Structure, Function and Potential Application Studies of Selenium Substituted Nucleic Acids

STRUCTURE, FUNCTION AND POTENTIAL APPLICANTION STUDIES OF SELENIUM SUBSTITUTED NUCLEIC ACIDS

by

ZIYUAN FANG

Under the Direction of Zhen Huang, PhD

ABSTRACT

Nucleic Acids are the most important macromolecules in living systems. They directly participate in storing, transferring and expressing the genetic information, thereby controlling and regulating the function of living systems. Therefore, nucleic acid research will give us molecular insights into lives. Meanwhile, many novel research methods based on nucleic acids have been developed. Moreover, nucleic acid-based therapeutics have been developed rapidly over the past several years, leading to a revolution on drug discovery and disease study at the molecular level and the genetic level. Consequently, the structure and function studies on nucleic acids and protein-nucleic acid complexes have attracted tremendous research attention, especially in medical science and biological chemistry.

Since oxygen and selenium are in the same main group in the periodic table, we anticipate oxygen atoms on nucleic acids can be replaced by selenium atoms without significant perturbations on structure and function. Previously, we successfully completed the chemical and enzymatic synthesis of Se-modified nucleic acids (SeNA) with selenium substitutions at various positions. Indeed, we have demonstrated that the Se-modifications would not lead to significant changes in the structure. In addition, SeNA can largely facilitate nucleic acid-protein X-ray crystallography, and its unique properties may lead to many other potential applications, such as gene silencing, therapeutics and cellular imaging. This dissertation has explored potential applications of SeNA, including therapeutic and biochemical research, via structural and functional studies. We demonstrated that SeNA has great potential for antisense drug discovery because of its stronger nuclease resistance, better compatibility with RNase H and more efficient gene silencing in cells. SeNA could also be applied in cellular imaging *via* micro-X-ray fluorescence (microXRF), which allowed us to establish a novel approach to image mRNA transportation in cells. With this approach, we have successfully monitored the GFP mRNA in HeLa cell. Furthermore, we found that SeNA doesn't perturb the structure of DNA/DNA polymerase complex, which allows an accurate structural determination of protein-nucleic acid complexes and provides a bright future on X-ray crystallography of nucleic acids and their protein complexes.

INDEX WORDS: Nucleic acid, Selenium, X-ray crystallography, Gene silencing, Biosensor.

STRUCTURE, FUNCTION AND POTENTIAL APPLICANTION STUDIES OF SELENIUM SUBSTITUTED NUCLEIC ACIDS

by

ZIYUAN FANG

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in the College of Arts and Sciences

Georgia State University

2018

STRUCTURE, FUNCTION AND POTENTIAL APPLICANTION STUDIES OF SELENIUM SUBSTITUTED NUCLEIC ACIDS

by

ZIYUAN FANG

Committee Chair: Zhen Huang

Committee: Markus Germann

Jun Yin

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

December 2018

DEDICATION

This dissertation is dedicated to my parents, Hong Fang and Jingmei Zhang, who had instilled in me the strength of character that navigates me through all challenges. Thank you for your encouragement and support during my 5-year Ph.D. program. Thank you for your education and guidance before my Ph.D. program. Also, thank you for your unconditional love in my whole life.

I further dedicate this dissertation to my lovely girlfriend, Shiying Zhang. Thank you for being my cheerleader. Your patience, encouragement, and love support me to finish this work.

Also, this dissertation is dedicated to my aunt, Lei Fang, who guided me here to Atlanta starting my Ph.D. program. Thank you for your help introducing me here and supporting my life in Atlanta. Without you, I will never have such a wonderful memory of my Ph.D. life.

ACKNOWLEDGEMENTS

The work described here was all under the direction and guidance of my adviser, Dr. Zhen Huang. I would like to express my deepest gratitude to Dr. Huang for offering me this opportunity to work on this attractive project, giving me valuable suggestions, providing the lab with excellent atmosphere and educating me to be a qualified scientist. Both encouragement and criticism from him will benefit me a lot. Without your help, I am not able to complete this dissertation. Thank you for all you have done for me.

I would like to appreciate my dissertation committee members, Dr. Markus Germann and Dr. Jun Yin for your guidance, advice, comments, and questions. Without your help, I am not able to achieve my current position in my research work.

I would also thank my labmates for their support and discussion. I will thank Dr. Wen Zhang for his teaching on structural biology and cellular biology and discussing my projects when I have any difficulties. I will thank Dr. Abdur Rob for his teaching on biochemistry technical. I will also thank Cen Chen, Lin Qin, Fukang Yang, Shuo Liu, and Liuyi Shi for sample synthesis, preparing and purification. I will also show my appreciation to Dr. Julianne Caton-Williams, Dr. Jozef Salon, Dr. Manindar Kaur, Dr. Sibo Jiang, Dr. Huiyan Sun, Chuilun Kong, Yifei Wang, Razieh Esmaeili, Travon Haynes, Edwin Ogbonna, James Campbell, Xinghua Chen, Lingrui Zheng, Babatunde Bello, Mahesh Bhoodai, Huilun Li for their discussions and helps in the lab.

I will also thank the financial, academic and technical support from the Department of Chemistry and Department of Biology at Georgia State University. I would like to thank Dr. Siming Wang for her help on mass analysis, thank Yanyi Chen for his help on the facility of Department of Chemistry and thank Debby Walthall for her help on the facility of Department of Biology.

I thank the beamline staffs at Advanced Light Source (ALS) Beamline 8.2.1 and Beamline 8.2.2 at Lawrence Berkeley National Laboratory. I appreciate Dr. Kevin Royal and Dr. Anthony Rozales for their kindly help on my X-ray diffraction data collection. I will also thank Dr. Stacey Ortega for her help on our beamline time assignment and arrangement.

