insights, highlighting the importance of geometric distortions in error detection.

Our findings support this model, as indel errors of Pdel 1–4 and Pins 1–5 substrates fall within the range of contact with the KF polymerase. This conclusion compares with our previous results regarding KF proofreading of internal mispairs, which indicated that mismatches located within the last 4 base pairs from the primer terminus could be effectively corrected [18].

4.4 Acid Termination Method Offers Safer Alternative to Phenol/Chloroform Extraction in Glycosylase Assay

Our MS-based method for analyzing glycosylase activity, including quantification, prioritized the integrity of AP products. Loss of AP product signal on the spectrum due to breakage during the reaction process or MS analysis can introduce bias in quantification. As evidenced in our results, we effectively managed the UDG reaction without AP product breakdown.

It is known that *E. coli* UDG does not require divalent cations; thus, the reaction cannot be terminated by using chelating agents such as EDTA. Conventional gel-based glycosylase assays employ highly alkaline conditions to terminate the reaction and

doi:10.6342/NTU202401466

induce strand breaks at the AP site for gel analysis [156,170]. The recommended UDG inactivation method involves phenol/chloroform extraction, preventing DNA degradation. However, this protocol included a laborious phase separation of hazardous chemicals. We introduced the acid termination method with HCl, effectively inactivating UDG by lowering the reaction pH to 2 ± 0.5. Subsequent neutralization with Tris base to prevent DNA damage. MS analysis of the AP product within 30 hours revealed no signs of base loss or modification in the mass spectra (Figure 7A). Extended storage (beyond 10 days) of reaction products led to some instances of non-specific depurination of DNA or strand breaks at AP sites (data not shown).

4.5 Comparison of UDG Kinetic Parameters Using MALDI-TOF MS and Traditional Radioisotope Methods

It's worth noting that the kinetic parameters of UDG and enzyme units derived from the G-U reaction, analyzed using the MALDI-TOF MS method (Figure 9A and 9B), were remarkably similar to those obtained using the conventional ^{3}H-uracil labeled measurement developed by Lindahl et al [152]. However, our oligo-duplex with a single G-U substrate is quite different from the PBS1 plasmid-based multiple A-U substrate. Specifically, UDG demonstrated three times higher activity with G-U compared to A-U

doi:10.6342/NTU202401466

(Table 3, Initial rate and Kcat of G-U versus A-U). To reconcile this seemingly contradictory evidence, our explanation is as follows: The PBS1 substrate contains a very high density of multiple A-U (36% of nucleotides are U [88], compared to only 3% in our G-U substrate). It has also been shown that UDG is a very "processive" enzyme, meaning multiple uracil cleavages could have occurred in a single protein-DNA binding event [171]. Taken together, this would explain the lower specific activity of UDG towards the A-U substrate while generating similar results to a single G-U substrate. Nevertheless, the finding that there is an almost 1:1 correlation of UDG measurements between these two methods (Figure 11) suggests that the MALDI-TOF MS assay could be a very good replacement for the conventional radioisotope method.

4.6 The Comparison of MS-based and Fluorescence-labeled Assays for Endonuclease V Repair Pathway Study

We have conducted several studies to understand the Endo V repair pathway and have also demonstrated the importance of Pol I exonuclease activity in dI repair [19,74]. In this study, we further explore the possible interaction between Endo V and Pol I, and also establish a reliable detection method for Endo V enzymatic analysis.

doi:10.6342/NTU202401466

We designed two protocols for Endo V nicking assays. One is an MS-based high turnover protocol, and the other is a PAGE-based slow turnover protocol. A shorter substrate (21-nt/28-nt DNA duplex) easily denatures after Endo V nicking, causing Endo V to dissociate from the DNA and allowing turnover. This turnover activity enhances sensitivity in MS-based analysis which is 16 times higher than PAGE-based assay.

The other one is PAGE-based analysis, which used a 40-bp fluorescence-labeled substrate DNA duplex as a substrate. For 40-bp substrate, DNA remains a stable complex after nicking by Endo V with a melting temperature (Tm) of approximately 52.7 °C. The stable structure of the DNA enables Endo V to bind to DNA effectively, thus it is suitable for analyzing the interaction between Pol I and Endo V. This highlights that substrate design is crucial for optimizing enzymatic assay performance. Detailed comparison data between these two methods can be found in Table 8.

Several important factors need to be considered when designing experiments for denatured electrophoresis. The resolution of fluorescence relies on the quality of the polyacrylamide gel. A well-prepared polyacrylamide gel results in lower background noise and improved resolution, particularly when dealing with low amounts of products. The background signal directly impacts the quantification in electrophoresis, leading to

doi:10.6342/NTU202401466

variations in analysis from batch to batch. In comparison to MS-based analysis, it is easier to distinguish the products from the signal when the signal-to-noise ratio is around 100, and the resolution is 1 Dalton [18]. Potassium adducts can be observed in the MS spectra, however, the signal from these adducts does not interfere with the quantification. The denatured electrophoresis can only analyze 10 to 15 samples on the same gel and four gels for the same batch. Yet automated MALDI-TOF MS allowed for at least 300 samples to be analyzed in the same batch within 2 hours [151].

The denatured electrophoresis has been widely used in many studies, but it still has limitations. An appropriate approach reduced the artifacts that may impact the outcome. Recently, the monofunctional glycosylase in mammalian -hSMUG1 was reported which processed incision activity [172]. The incision products were resolved by denature electrophoresis, which easily breaks down the AP products during the electrophoresis. Denaturing reagents such as formamide and urea, used in the denaturing gel electrophoresis protocol, may cause AP site breakage. The 3’ phosphodiester bonds of the AP site associated with the open aldehyde form are labile and can be hydrolyzed by a β-elimination reaction in which the pentose carbon beta to the aldehyde is activated at alkaline pH and elevated temperature [173]. We mimic denaturing electrophoresis conditions and treat the AP products in monofunctional glycosylase assay, then

doi:10.6342/NTU202401466

subjected the products to MS analysis. The results indicated that the denaturing reagent and the heating process caused non-enzymatic breakage, which could be prevented by using MALDI-TOF MS analysis [174].

4.7 The Turnover Activity of Endonuclease V

In part three of the study of Endo V, we focused on the turnover activity of Endo V. Endo V was initially purified as the deoxyinosine 3' endonuclease from *E. coli* in 1994 [131]. The endonuclease activity of the purified protein was determined by utilizing the PM2 DNA substrate containing dI. The PM2 circular DNA was 10,079 bp long with a GC content of 42.2% [175], and deoxyinosine was incorporated into the DNA to form I-C pairing through nick translation [176]. The unit definition of Endo V was defined as the enzyme amount required to cleavage 1 pmol of dI substrate per minute. According to our findings, the I-C substrate is not a good substrate for analyzing Endo V nicking activity, we suggest that the Endo V activity may underestimated under such reaction conditions.

The native endonuclease V exhibits poor solubility at physiological salt concentrations. Moreover, the low concentration in the cell and its instability during purification make the production of homogeneous preparations of endonuclease V

doi:10.6342/NTU202401466

challenging. Purification of the native Endo V was challenging until the complete sequence of the *nfi* gene was resolved [120]. However, the genetically engineered Endo V could be easily obtained. Commercialized Endo V is a fusion protein with a maltose-binding protein tag. Maltose-binding protein (MBP) is widely utilized as a fusion partner for generating recombinant proteins in bacteria [177]. The fusion does not affect the enzymatic activity and the binding affinity of Endo V to the dI lesion after nicking [129]. However, the extent to which this fusion protein may influence binding and turnover compared to the native enzyme remains unknown.

We further noticed that the difference of the substrate, the close circular heteroduplex is more biologically relevant and was used in the previous studies [134,135]. In our study, we obtained different results by using oligonucleotide substrates. We observed that Endo V exhibited nicking preference as follows: I-G > I-A ≈ I-T > I-C. The nicking activity was extremely high on the I-G substrate, it could be caused by the thermostability of the DNA duplex. According to the nearest-neighbor free energy models, a generalized stability trend was derived: I-C > I-A > I-T ≈ I-G > II from which finds widespread use in practical applications [178,179].

doi:10.6342/NTU202401466

Endo V binds DNA tightly through the interaction of the lesion-recognition pocket and the inosine [180]. Therefore, it may present a signal for the polymerase I to engage in the dI repair pathway. To understand the interplay of the Endo V and Pol I in the dI repair pathway, a turnover assay was conducted. The turnover activity of Endo V was promoted by the Klenow Fragment (KF) with proficient 3’ to 5’ exonuclease activity, which was higher than the KF exonuclease-deficient form. Based on our findings, we suggested that the physical interactions between Endo V and KF played an important role in Endo V turnover activity. Pol I removed deoxyinosine (dI) via its 3’ to 5’ exonuclease active site. Subsequently, Endo V dissociates from the DNA duplex due to the destabilization of the DNA structure.

In addition, the result of 3’ to 5’ exonuclease deficient KF suggested physical interaction between Pol I and Endo V also promoted turnover with less extend. It’s suggested that an additional mechanism, such as protein-protein interactions by direct physical contact is involved. The co-immunoprecipitation experiment of Pol I and Endo V in the presence of deoxyinosine substrate should provide further insight into this dI repair pathway.

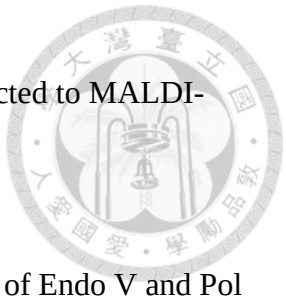
doi:10.6342/NTU202401466

Chapter V Conclusion

In conclusion, our study employed MALDI-TOF MS with unlabeled DNA substrates to investigate proofreading enzymes' activity, such as UDG, and Endo V. We found that Pol I effectively proofread insertion/deletion errors near the primer terminus. The indel error within 5-nt from the 3' terminus was effectively proofread. Based on these results, we proposed a model of Pol I proofreading capacity. To confirm these models, co-crystallography of DNA fitting in the exonuclease site is needed.

The MS-based UDG analysis demonstrates fast and high throughput advantages, particularly in inhibitor screening tests, which can be applied to pharmaceutical use. In cancer treatment, the effectiveness of anti-cancer drugs like 5-FU is related to the activity of uracil DNA glycosylase. The combination treatment of glycosylase inhibitors and 5-FU may overcome drug resistance. Therefore, establishing a high-throughput platform for inhibitor screening is essential. Furthermore, the results indicated that the G-U substrate is one of the most suitable substrates for UDG assays. This platform also could be adapted to bi-functional glycosylases such as formamidopyrimidine DNA glycosylase (Fpg) and 8-oxoguanine glycosylase (OGG1). Bi-functional glycosylase cleaves the AP site, followed by DNA lesion removal. By designing proper DNA substrates, bi-functional glycosylase can cleave the AP site, followed by DNA lesion

doi:10.6342/NTU202401466



removal. This results in two separated fragments, which can be subjected to MALDI-TOF MS for measuring the glycosylase activity.

In Endo V analysis, the results indicated the possible interaction of Endo V and Pol I. It proposed that the Endo V turnover activity could be promoted by either exonuclease proficient KF or exonuclease deficient KF, through the physical interaction between two proteins. The protein-protein interaction of Endo V and Pol I is await for the further investigation.

In conclusion, our method is expected to apply to other DNA repair proteins, providing a robust and high-throughput approach for analyzing and understanding the complex mechanisms of DNA repair. This work lays the foundation for further studies on the interactions and functions of various DNA repair enzymes, which are crucial for maintaining genomic integrity and developing effective therapeutic strategies.

doi:10.6342/NTU202401466

Figures

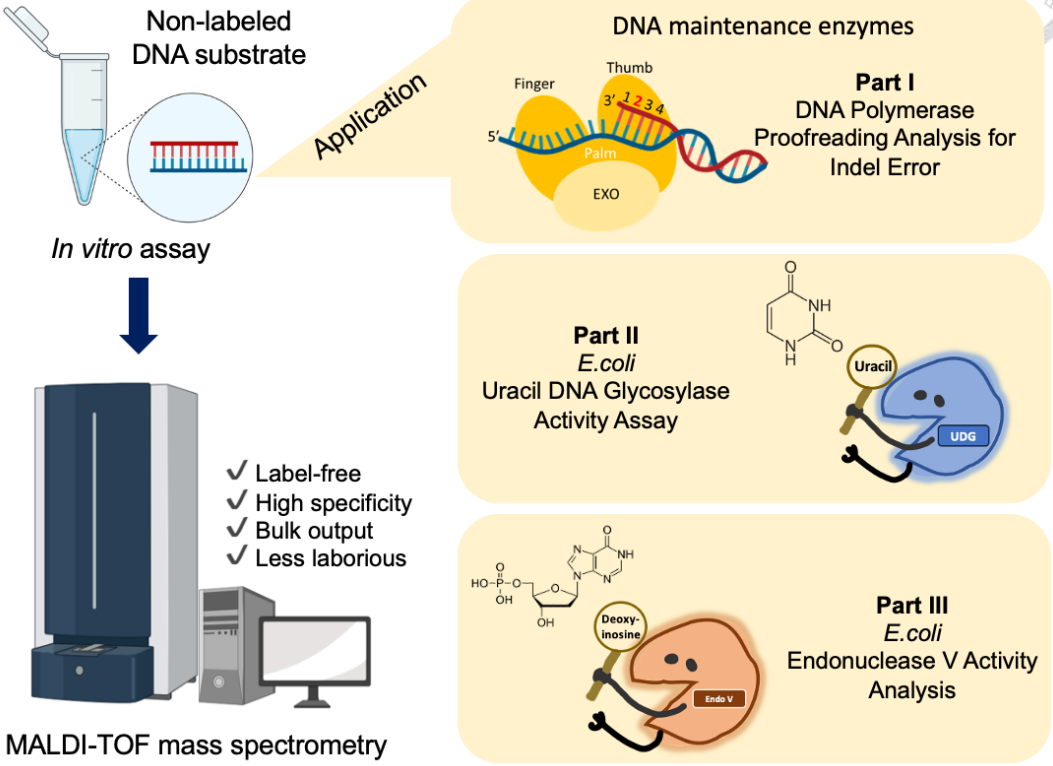


Figure 1. The graphic abstract of MALDI-TOF MS based analysis of DNA repair proteins

In this study, the MS-based analysis was developed to analyzed three DNA maintained protein included DNA polymerase I, uracil DNA glycosylase and Endonuclease V. Non-labeled DNA duplex was prepared in a 1.5 mL microcentrifuge tube and co-incubated with DNA repair protein included DNA polymerase I (Pol I), uracil DNA glycosylase (UDG) and endonuclease V (Endo V). The DNA was then subject to MS analysis.

doi:10.6342/NTU202401466

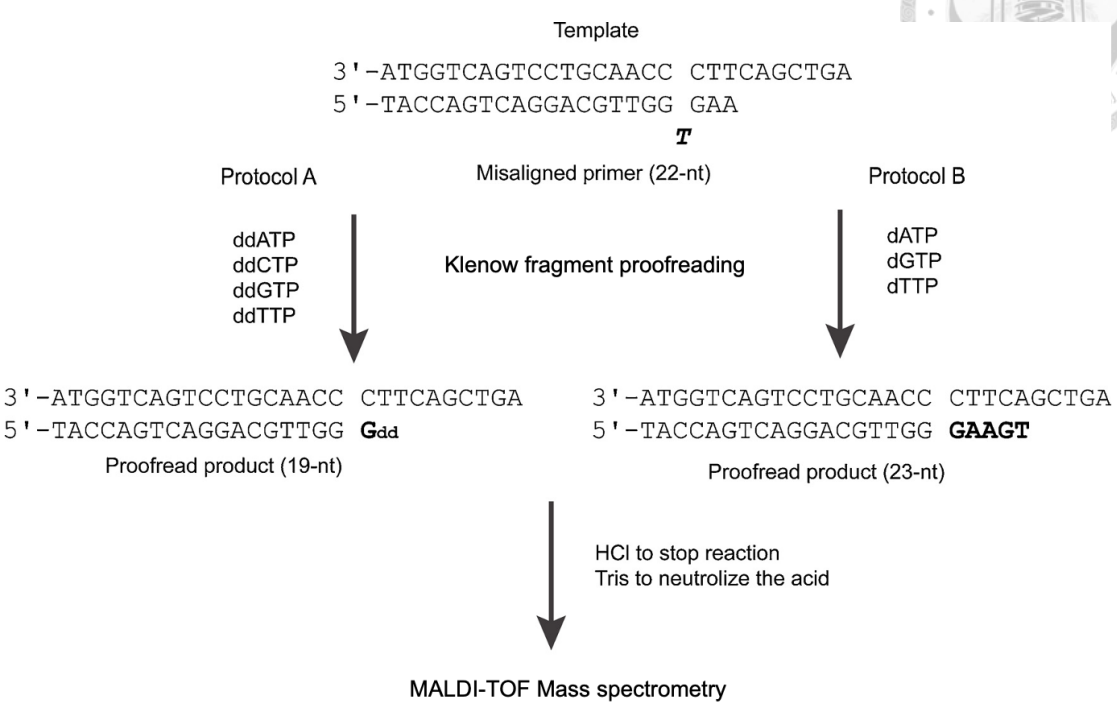


Figure 2. The scheme of the proofreading assay by MALDI-TOF MS analysis

Primers with an indel error (extra nucleotide in italic bold) were annealed as indicated to a template. The proofreading exonuclease of DNA polymerase identifies and removes the misaligned nucleotides. In protocol A, in the presence of four ddNTPs, the proofreading reaction terminates by synthesizing a single complementary base at the indel error site. In protocol B, in the presence of three dNTPs, the proofread primer can be extended to a specified length according to the template sequence. When analyzed using MALDI-TOF MS, the differences in mass between mismatched primers, excision products, and proofreading products can be distinguished.

doi:10.6342/NTU202401466

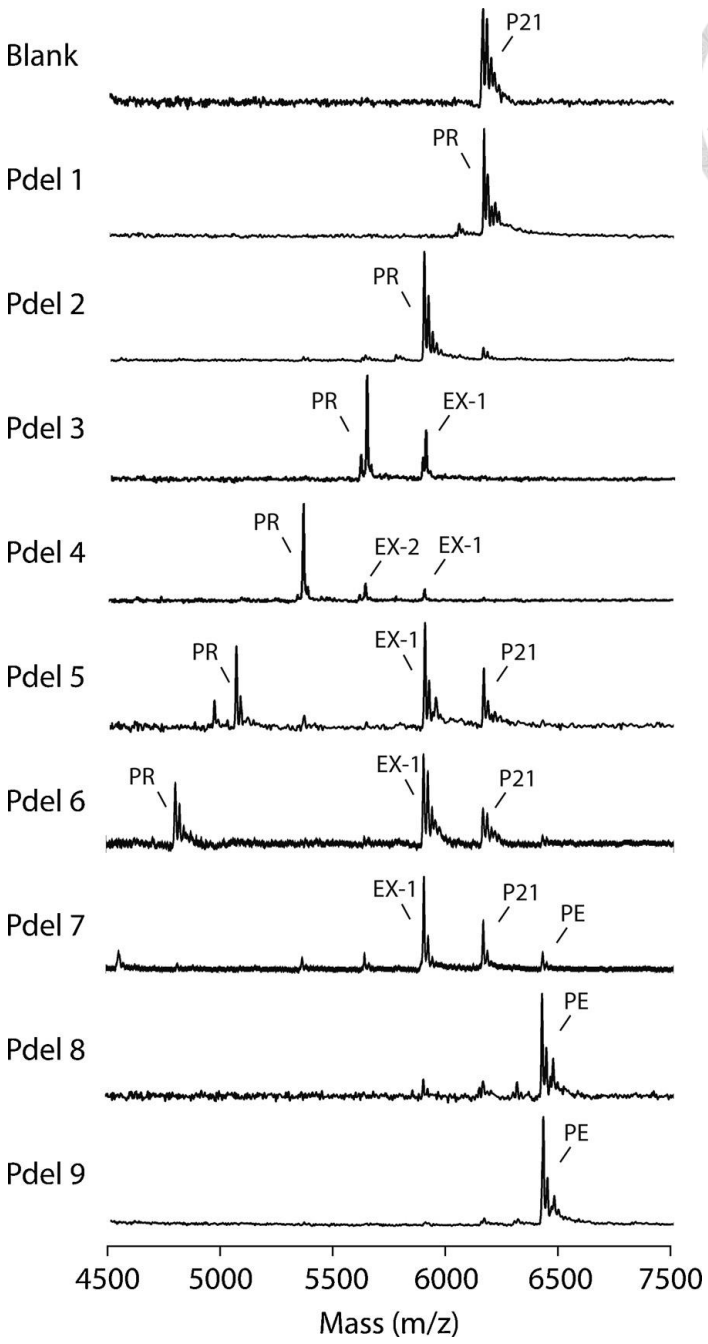


Figure 3. The single nucleotide deletion error from position 1 to 5 could be proofread by Klenow fragment effectively with 4ddNTPs

Proofreading assays were performed with 2 U (8.6 pmol) of KF and 50 pmol substrates at 37 °C for 20 minutes, following the procedure described in section 2.3. Blank refers to the enzyme blank of the P21 primer. Substrates labeled Pdel 1 to Pdel 9

doi:10.6342/NTU202401466

exhibit single nucleotide deletion errors in the primer, with the numbers indicating the loopout position relative to the 3’ end of the primer. P21, unrepaired primers; PR, proofread products; PE, primers; EX-1, EX-2, and EX-3, excision products, with the numbers indicating the nucleotides excised from the primers.

doi:10.6342/NTU202401466

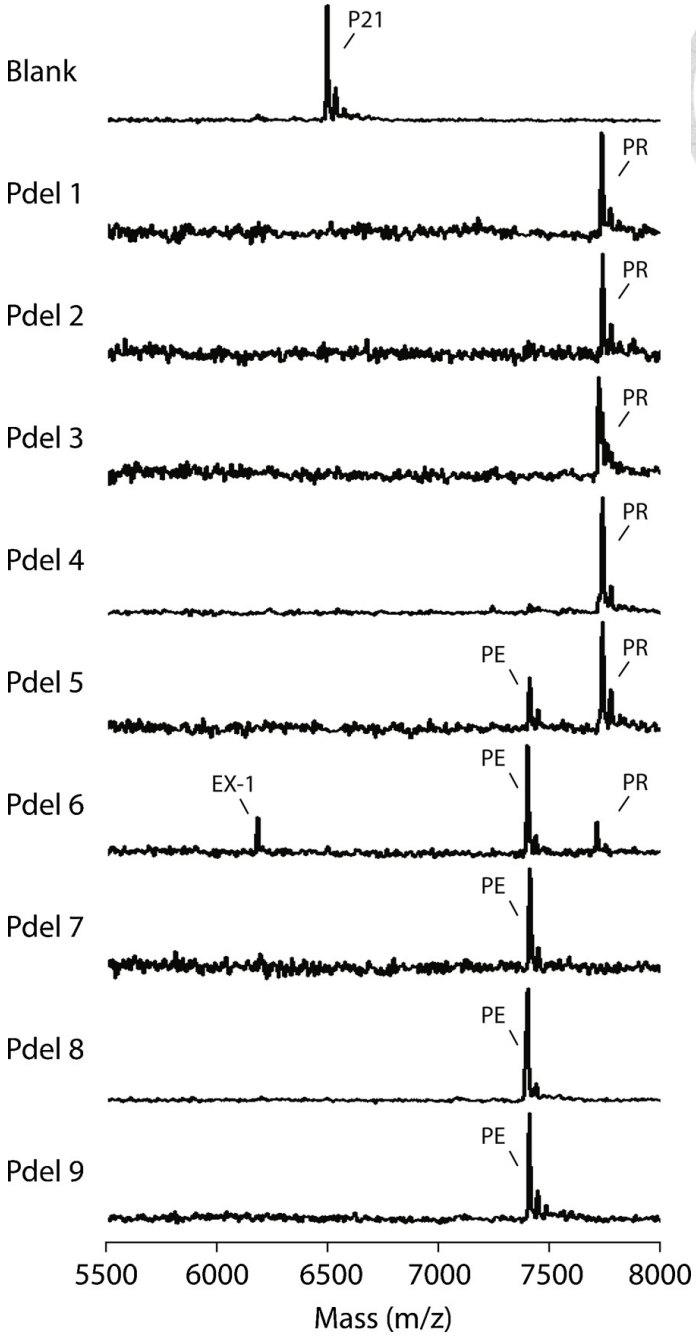


Figure 4. The single nucleotide deletion error from position 1 to 4 could be proofread by Klenow fragment effectively with 3dNTPs and ddCTP

Proofreading assays were performed with 2 U (8.6 pmol) of KF and 50 pmol substrates at 37 °C for 20 minutes, following the procedure described in section 2.3. Blank refers to the enzyme blank of the P21 primer. Substrates labeled Pdel 1 to Pdel 9

doi:10.6342/NTU202401466

exhibit single nucleotide deletion errors in the primer, with the numbers indicating the loopout position relative to the 3’ end of the primer. P21, unrepaired primers; PR, proofread products; PE, extended primers; EX-1, an excision product involving the removal of a single nucleotide from the primer.

doi:10.6342/NTU202401466

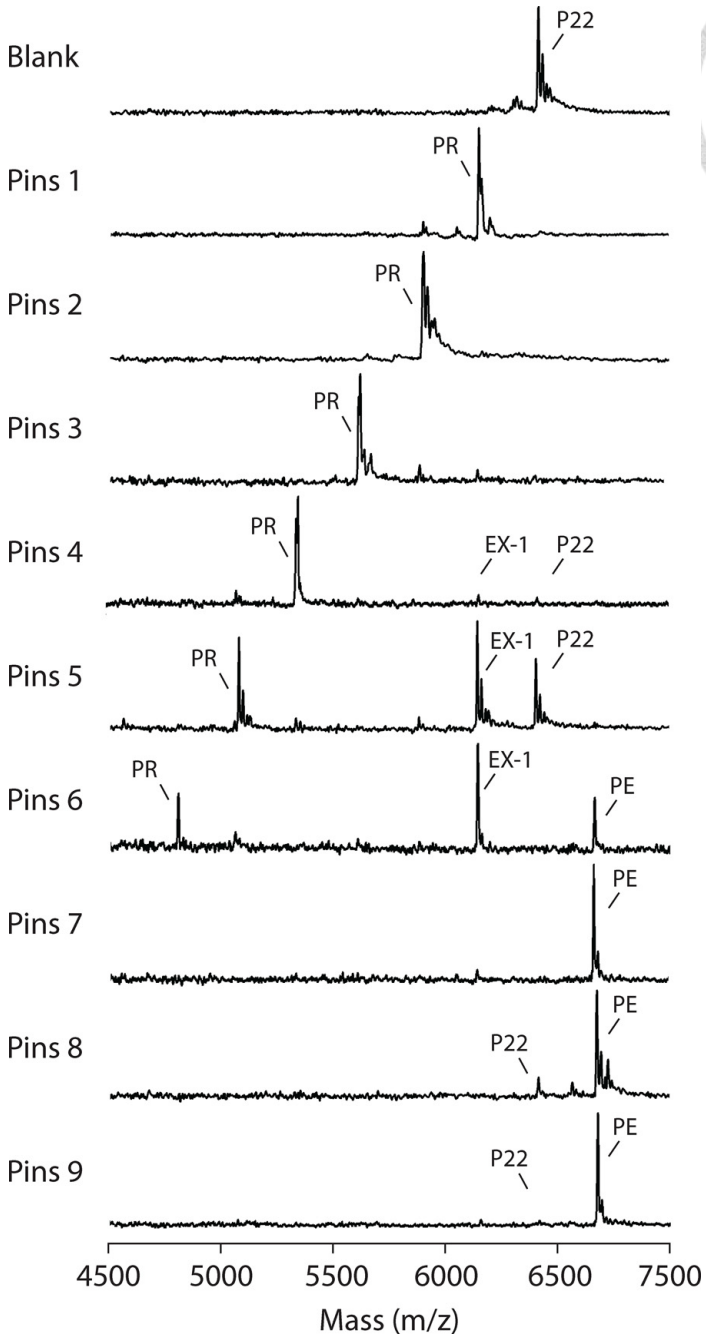


Figure 5. The single nucleotide insertion error from position 1 to 5 could be proofread by Klenow fragment effectively with 4ddNTPs

Proofreading assays were performed with 2 U (8.6 pmol) of KF and 50 pmol substrates at 37 °C for 20 minutes, following the procedure described in section 2.3. Blank refers to the enzyme blank of the P22 primer. Substrates labeled Pins 1 to Pins 9

doi:10.6342/NTU202401466

exhibit single nucleotide insertion errors in the primer, with the numbers indicating the loopout position relative to the 3’ end of the primer. P22, unrepaired primers; PR, proofread products; PE, primers; EX-1 and EX-2, excision products, with the numbers indicating the nucleotides excised from the primers.

doi:10.6342/NTU202401466

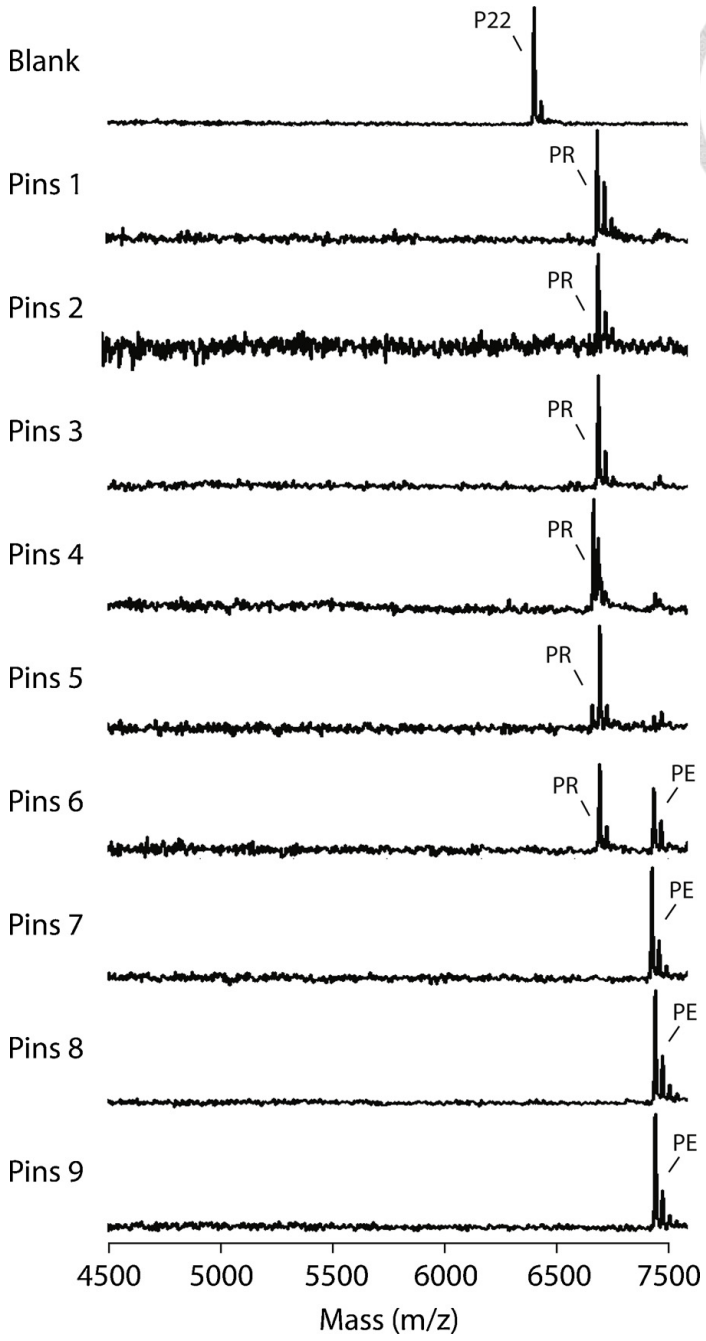


Figure 6. The single nucleotide insertion error from position 1 to 5 could be proofread by Klenow fragment effectively with 3dNTPs

Proofreading assays were performed with 2 U (8.6 pmol) of KF and 50 pmol substrates at 37 °C for 20 minutes, following the procedure described in section 2.3. Blank refers to the enzyme blank of the P22 primer. Substrates labeled Pins 1 to Pins 9

doi:10.6342/NTU202401466

exhibit single nucleotide deletion errors in the primer, with the numbers indicating the loop-out position relative to the 3’ end of the primer. P21, unrepaired primers; PR, proofread products; PE, extended primers.

doi:10.6342/NTU202401466

Figure 7. The scheme of DNA Polymerase I proofreading capacity for indel errors

In the event of an indel error arising from polymerase mis-incorporation or primer-template misalignment, DNA synthesis undergoes a temporary pause to facilitate the transfer of the primer to the 3'-5' exonuclease site (exo, highlighted in red) for proofreading. Indel errors occurring within the primer-template junction and close to the polymerase (numbers in red) trigger the primer transfer to the exonuclease site for proofreading. Indel errors situated in a region marginally in contact with the polymerase (numbers in orange) undergo partial proofreading. Indel errors positioned farther away from the primer's 3’ terminus and not in direct contact with the polymerase (numbers in black) result in the primers remaining in the polymerase site, escaping the proofreading process.

doi:10.6342/NTU202401466

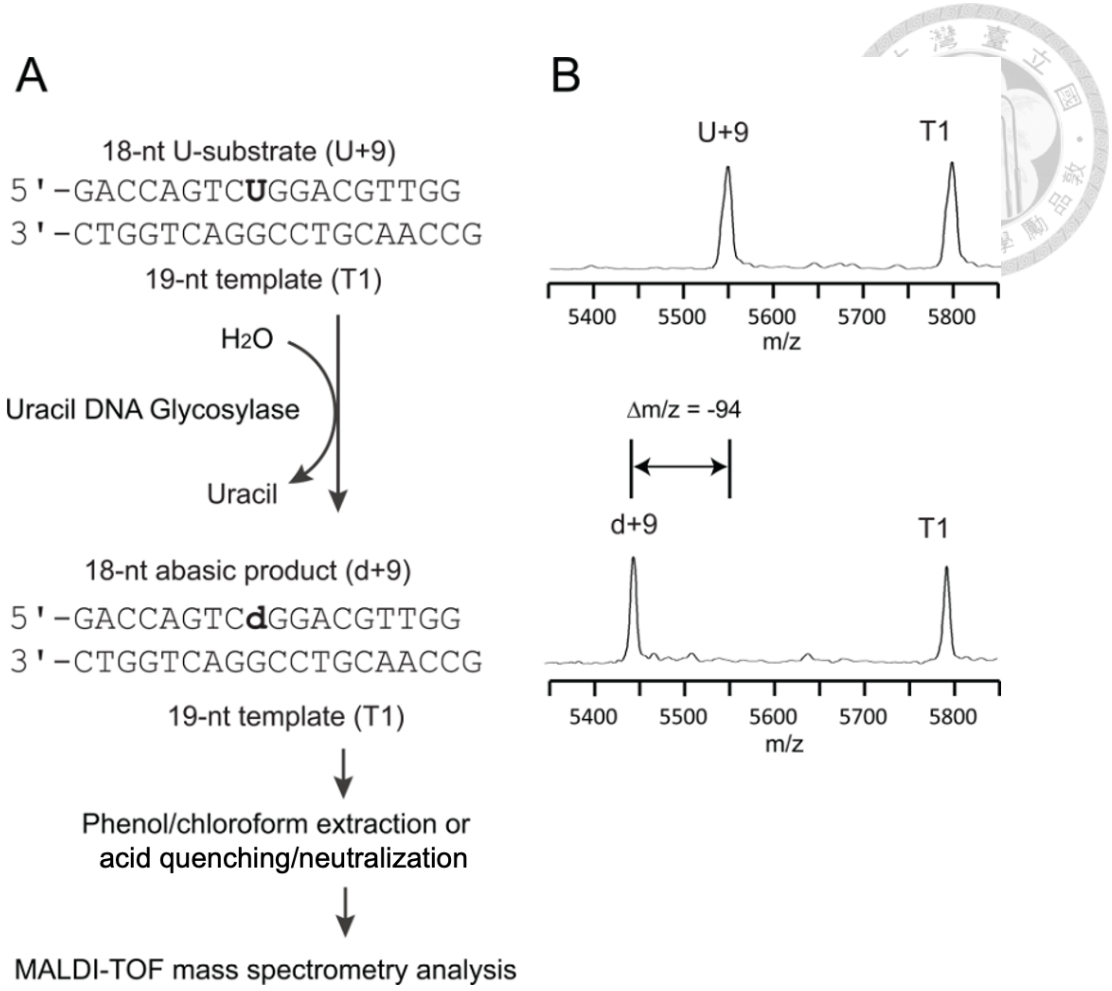


Figure 8. Model system for DNA uracil glycosylase assay

(A) The example of UDG reaction. The lesion strand, containing a single uridine (U + 9), was annealed as to a template (T1), forming a G-U substrate (U in bold). Uracil is removed by Uracil-DNA glycosylase creating an abasic products (d+9). When analyzed by MALDI-TOF MS, the distinction in m/z values between U-containing DNA and AP products can be resolved, as shown in panel (B). (B) The MS spectrum of UDG assay. A 18-nt DNA containing a single uracil (U+9) annealed to a 19-nt template (T1) was examined. In the upper spectrum, an enzyme blank of 50 pmol substrate in a 40 μL reaction buffer was subject to MS analysis. In the lower spectrum, a complete

doi:10.6342/NTU202401466

enzyme digestion of 50 pmol substrate in a 10 μL reaction with 0.5 unit UDG at 37 °C for 30 minutes. T1, the 19-nt G-containing template; U + 9, the 18-nt U-containing substrate; d+9, the 18-nt abasic product.

doi:10.6342/NTU202401466

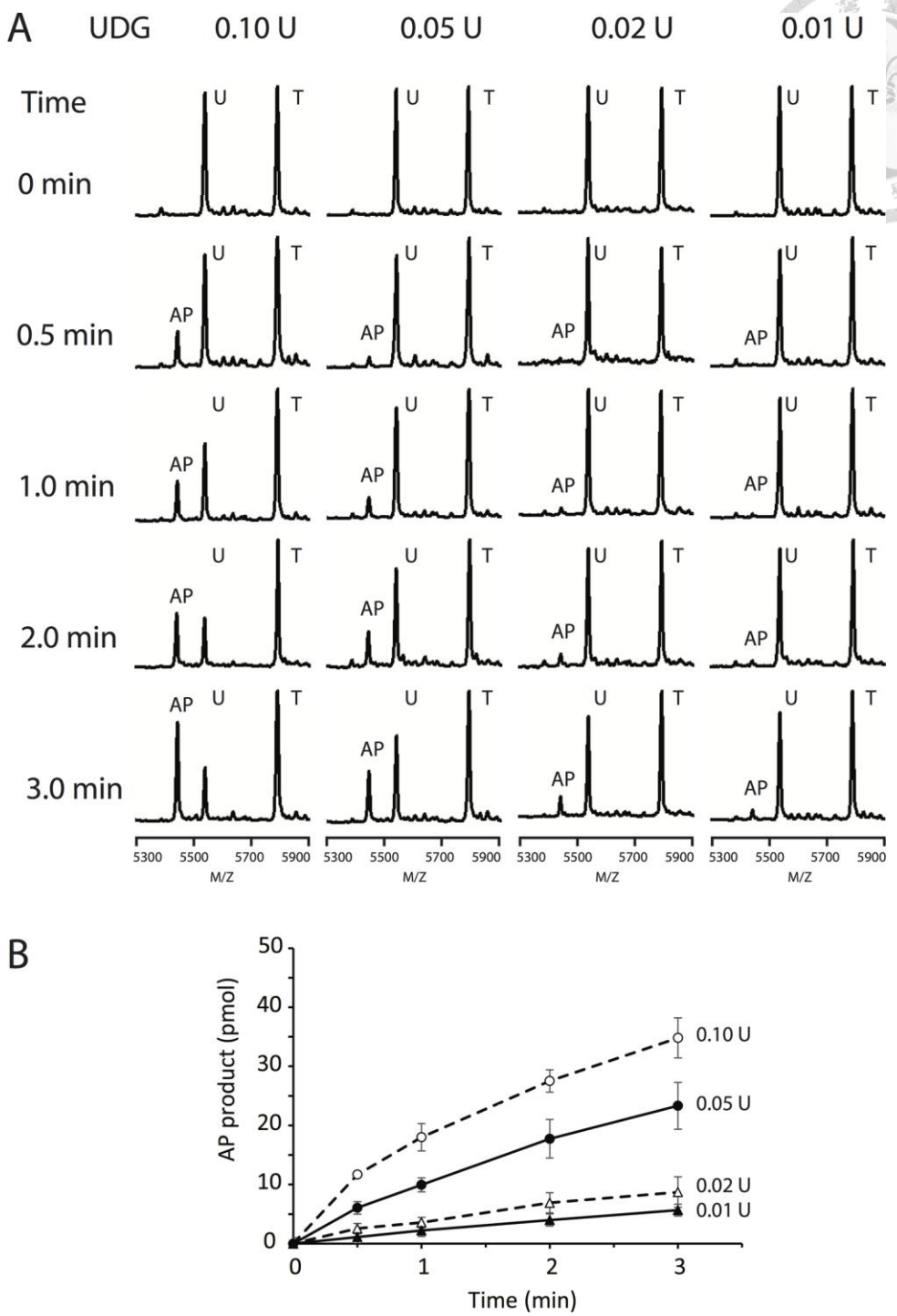


Figure 9. The UDG enzymatic activity measured by MS-based analysis showed time and dose dependency

(A) MALDI-TOF Mass spectrum allegorized the concentration-dependent processing of G-U substrates by UDG. 50 pmol of G-U DNA substrate was incubated

doi:10.6342/NTU202401466

with UDG at concentrations of 0.01, 0.02, 0.05, and 0.1 unit at 37 °C. Aliquots (10 μL) were taken from the reaction mixture at 0, 0.5, 1, 2, and 3 minutes and quenched with phenol/chloroform extraction. T, T1 template; U, U+9 substrate; AP, the 18-nt abasic product. (B) The kinetic assay with different UDG concentration. The amount of product was plotted against time, with UDG concentrations of 0.01 unit (closed triangles with a solid line), 0.02 unit (open triangle with a dashed line), 0.05 unit (closed circle with a solid line), and 0.1 unit (open circles with a solid line). The points represent the average of three independent determinations, and the error bars indicate 1 S.D.

doi:10.6342/NTU202401466

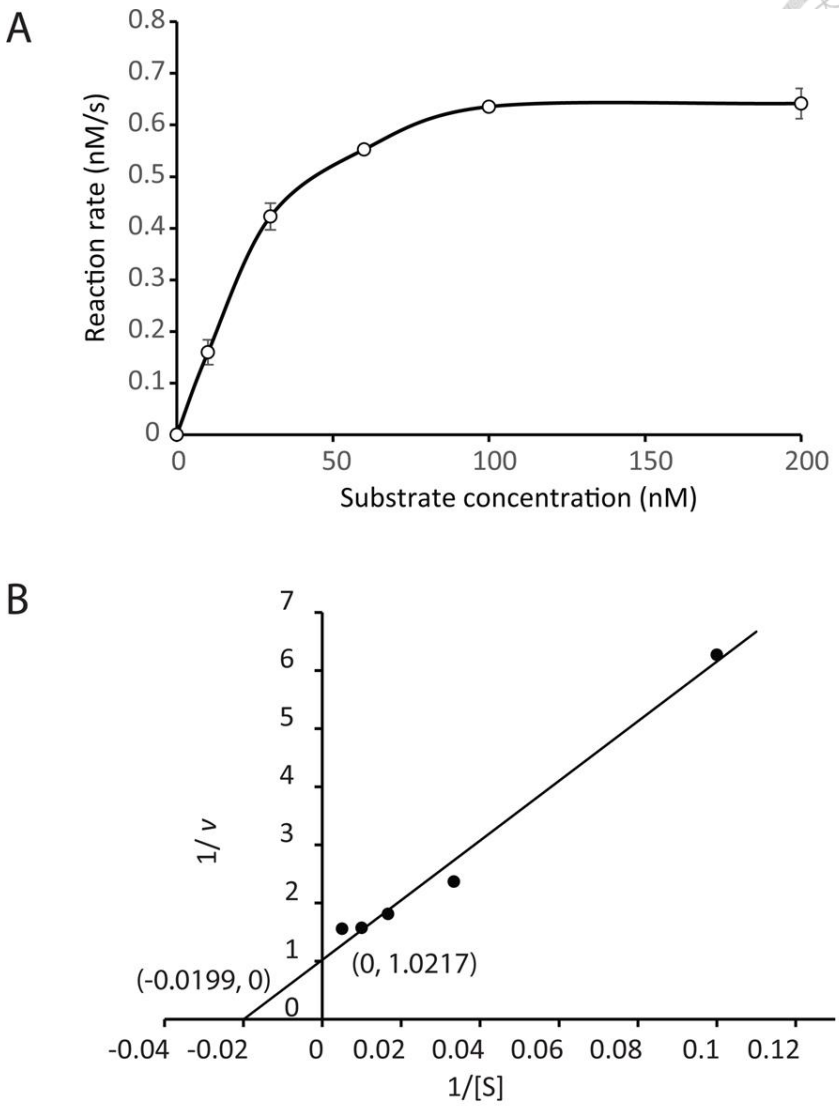


Figure 10. The UDG kinetic parameter determined by MALDI-TOF MS was comparable to traditional method.

Michaelis-Menten curve and Lineweaver-Burk plots to determine the K_m and k_{cat} for UDG enzyme-catalyzed excision of Uracil from DNA. V, reaction velocity (nM/s). The DNA glycosylase assay was performed using 0.1 Unit UDG and different concentrations (10, 30, 60, 100, 200 nM) of G-U substrate in 100 μL reaction for 30 second and 60 second.

doi:10.6342/NTU202401466

(A) Michaelis-Menten curve generated from three independent experiments. The error bars represent 1 SD.

(B) Lineweaver-Burk plot for the assay, the indicated intercepts permit calculation Km and Vmax.

doi:10.6342/NTU202401466

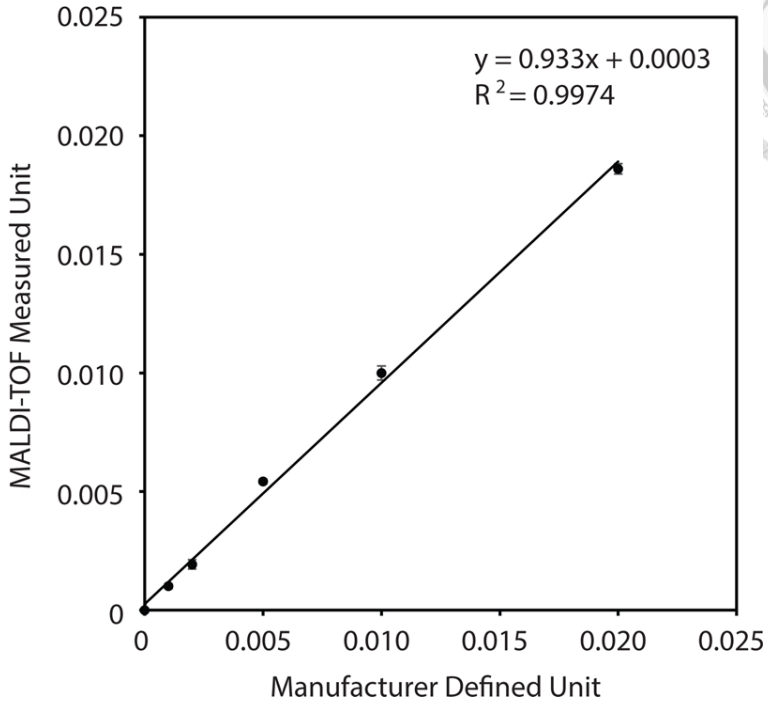


Figure 11. UDG assay by MALDI-TOF MS analysis was comparable to the manufacturer-assigned unit

Enzyme activity was defined through the removal of uracil, leading to the abasic site products. In a 10 μL reaction, 50 pmol (0.56 μg) of G-U substrate was incubated with diluted UDG at 37 °C for 30 minutes. The UDG was diluted with buffer [50% glycerol, 20 mM Tris-HCl (pH 7.5), 30 mM NaCl, 0.5 mM EDTA, 1 mM dithiothreitol] on ice. The unit was defined as the amount of enzyme that catalyzed the release of 60 pmol of uracil per minute. Error bars represent the standard deviation from six experiments.

doi:10.6342/NTU202401466

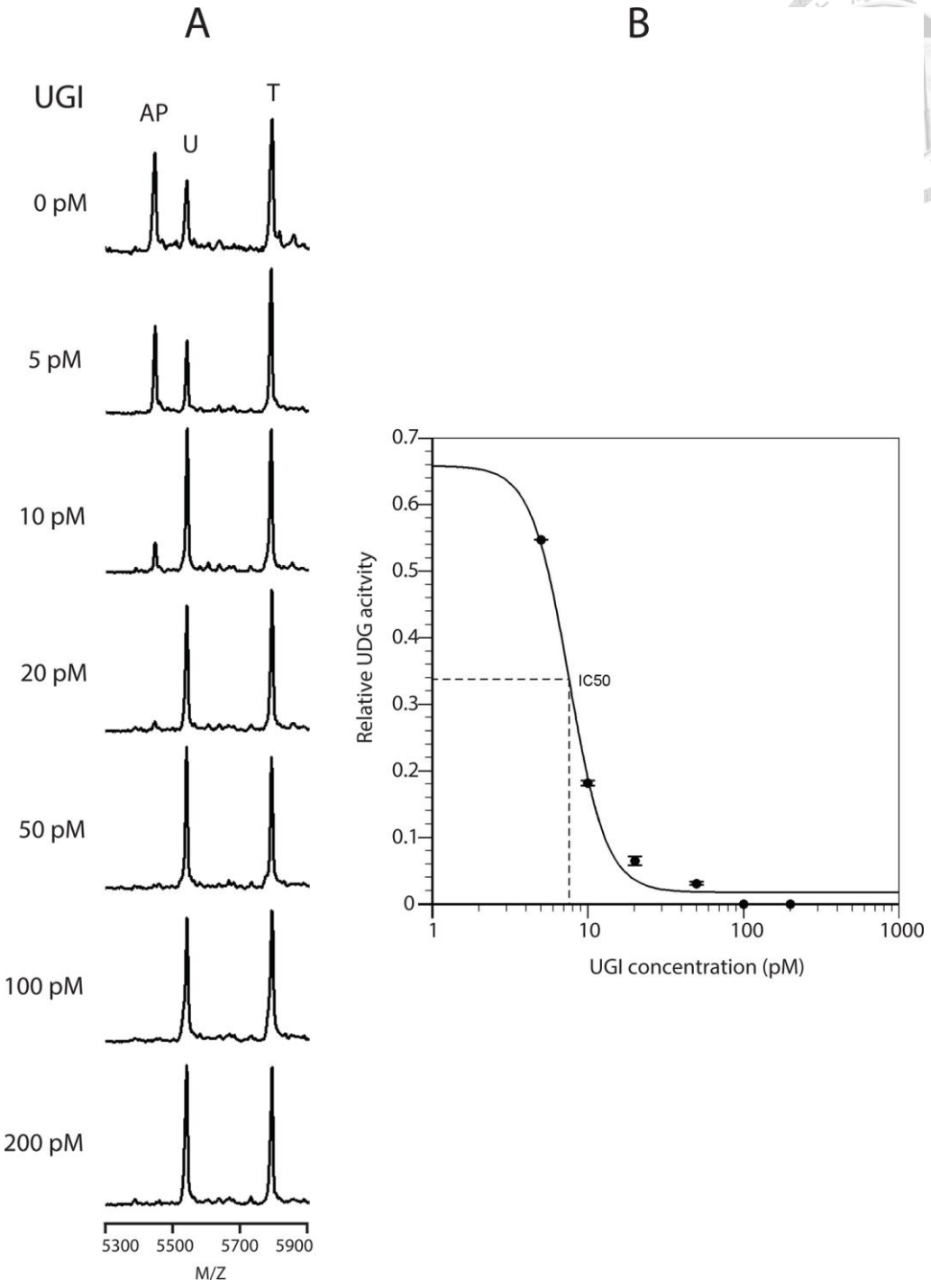


Figure 12. The inhibition curve of UGI

(A) In the 20 μL inhibition reactions, 0.05 Unit UDG (8 pM) and 50 pmol G: U substrate for 15 minutes, in the presence of UGI ranging from 0.001 Unit (5 pM) to 0.1 Unit (200 pM).

doi:10.6342/NTU202401466

(B) The inhibition curve of UGI. Each point represents the average of three independent determinations, with error bars indicating 1 S.D. The inhibition curve was fitted in the "Quest Graph™ IC50 Calculator" by AAT Bioquest, Inc. (June 19, 2020) and retrieved from https://www.aatbio.com/tools/ic50-calculator-v1.

doi:10.6342/NTU202401466

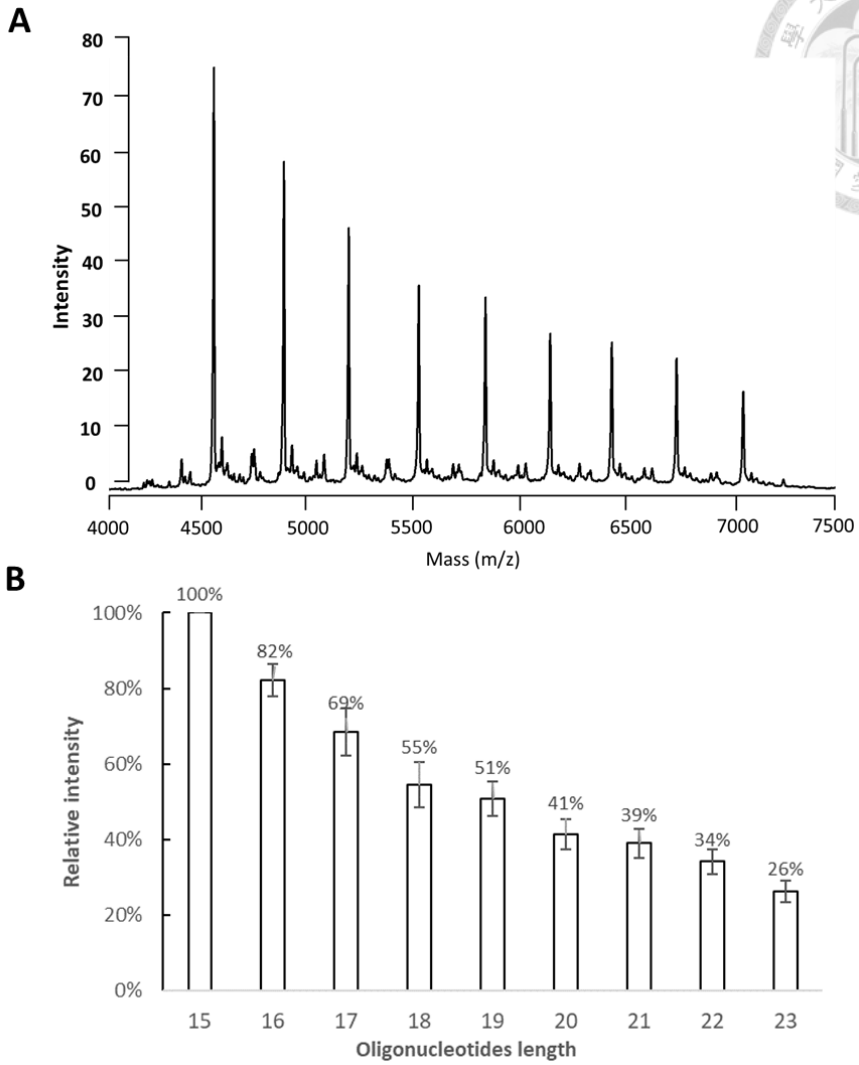
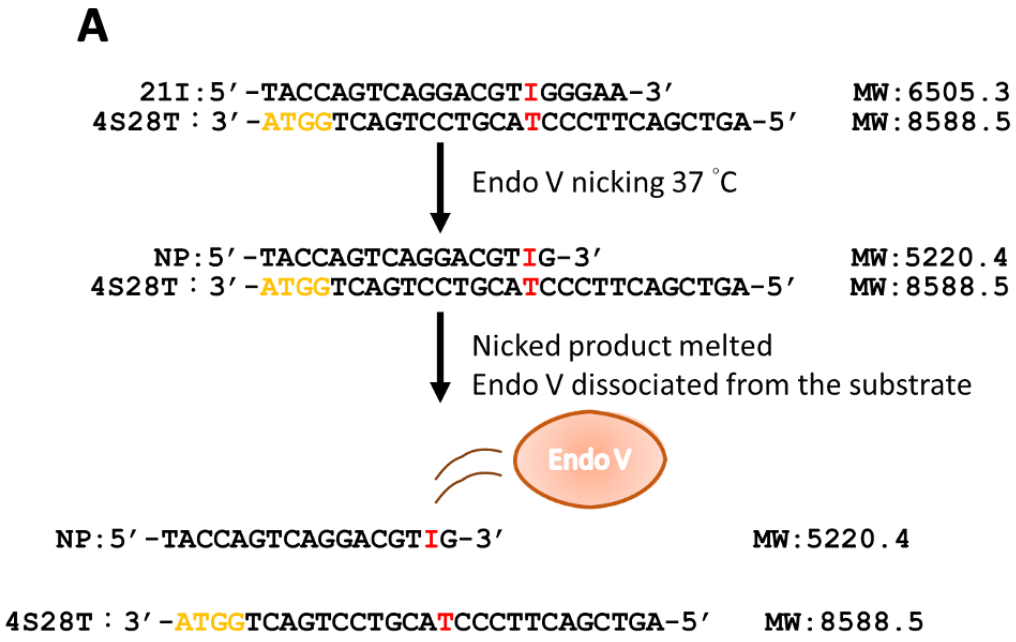
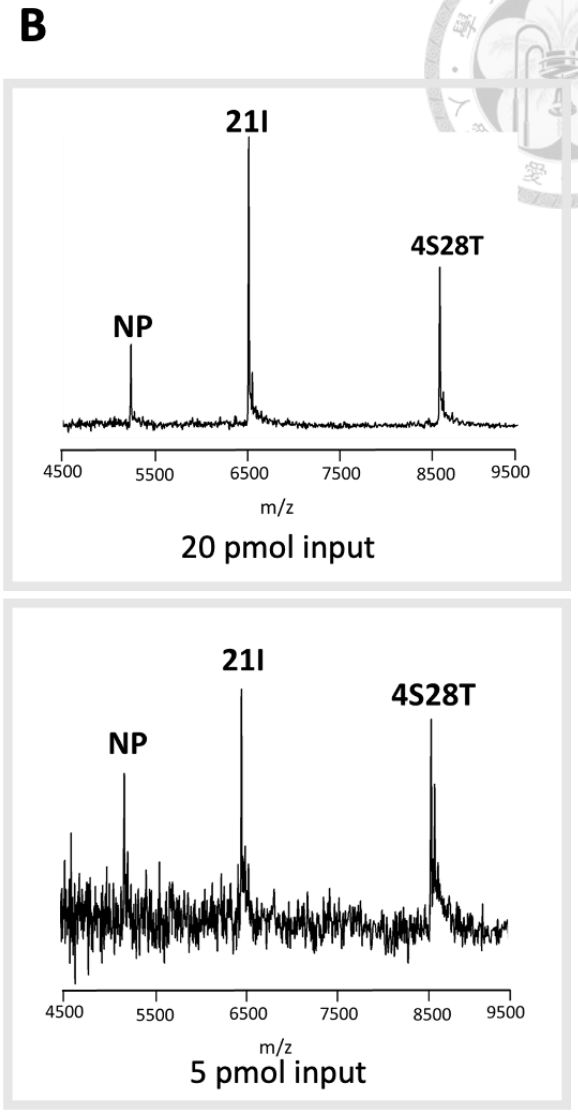


Figure 13. Signal intensity and molecular weight

Oligonucleotides from 15-nt to 23-nt were tested. (A) The spectrum of oligonucleotides. A mixture of 50 pmol of each oligonucleotide in 45 μL solutions was subjected to MS analysis. (B) The peak height of 15-nt was used as 100% for calculation. Relative signal intensity for each oligonucleotide was calculated using an average of four determinations and the error bars represent 1 S

doi:10.6342/NTU202401466



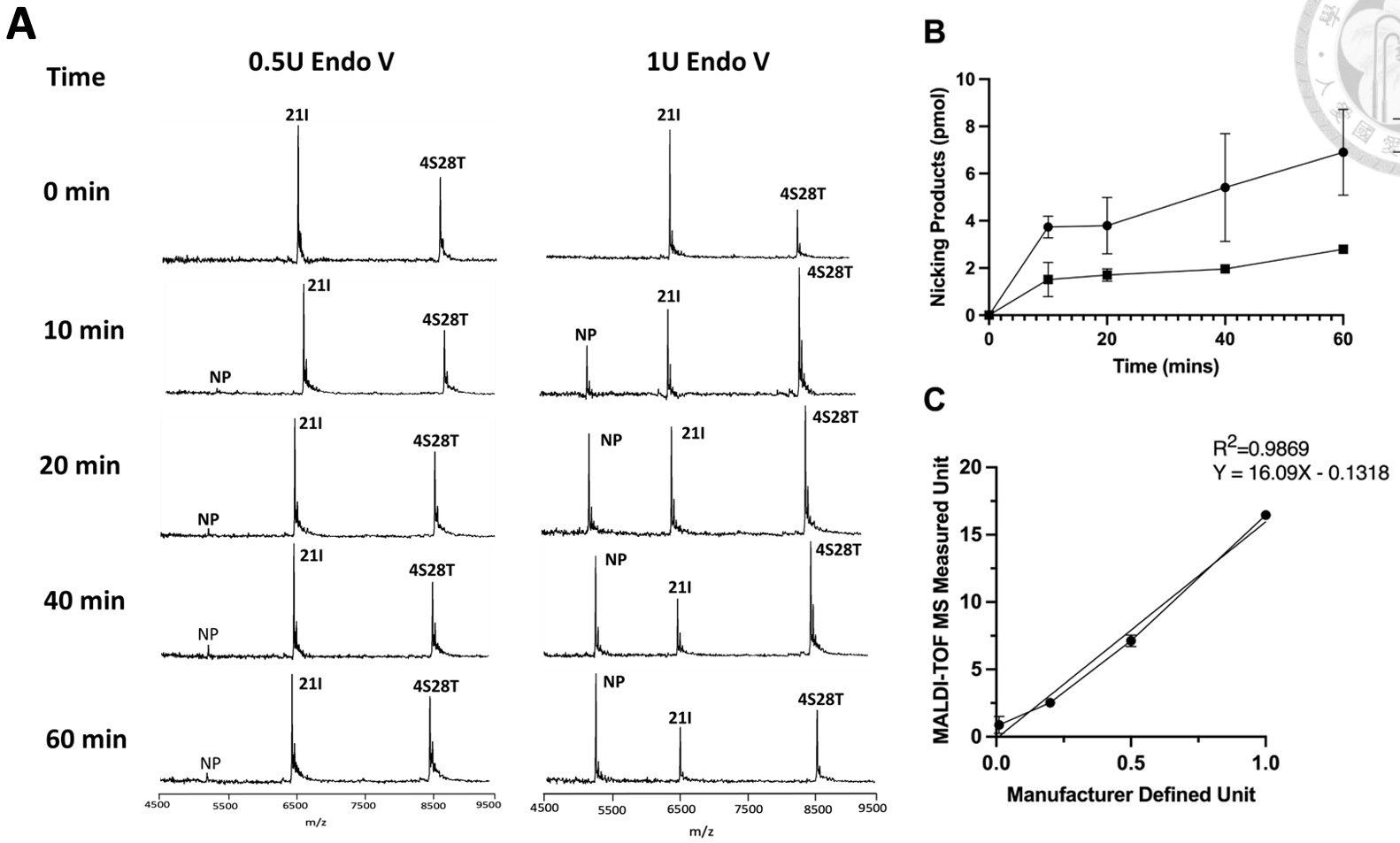


doi:10.6342/NTU202401466

Figure 14. Example of high turnover substrate for Endo V nicking activity analyzed by MALDI-TOF MS

The 21-IT (21I/4S28T) substrate was used as an example of a high-turnover substrate. (A) The scheme of Endo V nicking assay using a 21-IT duplex. Endo V nicked the short double-stranded dI substrate and then dissociated from the melted single-stranded product. (B) The signal-to-noise ratio of different reaction conditions. The nicking assay was conducted by 21-IT duplex incubated with 1U Endo V, 1X NEBuffer 4 (20 mM Tris-acetate, 50 mM potassium acetate, 10 mM Magnesium acetate, 1mM dithiothreitol, pH7.9 at 25°C). The reaction was performed at 37℃ for 10 minutes and followed by an acid quench/ Tris neutralization process. The signal-to-noise ratio of 20 pmol of 21-IT duplex input and 2U Endo V (upper) versus 5 pmol of 21-IT duplex input and 0.5 U Endo V were shown.

doi:10.6342/NTU202401466

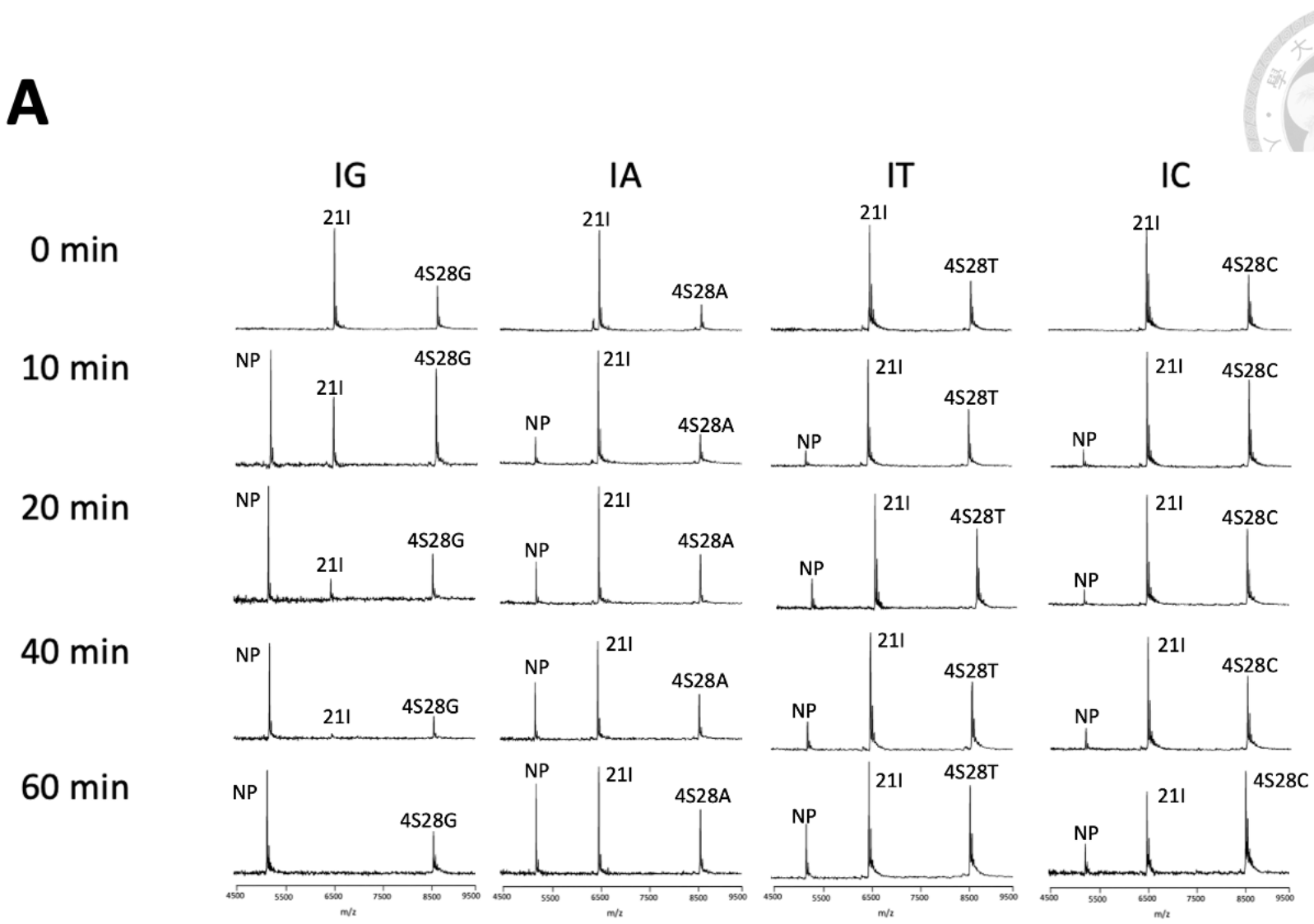


doi:10.6342/NTU202401466

Figure 15. Endo V enzymatic activity analysis by MALDI-TOF MS

(A) Nicking assay. The reaction contained 1X NEBuffer 4 (20 mM Tris-acetate, 50 mM potassium acetate, 10 mM Magnesium acetate, 1mM dithiothreitol, pH7.9 at 25°C), and 20 pmol of 21-IT substrate in 10 μL. The reaction was started by adding Endo V. The dose dependency of Endo V toward 21-IT substrate was observed. The reactions were scaled up 5 times and aliquots were quenched at 10, 20, 40, and 60 minutes for MALDI-MS analysis. (B) The nicking kinetic of Endo V. Endo V concentrations of 1 unit (circles) and 0.5 unit (squared). (C) Sensitivity assay. The Endo V was diluted with buffer [50% glycerol, 10 mM Tris-HCl (pH 7.4), 250 mM NaCl, 0.1 mM EDTA, 1 mM dithiothreitol, 0.15% Triton-X 100, 200 μg/ml BSA] on ice. Diluted Endo V incubated with 20 pmol of 21-IT and 1X NEbuffer 4 for 60 minutes at 37°C followed MS analysis. Each point was tested 3 times with independent measurements and the average was shown with 1±SD.

doi:10.6342/NTU202401466



doi:10.6342/NTU202401466

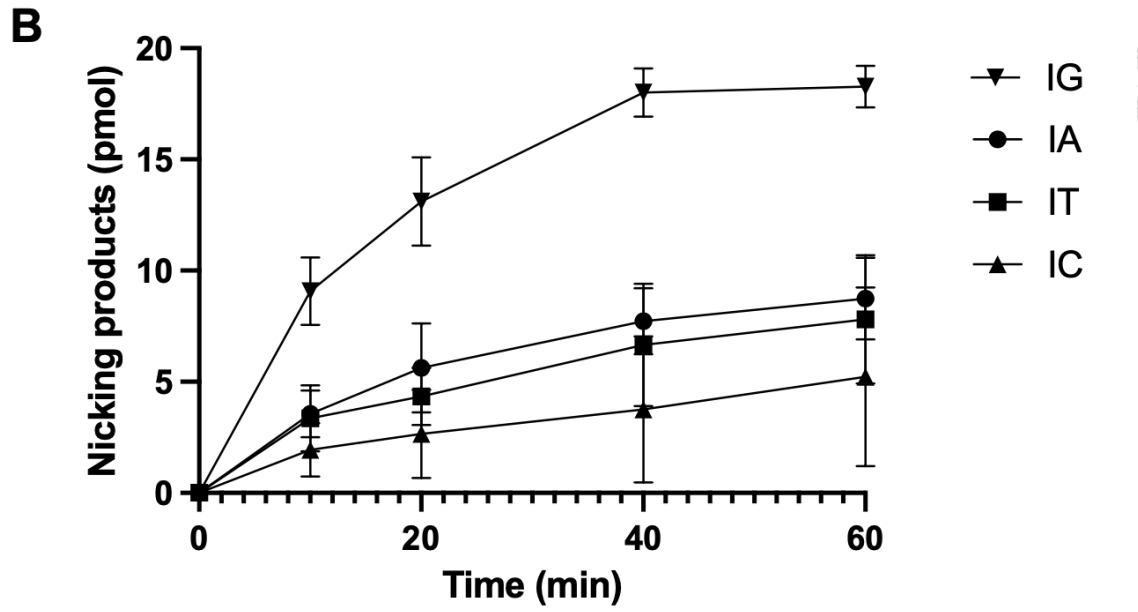


Figure 16. Endo V nicking specificity analysis using 21bp dI-substrate

(A) The nicking assay contained 1X NEBuffer 4 (20mM Tris-acetate, 50 mM potassium acetate, 10 mM magnesium acetate, 1 mM dithiothreitol, pH7.9 at 25°C), 20 pmol of substrate, 1 U of Endo V in a total reaction volume of 10 μL at 37°C. The reactions were scaled up 5 times and aliquots were quenched at 10, 20, 40, and 60 minutes for MALDI-MS analysis. (B) The time course analysis. Each point was an average of 4 independent measurements ± 1SD.

doi:10.6342/NTU202401466

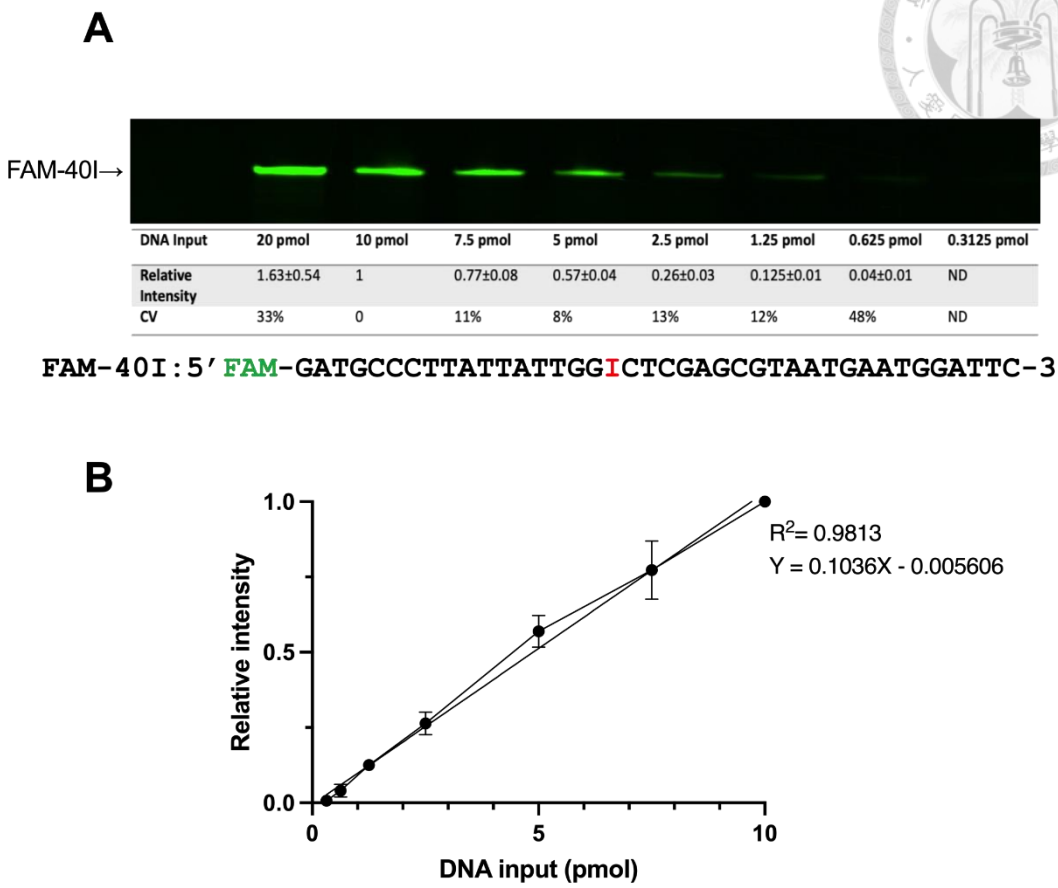


Figure 17. The detection limit of fluorescence-based gel assay

Measurement of the detection limit and the linearity range of the fluorescence detection assay. (A) Fluorescence gel analysis. Synthetic oligonucleotide contained single inosine (I, marked in red) with 5’ FMA modification (marked in green) was used in this study. The relative fluorescence intensity of each band was normalized with 10 pmol-band intensity. The fluorescence detection limit is 1.25 pmol (CV%=12%, results of 6 times independent experiments). (B) Detection lower limit determination. The correlation of relative intensity and DNA amount shown highly comparable, the

doi:10.6342/NTU202401466

R^2=0.9813 and the equation is Y=0.1036X-0.005606. Each point was an average of 6 independent measurements ± 1SD.

doi:10.6342/NTU202401466

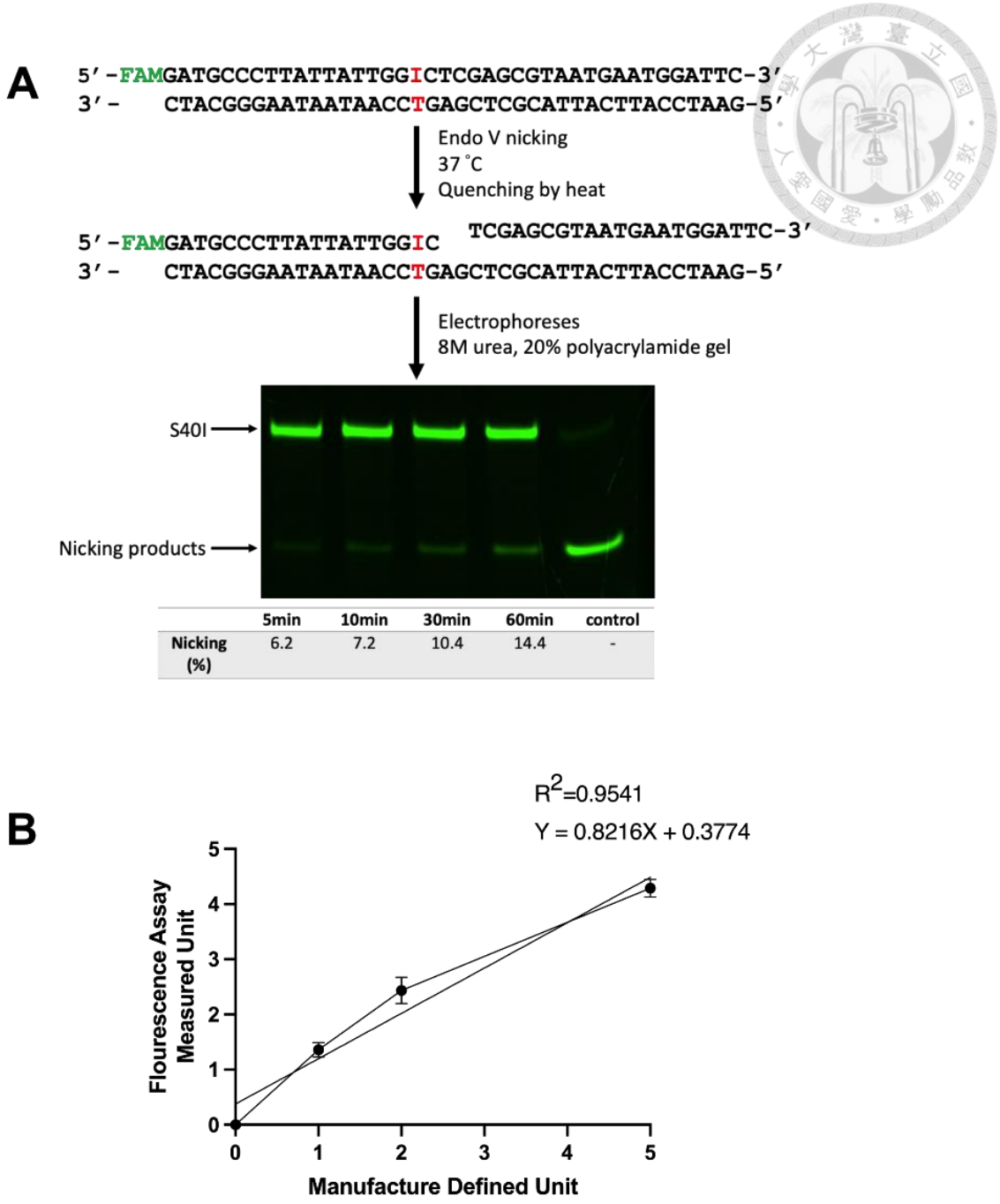


Figure 18. Endo V nicking assay using fluorescence-labeled substrate and sensitivity assay

(A)Endo V nicking assay. A 40-TI substrate was utilized in the Endo V nicking assay. The 19nt-Endo V nicking products were separated by electrophoresis and quantified using Image Lab Software (Bio-Rad). (B) Endo V selectivity assay. As defined by the

doi:10.6342/NTU202401466

manufacturer, one unit is the amount of enzyme required to cleave 1 pmol of a 34-nt oligonucleotide duplex containing a single deoxyinosine site in a total reaction volume of 10 μL within 1 hour at 37°C. The correlation plot indicates a high correlation between the two methods, with an R^2 value of 0.9541 and an equation of Y=0.8216X+0.3774. Each point was an average of 4 independent measurements ± 1SD.

doi:10.6342/NTU202401466

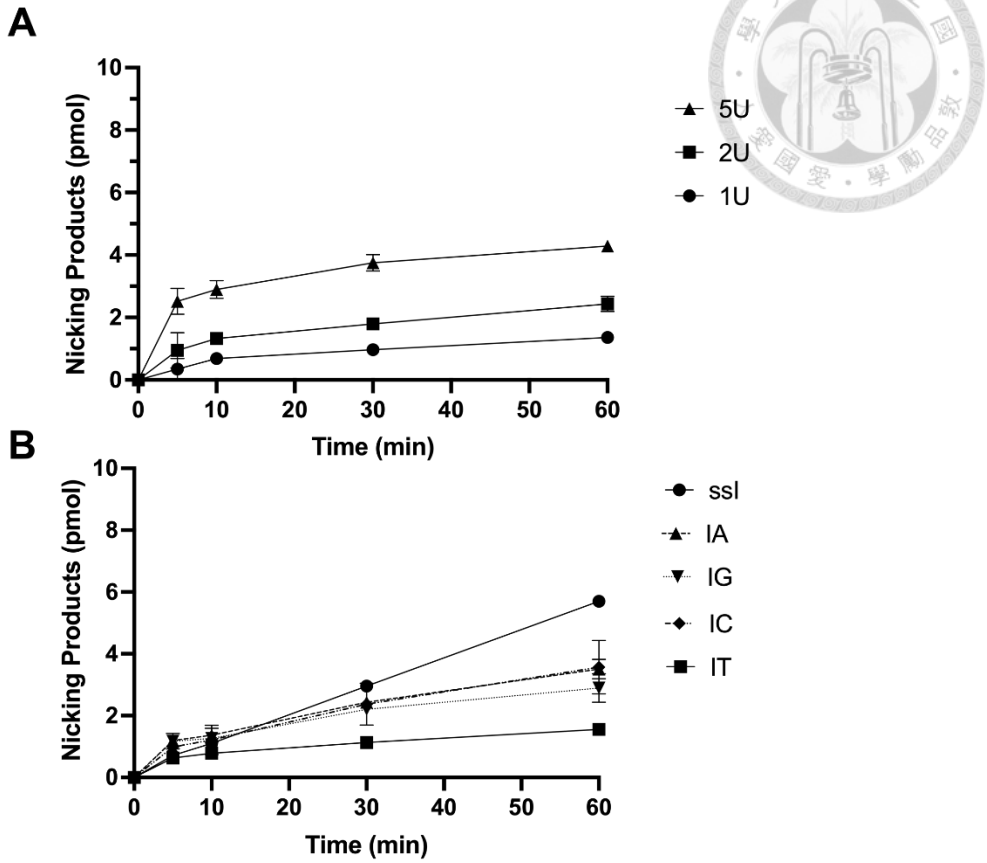


Figure 19. Endo V substrate specificity analysis using fluorescence-based gel assay

(A) Time course assay with various amounts of Endo V. A 40-TI DNA duplex was incubated with 1U, 2U, and 5U of Endo V at 37°C. Aliquots (10μl) were withdrawn from the reaction mixture at 0, 5, 10, 30, and 60 minutes, and the reaction was quenched by heat inactivation according to the manufacturer's instructions. Results showed that the reaction exhibited dose-dependency. (B) Substrate specificity of Endo V. 1U of Endo V incubated with 40-IT, 40-IA, 40-IC, 40-IG or FAM-40I substrate at 37°C for various time. The results of each time point was an average of 4 independent measurements ± 1SD.

doi:10.6342/NTU202401466

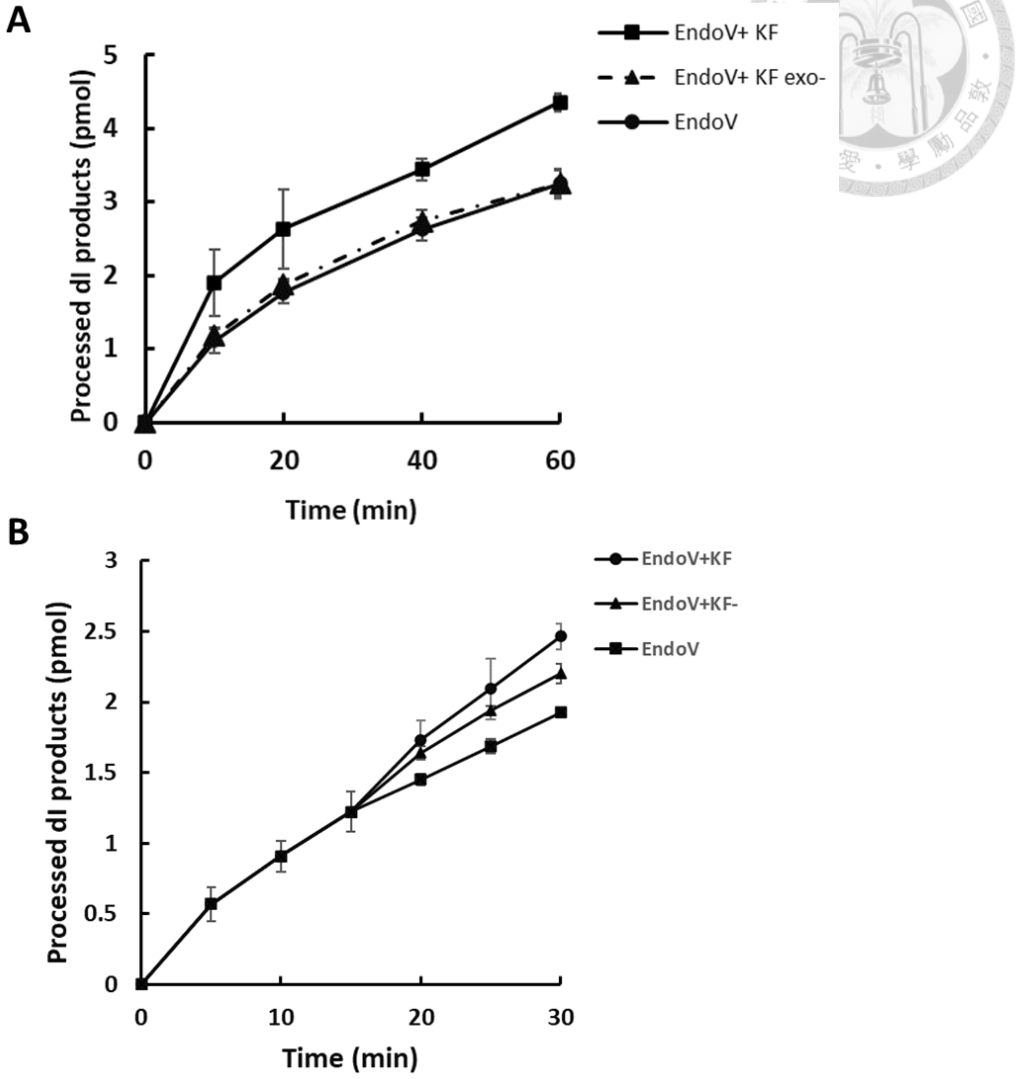


Figure 20. Pol I variants promote Endo V turnover using fluorescence-labeled gel assay

(A) Turnover assay contained 20 pmol fluorescence-labeled substrate (40-TI), 1X NEBuffer 4 (20 mM Tris-acetate, 50 mM potassium acetate, 10 mM magnesium acetate, 1 mM dithiothreitol, pH7.9 at 25°C), incubated with 1U Endo V (round circle with solid line), or co-incubated with 1U KF (rectangle with solid line) or 2U KF (exo-)

doi:10.6342/NTU202401466

(triangle with dot line). Aliquots (10μl) were withdrawn from the reaction mixture at 0, 10, 20, 40, and 60 minutes, and the reaction was quenched by heat inactivation according to the manufacturer's instructions. (B) Pol I spiked reaction contained 5 pmol fluorescence-labeled substrate (40-IT), 1X NEBuffer 4 (20 mM Tris-acetate, 50 mM potassium acetate, 10 mM magnesium acetate, 1 mM dithiothreitol, pH7.9 at 25°C), incubated with 2U Endo V (round circle with solid line) for 15 minutes, then spiked the reaction with 2U KF (circle) or 2U KF (exo-) (triangle), or no spiked (squared). Aliquots (10μl) were withdrawn from the reaction mixture at 5, 10, 15, 20, 25, and 30 minutes, and the reaction was quenched by heat inactivation according to the manufacturer's instructions. Each point was an average of 4 independent measurements ± 1SD.

doi:10.6342/NTU202401466

Tables

Table 1 The comparison of uracil DNA glycosylase detection assay

UDG Assays	Weakness	References
Radioactive labeling	Significant hazards due to radioactive materials. Time-consuming and labor-intensive preparation process.	[97]
Fluoresce Labeled Oligonucleotide	Requires careful consideration of the reaction buffer. Analysis involves restriction enzyme digestion and gel electrophoresis.	[98]
Fluorescent Resonance Energy Transfer (FRET) Based Analysis	Specific labeling requirements limit substrate design. High background levels can be problematic.	[99]
Colorimetric Based Method with G-Quadruplex DNA Probe	Precipitation of redox substances may cause high background.	[100, 101]
Electrochemical Based Method	Preparation of graphene-modified GC electrode is time-consuming and challenging.	[102]
PCR-Based Assay	Low specificity.	[103]
Enzyme-Free Electrochemical Biosensor-Based UDG Detection Assay	Tedious substrate preparation limits application.	[104]

doi:10.6342/NTU202401466

Table 2 sequences of template-primers containing a single nucleotide deletion or insertion error

Substrates	Sequences	4 ddNTPs[a]		3 dNTPs[b]	
		PR	PE	PR	PE
	/**C**\				
Pdel 1	3'-ATGGTCAGTCCTGCAACCCT TCAGCTGA	98%	--[c]	100%	--
P21A	5'-TACCAGTCAGGACGTTGGGA A				
	/**C**\				
Pdel 2	3'-ATGGTCAGTCCTGCAACCC TTCAGCTGA	88%	--	100%	--
P21A	5'-TACCAGTCAGGACGTTGGG AA				
	/**T**\				
Pdel 3	3'-ATGGTCAGTCCTGCAACC CTTCAGCTGA	60%	--	100%	--
P21A	5'-TACCAGTCAGGACGTTGG GAA				
	/**T**\				
Pdel 4	3'-ATGGTCAGTCCTGCAAC CCTTCAGCTGA	52%	--	100%	--
P21A	5'-TACCAGTCAGGACGTTG GGAA				
	/**T**\				
Pdel 5	3'-ATGGTCAGTCCTGCAA CCCTTCAGCTGA	29%	--	56%	44%
P21A	5'-TACCAGTCAGGACGTT GGGAA				
	/**T**\				
Pdel 6	3'-ATGGTCAGTCCTGCA ACCCTTCAGCTGA	32%	10%	6%	61%
P21A	5'-TACCAGTCAGGACGT TGGGAA				
	/**T**\				
Pdel 7	3'-ATGGTCAGTCCTGC AACCCTTCAGCTGA	--	22%	-	92%
P21A	5'-TACCAGTCAGGACG TTGGGAA				
	/**T**\				
Pdel 8	3'-ATGGTCAGTCCTG CAACCCTTCAGCTGA	--	77%	-	97%
P21A	5'-TACCAGTCAGGAC GTTGGGAA				
	/**T**\				
Pdel 9	3'-ATGGTCAGTCCT GCAACCCTTCAGCTGA	--	84%	-	95%
P21A	5'-TACCAGTCAGGA CGTTGGGAA				
T28	3'-ATGGTCAGTCCTGCAACCCT TCAGCTGA				
Pins 1	5'-TACCAGTCAGGACGTTGGGA A	69%	--	100%	--
	T/				
T28	3'-ATGGTCAGTCCTGCAACCC TTCAGCTGA				
Pins 2	5'-TACCAGTCAGGACGTTGGG AA	82%	--	100%	--
	T/				
T28	3'-ATGGTCAGTCCTGCAACC CTTCAGCTGA				
Pins 3	5'-TACCAGTCAGGACGTTGG GAA	51%	--	100%	--
	T/				
T28	3'-ATGGTCAGTCCTGCAAC CCTTCAGCTGA				
Pins 4	5'-TACCAGTCAGGACGTTG GGAA	13%	--	100%	--
	T/				
T28	3'-ATGGTCAGTCCTGCAA CCCTTCAGCTGA				
Pins 5	5'-TACCAGTCAGGACGTT GGGAA	13%	--	100%	--
	C/				
T28	3'-ATGGTCAGTCCTGCA ACCCTTCAGCTGA				
Pins 6	5'-TACCAGTCAGGACGT TGGGAA	3%	11%	53%	47%
	C/				
T28	3'-ATGGTCAGTCCTGC AACCCTTCAGCTGA				
Pins 7	5'-TACCAGTCAGGACG TTGGGAA	--	75%	--	100%
	C/				
T28	3'-ATGGTCAGTCCTG CAACCCTTCAGCTGA				
Pins 8	5'-TACCAGTCAGGAC GTTGGGAA	--	87%	--	100%
	T/				
T28	3'-ATGGTCAGTCCT GCAACCCTTCAGCTGA				
Pins 9	5'-TACCAGTCAGGA CGTTGGGAA	--	83%	--	100%
	T/				

[a] results from 4 ddNTPs reaction. PR, proofreading; PE, primer extension without editing.

[b] Results from 3 dNTPs reaction.

[c] -- below detection limit

doi:10.6342/NTU202401466

Table 3 U-containing substrate compositions and UDG specificities

DNA substrates	Oligo-nucleotides[a]	DNA Sequence[b]	Initial rate[c] pmol/sec	*k*cat (s^{-1})
ssU+9	U+9	5′–GACCAGTC**U**GGACGTTGG–3′	0.560 ± 0.098	35.0
ssU+3	U+3	5′–GA**U**CAGTCCGGACGTTGG–3′	0.756 ± 0.083	47.3
ssU-3	U-3	5′–GACCAGTCCGGACGT**U**GG–3′	0.448 ± 0.087	28.0
G:U	T1	3′–CTGGTCAGGCCTGCAACCG–5′	0.149 ± 0.046	9.3
	U+9	5′–GACCAGTC**U**GGACGTTGG–3′		
A:U	T2	3′–CTGGTCAGACCTGCAACCG–5′	0.053 ± 0.018	3.3
	U+9	5′–GACCAGTC**U**GGACGTTGG–3′		
G:U5'+3	T1	3′–CTGGTCAGGCCTGCAACCG–5′	0.256 ± 0.044	16.0
	U+3	5′–GA**U**CAGTCCGGACGTTGG–3′		
G:U5'+2	T3	3′–CGGGTCAGGCCTGCAACCG–5′	0.309	19.3
	U+2	5′–G**U**CCAGTCCGGACGTTGG–3′		
G:U5'+1	T4	3′–GTGGTCAGGCCTGCAACCG–5′	ND[d]	ND
	U+1	5′–**U**ACCAGTCCGGACGTTGG–3′		
G:U3'-3	T5	3′–CTGGTCAGGCCTGCAGCCG–5′	0.138 ± 0.015	8.6
	U-3	5′–GACCAGTCCGGACGT**U**GG–3′		
G:U3'-2	T6	3′–CTGGTCAGGCCTGCAAGC–5′	ND	ND
	U-2	5′–GACCAGTCCGGACGTT**U**G–3′		
G:U3'-1	T7	3′–CTGGTCAGGCCTGCAACGG–5′	ND	ND
	U-1	5′–GACCAGTCCGGACGTTG**U**–3′		

[a]The 18-nt U-containing DNA were designated by U±N, The '+ ' is counting the uracil position from 5' terminus, and the '-' is counting the uracil position from 3' terminus; for example, U+9 is the 9th position from 5' terminus, U-3 is the 3rd position from 3' terminus. The complementary 19-nt template of T1 to T7 would paired with the U-containing DNA forming G-U or A-U mismatches as indicated.

[b]Uracil bases are in bold.

[c]Reactions were using 50 pmol U substrates, 0.05U UDG, in 10μl as described in the Materials and Methods.

[d] ND, not detectable.

doi:10.6342/NTU202401466

Table 4 The sequence for MS normalization

Length	DNA sequence
15-nt	5′-CGTAAGTCAGGATCC-3′
16-nt	5′-CGTAAGTCAGGATCCG-3′
17-nt	5′-CGTAAGTCAGGATCCGT-3′
18-nt	5′-CGTAAGTCAGGATCCGTG-3′
19-nt	5′-CGTAAGTCAGGATCCGTGA-3′
20-nt	5′-CGTAAGTCAGGATCCGTGAT-3′
21-nt	5′-CGTAAGTCAGGATCCGTGATC-3′
22-nt	5′-CGTAAGTCAGGATCCGTGATCT-3′
23-nt	5′-CGTAAGTCAGGATCCGTGATCTA-3′

doi:10.6342/NTU202401466

Table 5 dI-containing substrate for MS-based Endo V analysis

dI substrate	DNA sequence
21-IT	5′-TACCAGTCAGGACGT**I**GCTAA-3′
	3′-ATGGTCAGTCCTGCATCGATTCAGCTGA-5′
21-IA	5′-TACCAGTCAGGACGT**I**GCTAA-3′
	3′-ATGGTCAGTCCTGCAACGATTCAGCTGA-5′
21-IG	5′-TACCAGTCAGGACGT**I**GCTAA-3′
	3′-ATGGTCAGTCCTGCAGCGATTCAGCTGA-5′
21-IC	5′-TACCAGTCAGGACGT**I**GCTAA-3′
	3′-ATGGTCAGTCCTGCACCGATTCAGCTGA-5′
21-IA*	5′-TACCAGTCAGGACGT**I**GGGAA-3′
	3′-ATGGTCAGTCCTGCAACCCTTCAGCTGA-5′

The red-bold I: inosine.

Under lined: calibrator oligonucleotide.

doi:10.6342/NTU202401466

Table 6 dI-containing substrate for fluorescence-based Endo V analysis

dI substrate	DNA sequence
34-IT	FAM-5′-GATTTCATTTTTATT**I**ATAACTTTACTTATATTG-3′
	3′-CTAAAGTAAAAATAATTATTGAAATGAATATAAC-5′
FAM-40I	FAM-5′-GATGCCCTTATTATTGG**I**CTCGAGCGTAATGAATGGATTC-3′
40-IT	FAM-5′-GATGCCCTTATTATTGG**I**CTCGAGCGTAATGAATGGATTC-3′
	3′-CTACGGGAATAATAACCTGCCATGGCATTACTTACCTAAG-5′
40-IA	FAM-5′-GATGCCCTTATTATTGG**I**CTCGAGCGTAATGAATGGATTC-3′
	3′-CTACGGGAATAATAACCAGCCATGGCATTACTTACCTAAG-5′
40-IC	FAM-5′-GATGCCCTTATTATTGG**I**CTCGAGCGTAATGAATGGATTC-3′
	3′-CTACGGGAATAATAACCCGCCATGGCATTACTTACCTAAG-5′
40-IG	FAM-5′-GATGCCCTTATTATTGG**I**CTCGAGCGTAATGAATGGATTC-3′
	3′-CTACGGGAATAATAACCGGCCATGGCATTACTTACCTAAG-5′
21-IT-FAM	FAM-5′-TACCAGTCAGGACGT**I**GCTAA-3′
	3′-ATGGTCAGTCCTGCATCGATTCAGCTGA-5′

The red-bold I: inosine.

doi:10.6342/NTU202401466

Table 7 Example of absorbance monitoring at λ = 260 nm for substrate duplex preparation.

	T5	U-3	T5/U-3 duplex [a]
A260 at 95°C	0.247	0.220	0.468 [b]
A260 at 70°C	0.200	0.219	0.386 [c]
A260 at 25°C	0.183 [d]	0.205	0.375

[a] Hybridization of template and substrate oligonucleotides was following established protocol in reaction buffer of 10 mM $MgCl_2$, 10 mM Tris-Cl (pH 7.5), 1 μM of T5, U-3, and T5/U-3 duplex were subjected A260 measurement at temperatures indicated.

[b] At 95°C all oligonucleotide displayed typical random coil single-stranded A260 absorbance.

[c] At 70 °C, decreasing of A260 of T5/U-3 duplex correspondent to the formation of heteroduplex.

[d] Decreasing A260 of T5 may be attributed to forming hairpin type of structure.

doi:10.6342/NTU202401466

Table 8 Comparison of MS-based and PAGE-based analysis for Endo V

	MS-based	PAGE-based
Substrate	Non-labeled	Fluorescence-labeled
Length of substrate	15 to 32-nt	10~1000 bp (depends on polyacrylamide %)
Sample test	384 samples (batch)	12 samples (gels) (4 gels/ batch)
Resolution	1 Dalton	1 bp
Background	low	high

doi:10.6342/NTU202401466

References

1 Karas, M., Bachmann, D. & Hillenkamp, F. Influence of the wavelength in high-irradiance ultraviolet laser desorption mass spectrometry of organic molecules. *Analytical Chemistry.* **57** (14), 2935-2939, doi:10.1021/ac00291a042, (1985).

2 Karas, M. & Hillenkamp, F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical Chemistry.* **60** (20), 2299-2301, doi:10.1021/ac00171a028, (1988).

3 Yates, J. R., 3rd. Mass spectrometry and the age of the proteome. *J Mass Spectrom.* **33** (1), 1-19, doi:10.1002/(sici)1096-9888(199801)33:1<1::Aid-jms624>3.0.Co;2-9, (1998).

4 Pareige, C., Lefebvre-Ulrikson, W., Vurpillot, F. & Sauvage, X. in *Atom Probe Tomography* eds Williams Lefebvre-Ulrikson, François Vurpillot, & Xavier Sauvage) 123-154 (Academic Press, 2016).

5 Brock, A., Rodriguez, N. & Zare, R. N. Hadamard Transform Time-of-Flight Mass Spectrometry. *Analytical Chemistry.* **70** (18), 3735-3741, doi:10.1021/ac9804036, (1998).

6 Singhal, N., Kumar, M., Kanaujia, P. K. & Virdi, J. S. MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. *Front Microbiol.* **6** 791, doi:10.3389/fmicb.2015.00791, (2015).

7 Fuh, M. M., Heikaus, L. & Schlüter, H. MALDI mass spectrometry in medical research and diagnostic routine laboratories. *International Journal of Mass Spectrometry.* **416** 96-109, doi:10.1016/j.ijms.2016.10.004, (2017).

8 Coolen, M. W., Statham, A. L., Gardiner-Garden, M. & Clark, S. J. Genomic profiling of CpG methylation and allelic specificity using quantitative high-throughput mass spectrometry: critical evaluation and improvements. *Nucleic Acids Res.* **35** (18), e119, doi:10.1093/nar/gkm662, (2007).

9 Martisova, A. *et al.* DNA Methylation in Solid Tumors: Functions and Methods of Detection. *Int J Mol Sci.* **22** (8), doi:10.3390/ijms22084247, (2021).

10 Hwa Yun, B., Guo, J., Bellamri, M. & Turesky, R. J. DNA adducts: Formation, biological effects, and new biospecimens for mass spectrometric measurements in humans. *Mass Spectrom Rev.* **39** (1-2), 55-82, doi:10.1002/mas.21570, (2020).

doi:10.6342/NTU202401466

11 Garaguso, I., Halter, R., Krzeminski, J., Amin, S. & Borlak, J. Method for the rapid detection and molecular characterization of DNA alkylating agents by MALDI-TOF mass spectrometry. *Anal Chem.* **82** (20), 8573-8582, doi:10.1021/ac101568h, (2010).

12 Su, K. Y. *et al.* Mutational monitoring of EGFR T790M in cfDNA for clinical outcome prediction in EGFR-mutant lung adenocarcinoma. *PLoS One.* **13** (11), e0207001, doi:10.1371/journal.pone.0207001, (2018).

13 Haff, L. A. & Smirnov, I. P. Single-nucleotide polymorphism identification assays using a thermostable DNA polymerase and delayed extraction MALDI-TOF mass spectrometry. *Genome Res.* **7** (4), 378-388, doi:10.1101/gr.7.4.378, (1997).

14 Haff, L. A. & Smirnov, I. P. Multiplex genotyping of PCR products with MassTag-labeled primers. *Nucleic Acids Res.* **25** (18), 3749-3750, doi:10.1093/nar/25.18.3749, (1997).

15 Jurinke, C., Oeth, P. & van den Boom, D. MALDI-TOF mass spectrometry: a versatile tool for high-performance DNA analysis. *Mol Biotechnol.* **26** (2), 147-164, doi:10.1385/mb:26:2:147, (2004).

16 Blondal, T. *et al.* A novel MALDI-TOF based methodology for genotyping single nucleotide polymorphisms. *Nucleic Acids Res.* **31** (24), e155, doi:10.1093/nar/gng156, (2003).

17 Su, K. Y. *et al.* Application of single nucleotide extension and MALDI-TOF mass spectrometry in proofreading and DNA repair assay. *DNA Repair (Amst).* **61** 63-75, doi:10.1016/j.dnarep.2017.11.011, (2018).

18 Su, K. Y. *et al.* Proofreading and DNA Repair Assay Using Single Nucleotide Extension and MALDI-TOF Mass Spectrometry Analysis. *J Vis Exp.* (136), doi:10.3791/57862, (2018).

19 Su, K.-Y. *et al.* DNA polymerase I proofreading exonuclease activity is required for endonuclease V repair pathway both in vitro and in vivo. *DNA Repair.* **64** 59-67, doi:10.1016/j.dnarep.2018.02.005, (2018).

20 Duval, A. & Hamelin, R. Mutations at coding repeat sequences in mismatch repair-deficient human cancers: toward a new concept of target genes for instability. *Cancer Res.* **62** (1), 2447-2454 (2002).

21 Duval, A. *et al.* Frequent frameshift mutations of the TCF-4 gene in colorectal cancers with microsatellite instability. *Cancer Res.* **59** (2), 4213-4215 (1999).

22 Minetti, C. A., Remeta, D. P., Dickstein, R. & Breslauer, K. J. Energetic signatures of single base bulges: thermodynamic consequences and biological

doi:10.6342/NTU202401466

implications. *Nucleic Acids Res.* **38** (1), 97-116, doi:10.1093/nar/gkp1036, (2010).

23 Streisinger, G. *et al.* Frameshift mutations and the genetic code. This paper is dedicated to Professor Theodosius Dobzhansky on the occasion of his 66th birthday. *Cold Spring Harb Symp Quant Biol.* **31** 77-84, doi:10.1101/sqb.1966.031.01.014, (1966).

24 Garcia-Diaz, M., Bebenek, K., Krahn, J. M., Pedersen, L. C. & Kunkel, T. A. Structural analysis of strand misalignment during DNA synthesis by a human DNA polymerase. *Cell.* **124** (2), 331-342, doi:10.1016/j.cell.2005.10.039, (2006).

25 Ling, H., Boudsocq, F., Woodgate, R. & Yang, W. Crystal structure of a Y-family DNA polymerase in action: a mechanism for error-prone and lesion-bypass replication. *Cell.* **107** (1), 91-102, doi:10.1016/s0092-8674(01)00515-3, (2001).

26 Calos, M. P. & Miller, J. H. Genetic and sequence analysis of frameshift mutations induced by ICR-191. *J Mol Biol.* **153** (1), 39-64, doi:10.1016/0022-2836(81)90525-8, (1981).

27 McCulloch, S. D. & Kunkel, T. A. The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. *Cell Res.* **18** (1), 148-161, doi:10.1038/cr.2008.4, (2008).

28 Kroutil, L. C., Register, K., Bebenek, K. & Kunkel, T. A. Exonucleolytic proofreading during replication of repetitive DNA. *Biochemistry.* **35** (3), 1046-1053, doi:10.1021/bi952178h, (1996).

29 Joyce, C. M. & Steitz, T. A. Function and structure relationships in DNA polymerases. *Annu Rev Biochem.* **63** 777-822, doi:10.1146/annurev.bi.63.070194.004021, (1994).

30 Garcia-Diaz, M. & Bebenek, K. Multiple functions of DNA polymerases. *CRC Crit Rev Plant Sci.* **26** (2), 105-122, doi:10.1080/07352680701252817, (2007).

31 Bessman, M. J., Kornberg, A., Lehman, I. R. & Simms, E. S. Enzymic synthesis of deoxyribonucleic acid. *Biochim Biophys Acta.* **21** (1), 197-198, doi:10.1016/0006-3002(56)90127-5, (1956).

32 De Lucia, P. & Cairns, J. Isolation of an E. coli strain with a mutation affecting DNA polymerase. *Nature.* **224** (5225), 1164-1166, doi:10.1038/2241164a0, (1969).

doi:10.6342/NTU202401466

33 Makiela-Dzbenska, K. *et al.* Role of Escherichia coli DNA polymerase I in chromosomal DNA replication fidelity. *Mol Microbiol.* **74** (5), 1114-1127, doi:10.1111/j.1365-2958.2009.06921.x, (2009).

34 Knippers, R. DNA polymerase II. *Nature.* **228** (5276), 1050-1053, doi:10.1038/2281050a0, (1970).

35 Escarceller, M. *et al.* Involvement of Escherichia coli DNA polymerase II in response to oxidative damage and adaptive mutation. *J Bacteriol.* **176** (20), 6221-6228, doi:10.1128/jb.176.20.6221-6228.1994, (1994).

36 Qiu, Z. & Goodman, M. F. The Escherichia coli polB locus is identical to dinA, the structural gene for DNA polymerase II. Characterization of Pol II purified from a polB mutant. *J Biol Chem.* **272** (13), 8611-8617, doi:10.1074/jbc.272.13.8611, (1997).

37 Bonner, C. A., Hays, S., McEntee, K. & Goodman, M. F. DNA polymerase II is encoded by the DNA damage-inducible dinA gene of Escherichia coli. *Proc Natl Acad Sci U S A.* **87** (19), 7663-7667, doi:10.1073/pnas.87.19.7663, (1990).

38 Iwasaki, H., Nakata, A., Walker, G. C. & Shinagawa, H. The Escherichia coli polB gene, which encodes DNA polymerase II, is regulated by the SOS system. *J Bacteriol.* **172** (11), 6268-6273, doi:10.1128/jb.172.11.6268-6273.1990, (1990).

39 Banach-Orlowska, M., Fijalkowska, I. J., Schaaper, R. M. & Jonczyk, P. DNA polymerase II as a fidelity factor in chromosomal DNA synthesis in Escherichia coli. *Mol Microbiol.* **58** (1), 61-70, doi:10.1111/j.1365-2958.2005.04805.x, (2005).

40 Gefter, M. L., Hirota, Y., Kornberg, T., Wechsler, J. A. & Barnoux, C. Analysis of DNA polymerases II and 3 in mutants of Escherichia coli thermosensitive for DNA synthesis. *Proc Natl Acad Sci U S A.* **68** (12), 3150-3153, doi:10.1073/pnas.68.12.3150, (1971).

41 Birge, E. A. in *Bacterial and Bacteriophage Genetics* (ed Edward A. Birge) 19-63 (Springer New York, 2000).

42 Yao, N. Y., Georgescu, R. E., Finkelstein, J. & O'Donnell, M. E. Single-molecule analysis reveals that the lagging strand increases replisome processivity but slows replication fork progression. *Proc Natl Acad Sci U S A.* **106** (32), 13236-13241, doi:10.1073/pnas.0906157106, (2009).

43 Lewis, J. S., Jergic, S. & Dixon, N. E. in *The Enzymes* Vol. 39 eds Laurie S. Kaguni & Marcos Túlio Oliveira) 31-88 (Academic Press, 2016).

doi:10.6342/NTU202401466

44 Tang, M. *et al.* Roles of E. coli DNA polymerases IV and V in lesion-targeted and untargeted SOS mutagenesis. *Nature.* **404** (6781), 1014-1018, doi:10.1038/35010020, (2000).

45 Goodman, M. F. Error-prone repair DNA polymerases in prokaryotes and eukaryotes. *Annu Rev Biochem.* **71** 17-50, doi:10.1146/annurev.biochem.71.083101.124707, (2002).

46 Kim, S. R., Matsui, K., Yamada, M., Gruz, P. & Nohmi, T. Roles of chromosomal and episomal dinB genes encoding DNA pol IV in targeted and untargeted mutagenesis in Escherichia coli. *Mol Genet Genomics.* **266** (2), 207-215, doi:10.1007/s004380100541, (2001).

47 Wagner, J. *et al.* The dinB gene encodes a novel E. coli DNA polymerase, DNA pol IV, involved in mutagenesis. *Mol Cell.* **4** (2), 281-286, doi:10.1016/s1097-2765(00)80376-7, (1999).

48 Witkin, E. M., McCall, J. O., Volkert, M. R. & Wermundsen, I. E. Constitutive expression of SOS functions and modulation of mutagenesis resulting from resolution of genetic instability at or near the recA locus of Escherichia coli. *Mol Gen Genet.* **185** (1), 43-50, doi:10.1007/bf00333788, (1982).

49 Wrzesiński, M., Nowosielska, A., Nieminuszczy, J. & Grzesiuk, E. Effect of SOS-induced Pol II, Pol IV, and Pol V DNA polymerases on UV-induced mutagenesis and MFD repair in Escherichia coli cells. *Acta Biochim Pol.* **52** (1), 139-147 (2005).

50 Joyce, C. M. & Grindley, N. D. Method for determining whether a gene of Escherichia coli is essential: application to the polA gene. *J Bacteriol.* **158** (2), 636-643, doi:10.1128/jb.158.2.636-643.1984, (1984).

51 Ollis, D. L., Brick, P., Hamlin, R., Xuong, N. G. & Steitz, T. A. Structure of large fragment of Escherichia coli DNA polymerase I complexed with dTMP. *Nature.* **313** (6005), 762-766, doi:10.1038/313762a0, (1985).

52 Freemont, P. S., Ollis, D. L., Steitz, T. A. & Joyce, C. M. A domain of the Klenow fragment of Escherichia coli DNA polymerase I has polymerase but no exonuclease activity. *Proteins.* **1** (1), 66-73, doi:10.1002/prot.340010111, (1986).

53 Polesky, A. H., Dahlberg, M. E., Benkovic, S. J., Grindley, N. D. & Joyce, C. M. Side chains involved in catalysis of the polymerase reaction of DNA polymerase I from Escherichia coli. *J Biol Chem.* **267** (12), 8417-8428 (1992).

doi:10.6342/NTU202401466

54 Polesky, A. H., Steitz, T. A., Grindley, N. D. & Joyce, C. M. Identification of residues critical for the polymerase activity of the Klenow fragment of DNA polymerase I from Escherichia coli. *J Biol Chem.* **265** (24), 14579-14591 (1990).

55 Catalano, C. E., Allen, D. J. & Benkovic, S. J. Interaction of Escherichia coli DNA polymerase I with azidoDNA and fluorescent DNA probes: identification of protein-DNA contacts. *Biochemistry.* **29** (15), 3612-3621, doi:10.1021/bi00467a004, (1990).

56 Kuchta, R. D., Mizrahi, V., Benkovic, P. A., Johnson, K. A. & Benkovic, S. J. Kinetic mechanism of DNA polymerase I (Klenow). *Biochemistry.* **26** (25), 8410-8417, doi:10.1021/bi00399a057, (1987).

57 Derbyshire, V. *et al.* Genetic and crystallographic studies of the 3',5'-exonucleolytic site of DNA polymerase I. *Science.* **240** (4849), 199-201, doi:10.1126/science.2832946, (1988).

58 Beese, L. S., Derbyshire, V. & Steitz, T. A. Structure of DNA polymerase I Klenow fragment bound to duplex DNA. *Science.* **260** (5106), 352-355, doi:10.1126/science.8469987, (1993).

59 Joyce, C. M. & Steitz, T. A. DNA polymerase I: from crystal structure to function via genetics. *Trends in Biochemical Sciences.* **12** 288-292, doi:10.1016/0968-0004(87)90143-5, (1987).

60 Allen, D. J., Darke, P. L. & Benkovic, S. J. Fluorescent oligonucleotides and deoxynucleotide triphosphates: preparation and their interaction with the large (Klenow) fragment of Escherichia coli DNA polymerase I. *Biochemistry.* **28** (11), 4601-4607, doi:10.1021/bi00437a014, (1989).

61 Turner, R. M., Jr., Grindley, N. D. & Joyce, C. M. Interaction of DNA polymerase I (Klenow fragment) with the single-stranded template beyond the site of synthesis. *Biochemistry.* **42** (8), 2373-2385, doi:10.1021/bi026566c, (2003).

62 Patel, P. H., Kawate, H., Adman, E., Ashbach, M. & Loeb, L. A. A single highly mutable catalytic site amino acid is critical for DNA polymerase fidelity. *J Biol Chem.* **276** (7), 5044-5051, doi:10.1074/jbc.M008701200, (2001).

63 Steitz, T. A. DNA polymerases: structural diversity and common mechanisms. *J Biol Chem.* **274** (25), 17395-17398, doi:10.1074/jbc.274.25.17395, (1999).

64 Ludmann, S. & Marx, A. Getting it Right: How DNA Polymerases Select the Right Nucleotide. *Chimia (Aarau).* **70** (3), 203-206, doi:10.2533/chimia.2016.203, (2016).

doi:10.6342/NTU202401466

65 Xia, S. & Konigsberg, W. H. RB69 DNA polymerase structure, kinetics, and fidelity. *Biochemistry.* **53** (17), 2752-2767, doi:10.1021/bi4014215, (2014).
66 Wang, M. *et al.* Insights into base selectivity from the 1.8 Å resolution structure of an RB69 DNA polymerase ternary complex. *Biochemistry.* **50** (4), 581-590, doi:10.1021/bi101192f, (2011).
67 Lovett, S. T. The DNA Exonucleases of Escherichia coli. *EcoSal Plus.* **4** (2), doi:10.1128/ecosalplus.4.4.7, (2011).
68 Wong, I., Patel, S. S. & Johnson, K. A. An induced-fit kinetic mechanism for DNA replication fidelity: direct measurement by single-turnover kinetics. *Biochemistry.* **30** (2), 526-537, doi:10.1021/bi00216a030, (1991).
69 Green, M. R. & Sambrook, J. E. coli DNA Polymerase I and the Klenow Fragment. *Cold Spring Harb Protoc.* **2020** (5), 100743, doi:10.1101/pdb.top100743, (2020).
70 Cowart, M., Gibson, K. J., Allen, D. J. & Benkovic, S. J. DNA substrate structural requirements for the exonuclease and polymerase activities of procaryotic and phage DNA polymerases. *Biochemistry.* **28** (5), 1975-1983, doi:10.1021/bi00431a004, (1989).
71 Freemont, P. S., Friedman, J. M., Beese, L. S., Sanderson, M. R. & Steitz, T. A. Cocrystal structure of an editing complex of Klenow fragment with DNA. *Proc Natl Acad Sci U S A.* **85** (23), 8924-8928, doi:10.1073/pnas.85.23.8924, (1988).
72 Tippin, B., Kobayashi, S., Bertram, J. G. & Goodman, M. F. To slip or skip, visualizing frameshift mutation dynamics for error-prone DNA polymerases. *J Biol Chem.* **279** (44), 45360-45368, doi:10.1074/jbc.M408600200, (2004).
73 Dangerfield, T. L., Kirmizialtin, S. & Johnson, K. A. Substrate specificity and proposed structure of the proofreading complex of T7 DNA polymerase. *J Biol Chem.* **298** (3), 101627, doi:10.1016/j.jbc.2022.101627, (2022).
74 Lee, C. C. *et al.* The excision of 3' penultimate errors by DNA polymerase I and its role in endonuclease V-mediated DNA repair. *DNA Repair (Amst).* **12** (11), 899-911, doi:10.1016/j.dnarep.2013.08.003, (2013).
75 Krokan, H. E., Drabløs, F. & Slupphaug, G. Uracil in DNA--occurrence, consequences and repair. *Oncogene.* **21** (58), 8935-8948, doi:10.1038/sj.onc.1205996, (2002).
76 Hagen, L. *et al.* Genomic uracil and human disease. *Exp Cell Res.* **312** (14), 2666-2672, doi:10.1016/j.yexcr.2006.06.015, (2006).
77 Kuraoka, I. *et al.* Effects of endogenous DNA base lesions on transcription elongation by mammalian RNA polymerase II. Implications for transcription-

doi:10.6342/NTU202401466

coupled DNA repair and transcriptional mutagenesis. *J Biol Chem.* **278** (9), 7294-7299, doi:10.1074/jbc.M208102200, (2003).

78 Tye, B. K., Chien, J., Lehman, I. R., Duncan, B. K. & Warner, H. R. Uracil incorporation: a source of pulse-labeled DNA fragments in the replication of the Escherichia coli chromosome. *Proc Natl Acad Sci U S A.* **75** (1), 233-237, doi:10.1073/pnas.75.1.233, (1978).

79 Mosbaugh, D. W. & Bennett, S. E. Uracil-excision DNA repair. *Prog Nucleic Acid Res Mol Biol.* **48** 315-370, doi:10.1016/s0079-6603(08)60859-4, (1994).

80 Ingraham, H. A., Dickey, L. & Goulian, M. DNA fragmentation and cytotoxicity from increased cellular deoxyuridylate. *Biochemistry.* **25** (11), 3225-3230, doi:10.1021/bi00359a022, (1986).

81 Duncan, B. K. & Miller, J. H. Mutagenic deamination of cytosine residues in DNA. *Nature.* **287** (5782), 560-561, doi:10.1038/287560a0, (1980).

82 Lindahl, T. Instability and decay of the primary structure of DNA. *Nature.* **362** (6422), 709-715, doi:10.1038/362709a0, (1993).

83 Frederico, L. A., Kunkel, T. A. & Shaw, B. R. A sensitive genetic assay for the detection of cytosine deamination: determination of rate constants and the activation energy. *Biochemistry.* **29** (10), 2532-2537, doi:10.1021/bi00462a015, (1990).

84 Sousa, M. M., Krokan, H. E. & Slupphaug, G. DNA-uracil and human pathology. *Mol Aspects Med.* **28** (3-4), 276-306, doi:10.1016/j.mam.2007.04.006, (2007).

85 Jacobs, A. L. & Schär, P. DNA glycosylases: in DNA repair and beyond. *Chromosoma.* **121** (1), 1-20, doi:10.1007/s00412-011-0347-4, (2012).

86 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.* **66** (6), 981-993, doi:10.1007/s00018-009-8736-z, (2009).

87 Li, M. & Wilson, D. M., 3rd. Human apurinic/apyrimidinic endonuclease 1. *Antioxid Redox Signal.* **20** (4), 678-707, doi:10.1089/ars.2013.5492, (2014).

88 Lindahl, T. Uracil-DNA glycosylase from Escherichia coli. *Methods Enzymol.* **65** (1), 284-290, doi:10.1016/s0076-6879(80)65038-1, (1980).

89 Richardson, C. C., Lehman, I. R. & Kornberg, A. A DEOXYRIBONUCLEIC ACID PHOSPHATASE-EXONUCLEASE FROM ESCHERICHIA COLI. II. CHARACTERIZATION OF THE EXONUCLEASE ACTIVITY. *J Biol Chem.* **239** 251-258 (1964).

doi:10.6342/NTU202401466

90 Bailly, V. & Verly, W. G. The multiple activities of Escherichia coli endonuclease IV and the extreme lability of 5'-terminal base-free deoxyribose 5-phosphates. *Biochem J.* **259** (3), 761-768, doi:10.1042/bj2590761, (1989).

91 Lyamichev, V., Brow, M. A. & Dahlberg, J. E. Structure-specific endonucleolytic cleavage of nucleic acids by eubacterial DNA polymerases. *Science.* **260** (5109), 778-783, doi:10.1126/science.7683443, (1993).

92 Chen, R., Wang, H. & Mansky, L. M. Roles of uracil-DNA glycosylase and dUTPase in virus replication. *J Gen Virol.* **83** (Pt 10), 2339-2345, doi:10.1099/0022-1317-83-10-2339, (2002).

93 Olsen, L. C., Aasland, R., Wittwer, C. U., Krokan, H. E. & Helland, D. E. Molecular cloning of human uracil-DNA glycosylase, a highly conserved DNA repair enzyme. *Embo j.* **8** (10), 3121-3125, doi:10.1002/j.1460-2075.1989.tb08464.x, (1989).

94 Chung, J. H. *et al.* A novel uracil-DNA glycosylase family related to the helix-hairpin-helix DNA glycosylase superfamily. *Nucleic Acids Res.* **31** (8), 2045-2055, doi:10.1093/nar/gkg319, (2003).

95 Duncan, B. K. & Weiss, B. Specific mutator effects of ung (uracil-DNA glycosylase) mutations in Escherichia coli. *J Bacteriol.* **151** (2), 750-755, doi:10.1128/jb.151.2.750-755.1982, (1982).

96 Lutsenko, E. & Bhagwat, A. S. The role of the Escherichia coli mug protein in the removal of uracil and 3,N(4)-ethenocytosine from DNA. *J Biol Chem.* **274** (43), 31034-31038, doi:10.1074/jbc.274.43.31034, (1999).

97 Lee, H. W., Dominy, B. N. & Cao, W. New family of deamination repair enzymes in uracil-DNA glycosylase superfamily. *J Biol Chem.* **286** (36), 31282-31287, doi:10.1074/jbc.M111.249524, (2011).

98 Nilsen, H. *et al.* Nuclear and mitochondrial uracil-DNA glycosylases are generated by alternative splicing and transcription from different positions in the UNG gene. *Nucleic Acids Res.* **25** (4), 750-755, doi:10.1093/nar/25.4.750, (1997).

99 Pettersen, H. S. *et al.* Uracil-DNA glycosylases SMUG1 and UNG2 coordinate the initial steps of base excision repair by distinct mechanisms. *Nucleic Acids Res.* **35** (12), 3879-3892, doi:10.1093/nar/gkm372, (2007).

100 Aravind, L. & Koonin, E. V. The alpha/beta fold uracil DNA glycosylases: a common origin with diverse fates. *Genome Biol.* **1** (4), Research0007, doi:10.1186/gb-2000-1-4-research0007, (2000).

doi:10.6342/NTU202401466

101 Oliver, B. in *Antibody Engineering* (ed Böldicke Thomas) Ch. 1 (IntechOpen, 2018).

102 Seal, G., Brech, K., Karp, S. J., Cool, B. L. & Sirover, M. A. Immunological lesions in human uracil DNA glycosylase: association with Bloom syndrome. *Proc Natl Acad Sci U S A.* **85** (7), 2339-2343, doi:10.1073/pnas.85.7.2339, (1988).

103 Duschinsky, R., Pleven, E. & Heidelberger, C. THE SYNTHESIS OF 5-FLUOROPYRIMIDINES. *Journal of the American Chemical Society.* **79** (16), 4559-4560, doi:10.1021/ja01573a087, (1957).

104 Longley, D. B., Harkin, D. P. & Johnston, P. G. 5-fluorouracil: mechanisms of action and clinical strategies. *Nat Rev Cancer.* **3** (5), 330-338, doi:10.1038/nrc1074, (2003).

105 Tanaka, M., Yoshida, S., Saneyoshi, M. & Yamaguchi, T. Utilization of 5-fluoro-2'-deoxyuridine triphosphate and 5-fluoro-2'-deoxycytidine triphosphate in DNA synthesis by DNA polymerases alpha and beta from calf thymus. *Cancer Res.* **41** (10), 4132-4135 (1981).

106 Fischer, F., Baerenfaller, K. & Jiricny, J. 5-Fluorouracil is efficiently removed from DNA by the base excision and mismatch repair systems. *Gastroenterology.* **133** (6), 1858-1868, doi:10.1053/j.gastro.2007.09.003, (2007).

107 Mauro, D. J., De Riel, J. K., Tallarida, R. J. & Sirover, M. A. Mechanisms of excision of 5-fluorouracil by uracil DNA glycosylase in normal human cells. *Mol Pharmacol.* **43** (6), 854-857 (1993).

108 Andersen, S. *et al.* Incorporation of dUMP into DNA is a major source of spontaneous DNA damage, while excision of uracil is not required for cytotoxicity of fluoropyrimidines in mouse embryonic fibroblasts. *Carcinogenesis.* **26** (3), 547-555, doi:10.1093/carcin/bgh347, (2005).

109 An, Q., Robins, P., Lindahl, T. & Barnes, D. E. 5-Fluorouracil incorporated into DNA is excised by the Smug1 DNA glycosylase to reduce drug cytotoxicity. *Cancer Res.* **67** (3), 940-945, doi:10.1158/0008-5472.Can-06-2960, (2007).

110 Krokan, H. & Wittwer, C. U. Uracil DNa-glycosylase from HeLa cells: general properties, substrate specificity and effect of uracil analogs. *Nucleic Acids Res.* **9** (11), 2599-2613, doi:10.1093/nar/9.11.2599, (1981).

111 Kreklau, E. L. *et al.* A novel fluorometric oligonucleotide assay to measure O(6)-methylguanine DNA methyltransferase, methylpurine DNA glycosylase, 8-oxoguanine DNA glycosylase and abasic endonuclease activities: DNA repair status in human breast carcinoma cells overexpressing methylpurine DNA

doi:10.6342/NTU202401466

glycosylase. *Nucleic Acids Res.* **29** (12), 2558-2566, doi:10.1093/nar/29.12.2558, (2001).

112 Ono, T. *et al.* Direct fluorescence monitoring of DNA base excision repair. *Angew Chem Int Ed Engl.* **51** (7), 1689-1692, doi:10.1002/anie.201108135, (2012).

113 Gao, X. *et al.* A G-quadruplex DNA structure resolvase, RHAU, is essential for spermatogonia differentiation. *Cell Death Dis.* **6** (1), e1610, doi:10.1038/cddis.2014.571, (2015).

114 Zhao, H., Hu, W., Jing, J. & Zhang, X. One-step G-quadruplex-based fluorescence resonance energy transfer sensing method for ratiometric detection of uracil-DNA glycosylase activity. *Talanta.* **221** 121609, doi:10.1016/j.talanta.2020.121609, (2021).

115 Jiao, F. *et al.* A novel and label-free biosensors for uracil-DNA glycosylase activity based on the electrochemical oxidation of guanine bases at the graphene modified electrode. *Talanta.* **147** 98-102, doi:10.1016/j.talanta.2015.09.045, (2016).

116 Squillaro, T. *et al.* A rapid, safe, and quantitative in vitro assay for measurement of uracil-DNA glycosylase activity. *J Mol Med (Berl).* **97** (7), 991-1001, doi:10.1007/s00109-019-01788-8, (2019).

117 Tian, G., Li, W., Liu, B., Xiao, M. & Xia, Q. An enzyme-free electrochemical biosensor based on NiCoP@PtCu nanozyme and multi-MNAzyme junctions for ultrasensitive Uracil-DNA glycosylase detection. *Sensors and Actuators B: Chemical.* **379** 133224, doi:10.1016/j.snb.2022.133224, (2023).

118 Myrnes, B., Guddal, P. H. & Krokan, H. Metabolism of dITP in HeLa cell extracts, incorporation into DNA by isolated nuclei and release of hypoxanthine from DNA by a hypoxanthine-DNA glycosylase activity. *Nucleic Acids Res.* **10** (12), 3693-3701, doi:10.1093/nar/10.12.3693, (1982).

119 Karran, P. & Lindahl, T. Hypoxanthine in deoxyribonucleic acid: generation by heat-induced hydrolysis of adenine residues and release in free form by a deoxyribonucleic acid glycosylase from calf thymus. *Biochemistry.* **19** (26), 6005-6011, doi:10.1021/bi00567a010, (1980).

120 Taghizadeh, K. *et al.* Quantification of DNA damage products resulting from deamination, oxidation and reaction with products of lipid peroxidation by liquid chromatography isotope dilution tandem mass spectrometry. *Nat Protoc.* **3** (8), 1287-1298, doi:10.1038/nprot.2008.119, (2008).

doi:10.6342/NTU202401466

121 Behmanesh, M. *et al.* Characterization of the structure and expression of mouse Itpa gene and its related sequences in the mouse genome. *DNA Res.* **12** (1), 39-51, doi:10.1093/dnares/12.1.39, (2005).

122 Lin, S. *et al.* Cloning, expression, and characterization of a human inosine triphosphate pyrophosphatase encoded by the itpa gene. *J Biol Chem.* **276** (22), 18695-18701, doi:10.1074/jbc.M011084200, (2001).

123 Bjelland, S., Birkeland, N. K., Benneche, T., Volden, G. & Seeberg, E. DNA glycosylase activities for thymine residues oxidized in the methyl group are functions of the AlkA enzyme in Escherichia coli. *J Biol Chem.* **269** (48), 30489-30495 (1994).

124 Cunningham, R. P. DNA glycosylases. *Mutation Research/DNA Repair.* **383** (3), 189-196, doi:10.1016/S0921-8777(97)00008-6, (1997).

125 Zheng, L., Tsai, B. & Gao, N. Structural and mechanistic insights into the DNA glycosylase AAG-mediated base excision in nucleosome. *Cell Discov.* **9** (1), 62, doi:10.1038/s41421-023-00560-0, (2023).

126 Saparbaev, M. & Laval, J. Excision of hypoxanthine from DNA containing dIMP residues by the Escherichia coli, yeast, rat, and human alkylpurine DNA glycosylases. *Proc Natl Acad Sci U S A.* **91** (13), 5873-5877, doi:10.1073/pnas.91.13.5873, (1994).

127 Hedglin, M. & O'Brien, P. J. Human alkyladenine DNA glycosylase employs a processive search for DNA damage. *Biochemistry.* **47** (44), 11434-11445, doi:10.1021/bi801046y, (2008).

128 Gates, F. T. & Linn, S. Endonuclease from Escherichia coli that acts specifically upon duplex DNA damaged by ultraviolet light, osmium tetroxide, acid, or x-rays. *Journal of Biological Chemistry.* **252** (9), 2802-2807, doi:10.1016/S0021-9258(17)40433-9, (1977).

129 Lee, C. C. *et al.* Endonuclease V-mediated deoxyinosine excision repair in vitro. *DNA Repair (Amst).* **9** (10), 1073-1079, doi:10.1016/j.dnarep.2010.07.007, (2010).

130 Lee, C.-C. *et al.* Deoxyinosine repair in nuclear extracts of human cells. *Cell & Bioscience.* **5** (1), 52, doi:10.1186/s13578-015-0044-8, (2015).

131 Yao, M. & Kow, Y. W. Strand-specific cleavage of mismatch-containing DNA by deoxyinosine 3'-endonuclease from Escherichia coli. *J Biol Chem.* **269** (50), 31390-31396 (1994).

132 Gates, F. T., 3rd & Linn, S. Endonuclease V of Escherichia coli. *J Biol Chem.* **252** (5), 1647-1653 (1977).

doi:10.6342/NTU202401466

133 Guo, G., Ding, Y. & Weiss, B. nfi, the gene for endonuclease V in Escherichia coli K-12. *J Bacteriol.* **179** (2), 310-316, doi:10.1128/jb.179.2.310-316.1997, (1997).

134 Yao, M. & Kow, Y. W. Cleavage of insertion/deletion mismatches, flap and pseudo-Y DNA structures by deoxyinosine 3'-endonuclease from Escherichia coli. *J Biol Chem.* **271** (48), 30672-30676, doi:10.1074/jbc.271.48.30672, (1996).

135 Yao, M. & Kow, Y. W. Further characterization of Escherichia coli endonuclease V. Mechanism of recognition for deoxyinosine, deoxyuridine, and base mismatches in DNA. *J Biol Chem.* **272** (49), 30774-30779, doi:10.1074/jbc.272.49.30774, (1997).

136 Yao, M. & Kow, Y. W. Interaction of deoxyinosine 3'-endonuclease from Escherichia coli with DNA containing deoxyinosine. *J Biol Chem.* **270** (48), 28609-28616, doi:10.1074/jbc.270.48.28609, (1995).

137 Shapiro, R. & Pohl, S. H. The reaction of ribonucleosides with nitrous acid. Side products and kinetics. *Biochemistry.* **7** (1), 448-455, doi:10.1021/bi00841a057, (1968).

138 Mi, R., Abole, A. K. & Cao, W. Dissecting endonuclease and exonuclease activities in endonuclease V from Thermotoga maritima. *Nucleic Acids Research.* **39** (2), 536-544, doi:10.1093/nar/gkq791, (2011).

139 Huang, J., Lu, J., Barany, F. & Cao, W. Multiple Cleavage Activities of Endonuclease V from Thermotoga maritima: Recognition and Strand Nicking Mechanism. *Biochemistry.* **40** (30), 8738-8748, doi:10.1021/bi010183h, (2001).

140 Feng, H., Klutz, A. M. & Cao, W. Active site plasticity of endonuclease V from Salmonella typhimurium. *Biochemistry.* **44** (2), 675-683, doi:10.1021/bi048752j, (2005).

141 Fladeby, C. *et al.* The human homolog of Escherichia coli endonuclease V is a nucleolar protein with affinity for branched DNA structures. *PLoS One.* **7** (11), e47466, doi:10.1371/journal.pone.0047466, (2012).

142 Mi, R., Alford-Zappala, M., Kow, Y. W., Cunningham, R. P. & Cao, W. Human endonuclease V as a repair enzyme for DNA deamination. *Mutat Res.* **735** (1-2), 12-18, doi:10.1016/j.mrfmmm.2012.05.003, (2012).

143 Vik, E. S. *et al.* Endonuclease V cleaves at inosines in RNA. *Nat Commun.* **4** 2271, doi:10.1038/ncomms3271, (2013).

doi:10.6342/NTU202401466

144 Kong, X. Y. *et al.* Endonuclease V Regulates Atherosclerosis Through C-C Motif Chemokine Ligand 2-Mediated Monocyte Infiltration. *J Am Heart Assoc.* **10** (14), e020656, doi:10.1161/jaha.120.020656, (2021).

145 Kong, X. Y. *et al.* Deletion of Endonuclease V suppresses chemically induced hepatocellular carcinoma. *Nucleic Acids Res.* **48** (8), 4463-4479, doi:10.1093/nar/gkaa115, (2020).

146 Nawaz, M. S. *et al.* Regulation of Human Endonuclease V Activity and Relocalization to Cytoplasmic Stress Granules*. *Journal of Biological Chemistry.* **291** (41), 21786-21801, doi:10.1074/jbc.M116.730911, (2016).

147 Summer, H., Grämer, R. & Dröge, P. Denaturing urea polyacrylamide gel electrophoresis (Urea PAGE). *J Vis Exp.* (32), doi:10.3791/1485, (2009).

148 Baker, R. P. & Reha-Krantz, L. J. Identification of a transient excision intermediate at the crossroads between DNA polymerase extension and proofreading pathways. *Proc Natl Acad Sci U S A.* **95** (7), 3507-3512, doi:10.1073/pnas.95.7.3507, (1998).

149 Berdis, A. J. Mechanisms of DNA polymerases. *Chem Rev.* **109** (7), 2862-2879, doi:10.1021/cr800530b, (2009).

150 Brammer, K. W., Jones, A. S., Mian, A. M. & Walker, R. T. Study of the use of alkaline degradation of DNA derivatives as a procedure for the determination of nucleotide distribution. *Biochimica et Biophysica Acta (BBA) - Nucleic Acids and Protein Synthesis.* **166** (3), 732-734, doi:10.1016/0005-2787(68)90388-2, (1968).

151 Su, K. Y. *et al.* Pretreatment epidermal growth factor receptor (EGFR) T790M mutation predicts shorter EGFR tyrosine kinase inhibitor response duration in patients with non-small-cell lung cancer. *J Clin Oncol.* **30** (4), 433-440, doi:10.1200/jco.2011.38.3224, (2012).

152 Lindahl, T., Ljungquist, S., Siegert, W., Nyberg, B. & Sperens, B. DNA N-glycosidases: properties of uracil-DNA glycosidase from Escherichia coli. *Journal of Biological Chemistry.* **252** (10), 3286-3294, doi:10.1016/S0021-9258(17)40386-3, (1977).

153 Rashtchian, A., Buchman, G. W., Schuster, D. M. & Berninger, M. S. Uracil DNA glycosylase-mediated cloning of polymerase chain reaction-amplified DNA: application to genomic and cDNA cloning. *Anal Biochem.* **206** (1), 91-97, doi:10.1016/s0003-2697(05)80015-6, (1992).

154 Varshney, U. & van de Sande, J. H. Specificities and kinetics of uracil excision from uracil-containing DNA oligomers by Escherichia coli uracil DNA

doi:10.6342/NTU202401466

glycosylase. *Biochemistry.* **30** (16), 4055-4061, doi:10.1021/bi00230a033, (1991).

155 Delort, A. M. *et al.* Excision of uracil residues in DNA: mechanism of action of Escherichia coli and Micrococcus luteus uracil-DNA glycosylases. *Nucleic Acids Res.* **13** (2), 319-335, doi:10.1093/nar/13.2.319, (1985).

156 Varshney, U. & van de Sande, J. H. Specificities and kinetics of uracil excision from uracil-containing DNA oligomers by Escherichia coli uracil DNA glycosylase. *Biochemistry.* **30** (16), 4055-4061, doi:10.1021/bi00230a033, (1991).

157 Savva, R. & Pearl, L. H. Cloning and expression of the uracil-DNA glycosylase inhibitor (UGI) from bacteriophage PBS-1 and crystallization of a uracil-DNA glycosylase-UGI complex. *Proteins.* **22** (3), 287-289, doi:10.1002/prot.340220310, (1995).

158 Wang, Z. G., Smith, D. G. & Mosbaugh, D. W. Overproduction and characterization of the uracil-DNA glycosylase inhibitor of bacteriophage PBS2. *Gene.* **99** (1), 31-37, doi:10.1016/0378-1119(91)90030-f, (1991).

159 Biolabs, N. E. Unit definition of Endonuclease V (2024).

160 林靜宜. 應用質譜儀進行第五型核酸內切酶與第一型 DNA 聚合酶修復亞黃嘌呤之研究. *碩士論文*. doi:10.6342/NTU202103527, (2021).

161 林哲宇. 第五型核酸內切酶與第一型 DNA 聚合酶在亞黃嘌呤修復中交互作用之研究. *碩士論文*. doi:10.6342/NTU202100313, (2021).

162 Murray, K. K. *et al.* Definitions of terms relating to mass spectrometry (IUPAC Recommendations 2013). *Pure and Applied Chemistry.* **85** (7), 1515-1609, doi:10.1351/PAC-REC-06-04-06, (2013).

163 Oberacher, H., Parson, W., Mühlmann, R. & Huber, C. G. Analysis of Polymerase Chain Reaction Products by On-Line Liquid Chromatography−Mass Spectrometry for Genotyping of Polymorphic Short Tandem Repeat Loci. *Analytical Chemistry.* **73** (21), 5109-5115, doi:10.1021/ac010587f, (2001).

164 徐屾玨. 第一型 DNA 聚合酶校正機制之研究分析. *碩士論文*. doi:10.6342/NTU202303596, (2023).

165 Capson, T. L. *et al.* Kinetic characterization of the polymerase and exonuclease activities of the gene 43 protein of bacteriophage T4. *Biochemistry.* **31** (45), 10984-10994, doi:10.1021/bi00160a007, (1992).

166 Astatke, M., Ng, K., Grindley, N. D. & Joyce, C. M. A single side chain prevents Escherichia coli DNA polymerase I (Klenow fragment) from

doi:10.6342/NTU202401466

incorporating ribonucleotides. *Proc Natl Acad Sci U S A.* **95** (7), 3402-3407, doi:10.1073/pnas.95.7.3402, (1998).

167 Beese, L. S. & Steitz, T. A. Structural basis for the 3'-5' exonuclease activity of Escherichia coli DNA polymerase I: a two metal ion mechanism. *Embo j.* **10** (1), 25-33, doi:10.1002/j.1460-2075.1991.tb07917.x, (1991).

168 Garcia-Diaz, M. & Kunkel, T. A. Mechanism of a genetic glissando: structural biology of indel mutations. *Trends Biochem Sci.* **31** (4), 206-214, doi:10.1016/j.tibs.2006.02.004, (2006).

169 Derbyshire, V., Grindley, N. D. & Joyce, C. M. The 3'-5' exonuclease of DNA polymerase I of Escherichia coli: contribution of each amino acid at the active site to the reaction. *Embo j.* **10** (1), 17-24, doi:10.1002/j.1460-2075.1991.tb07916.x, (1991).

170 Tsai-Wu, J. J., Liu, H. F. & Lu, A. L. Escherichia coli MutY protein has both N-glycosylase and apurinic/apyrimidinic endonuclease activities on A.C and A.G mispairs. *Proc Natl Acad Sci U S A.* **89** (18), 8779-8783, doi:10.1073/pnas.89.18.8779, (1992).

171 Bennett, S. E., Sanderson, R. J. & Mosbaugh, D. W. Processivity of Escherichia coli and Rat Liver Mitochondrial Uracil-DNA Glycosylase Is Affected by NaCl Concentration. *Biochemistry.* **34** (18), 6109-6119, doi:10.1021/bi00018a014, (1995).

172 Alexeeva, M. *et al.* Excision of uracil from DNA by hSMUG1 includes strand incision and processing. *Nucleic Acids Res.* **47** (2), 779-793, doi:10.1093/nar/gky1184, (2019).

173 Brammer, K. W., Jones, A. S., Mian, A. M. & Walker, R. T. Study of the use of alkaline degradation of DNA derivative as a procedure for the determination of nucleotide distribution. *Biochim Biophys Acta.* **166** (3), 732-734, doi:10.1016/0005-2787(68)90388-2, (1968).

174 張倖林. 發展質譜儀進行 DNA 修復蛋白-hSMUG1 之分析. *碩士論文*. doi:10.6342/NTU202002939, (2020).

175 Männistö, R. H., Kivelä, H. M., Paulin, L., Bamford, D. H. & Bamford, J. K. The complete genome sequence of PM2, the first lipid-containing bacterial virus To Be isolated. *Virology.* **262** (2), 355-363, doi:10.1006/viro.1999.9837, (1999).

176 Kowalski, D. & Sanford, J. P. Action of mung bean nuclease on supercoiled PM2 DNA. *J Biol Chem.* **257** (13), 7820-7825 (1982).

doi:10.6342/NTU202401466

177 Lebendiker, M. & Danieli, T. Purification of proteins fused to maltose-binding protein. *Methods Mol Biol.* **681** 281-293, doi:10.1007/978-1-60761-913-0_15, (2011).

178 Öhrmalm, C. *et al.* Hybridization properties of long nucleic acid probes for detection of variable target sequences, and development of a hybridization prediction algorithm. *Nucleic Acids Research.* **38** (21), e195-e195, doi:10.1093/nar/gkq777, (2010).

179 Baumann, T., Arndt, K. M. & Müller, K. M. Directional cloning of DNA fragments using deoxyinosine-containing oligonucleotides and endonuclease V. *BMC Biotechnology.* **13** (1), 81, doi:10.1186/1472-6750-13-81, (2013).

180 Zhang, Z., Jia, Q., Zhou, C. & Xie, W. Crystal structure of E. coli endonuclease V, an essential enzyme for deamination repair. *Sci Rep.* **5** 12754, doi:10.1038/srep12754, (2015).

doi:10.6342/NTU202401466

Biography

The author was born October 4th ,1994 in in Taipei, Taiwan, R.O.C, the single child of Shyurng-rern Chang and Shin-chi Li. She graduated form Lishan High School in Taipei, Taiwan in 2013. She then enrolled at the Taipei Medical University at Taipei, Taiwan. She graduaated with Bachelor of Science degree in Medical Laboratory Science and Biotechnology in 2017. She then enrolled in the Department of Clinical Laboratory Science and Medical Biotechnology at National Taiwan University as a Graduate Fellow. She enrolled in Direct-Entry Doctoral Degree Programs in 2019 and awarded a 3-years-doctoral scholorship at National Taiwan University. Additional financil support for her graduate studies has been provided by Lungyen company and Good Liver Foundation, Taipei, Taiwan.

doi:10.6342/NTU202401466

Publications

1. Proofreading of single nucleotide insertion/deletion replication errors analyzed by MALDI-TOF mass spectrometry assay

Hui-Lan Chang, Kang-Yi Su, Steven D. Goodman, Neng-An Chou, Kuei Ching Lin, Wern-Cherng Cheng, Liang-In Lin, Sui-Yuan Chang, Woei-horng Fang*

DNA Repair(Amst), 2020, Apr :88: 102810. doi: 10.1016/j.dnarep.2020.102810.

2. Measurement of uracil-DNA glycosylase activity by matrix assisted laser desorption/ionization time-of-flight mass spectrometry technique

Hui-Lan Chang, Kang-Yi Su, Steven D. Goodman, Rong-Syuan Yen, Wern-Cherng Cheng, Ya-Chien Yang, Liang-In Lin, Sui-Yuan Chang, Woei-horng Fang*

DNA Repair(Amst), 2021, Jan :97: 103028. doi: 10.1016/j.dnarep.2020.103028.

3. Uracil-DNA Glycosylase Assay by Matrix-assisted Laser Desorption/Ionization Time-of-flight Mass Spectrometry Analysis

Hui-Lan Chang, Kang-Yi Su, Steven D. Goodman, Wern-Cherng Cheng, Liang-In Lin, Sui-Yuan Chang, Woei-horng Fang*

J Vis Exp. 2022 Apr 22:(182):10.3791/63089. doi: 10.3791/63089.

doi:10.6342/NTU202401466