I would like to thank Dr. Barry Lai from Argonne National Laboratory for his help on micro-X-ray fluorescence determination on our cell sample.

I thank Dr. Wei Yang from NIH for her kindly providing us the RNase H plasmids and offering help on RNase H expression and purification.

I would like to give special thanks to the Rapidata 2017 course at Stanford Synchrotron Radiation Lightsource of SLAC National Accelerator Laboratory. Thank you for teaching me basic knowledge on macromolecular X-ray diffraction data collection, data processing, and structure solution.

Last but not least, I would like to thank the funding and financial support from the National Institutes of Health and the Georgia Cancer Coalition.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

SeNA	Selenium modified nucleic acids
PDB	Protein data bank
MR	Molecular replacement
SIR	Single isomorphism replacement
MIR	Multiple isomorphism replacement
SAD	Single-wavelength anomalous diffraction
MAD	Multi-wavelength anomalous diffraction
DIS	Dimerization initiation site
PSe-DNA	Phosphoroselenoate DNAs
PS	Phosphorothioate
LNA	Locked nucleic acids
PMO	Phosphorodiamidate morpholino oligomer
PAGE	Polyacrylamide gel electrophoresis
T_m	Thermal melting temperature
ALS	Advanced Light Source
PBS	Phosphate saline buffer
mRNA	Massager RNA
XRF	X-ray fluorescence
GFP	Green fluorescence protein
Pol I	DNA Polymerase I
KF	Klenow fragment
Taq	*Thermus aquaticus* DNA polymerase I

Bst	*Bacillus stearothermophilus* DNA polymerase I
PCR	Polymerase chain reaction
CD	Circular dichroism

1 GENERAL INTRODUCTION

1.1 Nucleic Acids for Biofunction Studies and Drug Discovery

Nucleic acids are the storage and carrier of genetic information and directly participate in gene replication, transcription, and expression[1-2]. The control of nucleic acids can lead to the regulation of genetic information flow and gene expression. In addition, many non-coding RNAs, like microRNA and ribosomal RNA, participate in signal transduction directly by regulatory functions[3-4]. Moreover, nucleic acid-based therapeutic strategies are effective tools in drug discovery and disease study at the molecular level as well as the genetic level[5-6]. In these therapeutic strategies, the gene, instead of a protein product, is recognized as drug targets and the modified oligonucleotides (including DNAs and RNAs) are applied as drug candidates. Therefore, nucleic acids related research is essential for better understanding of diseases and identifying novel drug candidate and target for the development of therapeutics with high efficiency to treat some world-wide diseases. Because of the advantages of low toxicity, notable efficacy, and high specificity, numerous nucleic acid-based potential therapeutics are under development or in clinical trials, including nucleoside analog, antisense oligonucleotides, siRNAs, miRNAs, aptamers and ribozymes and DNazymes[7-8]. Instead of the traditional attacking of aberrant proteins, modulating expression of specific diseases' gene by these nucleic acids therapies can lead to better treatment. Research activities in this area require further development of new nucleic acid analogs with improved properties.

On the other hand, the unique structures of DNA and RNA provide nucleic acids multiple functions and properties. The negative charges on phosphate groups make nucleic acids soluble in water, and the hydrophobic nucleobases in the center can form Watson-Crick hydrogen bonds. Additionally, the complicated RNA secondary and tertiary structures generate enzyme-like

catalytic functions. Recently, X-ray crystallography has been the most powerful technical tool to provide researchers detailed 3-D structural information of nucleic acids at the atomic level, which is helpful for us to discover new drug candidate and understand nucleic acids related biological processes mechanisms, such as DNA replication, ribozyme catalysis, DNA transcription, and RNA translation[9-10]. Despite the bottle-necks of nucleic acid crystallography and phase problem, it is greatly promising for researchers to rely on it for nucleic acid function study and gene therapy development[11-12].

1.2 Selenium Substituted Nucleic Acids for Structural Determination

In the past half century, X-ray crystallographic technologies have largely facilitated the structure determination of biological macromolecules, including nucleic acids and protein-nucleic acid complexes. More and more structures of nucleic acids and nucleic acid-protein complexes have been determined. At the same time, there are many potential nucleic acid-based therapeutics in development, and some of them have already been approved by FDA[13]. These investigations make structural and functional studies of nucleic acid-protein complexes an increasingly active research field, thereby demanding new methodologies especially novel strategies for nucleic acid X-ray crystallography, such as the selenium-atom-specific replacement of oxygen atoms in nucleic acids[14].

However, compared with protein crystallography, the process of structure determination of nucleic acids and nucleic acid/protein complexes is much slower. **Figure 1.1** indicates the number of total released structures in protein data bank (PDB) since 2001. It is obvious that the increase of nucleic acids related structures determination is much slower than the structure of protein only. The reason is that there are some unique challenges for nucleic acid X-ray crystallography. The

first challenge is to obtain the crystals of nucleic acids. Proteins have a variety of structural and chemical functionalities on its surface to enable molecular interaction and could easily pack to crystals. However, the surface of nucleic acid has a lot of negative charges and repetitive phosphate functionalities. And there are also many nucleobases which are much alike with each other with aromatic rings. Compared with protein, these intrinsic charges, as well as its structural dynamics, make the crystal packing of DNA and RNA molecules much more difficult[14]. The crystallization condition and buffer screening process are labor intensive, and the formation of high-quality crystals could be hard and time-consuming. These issues have significantly slowed down the structure determination of nucleic acids and protein-nucleic acid complexes.

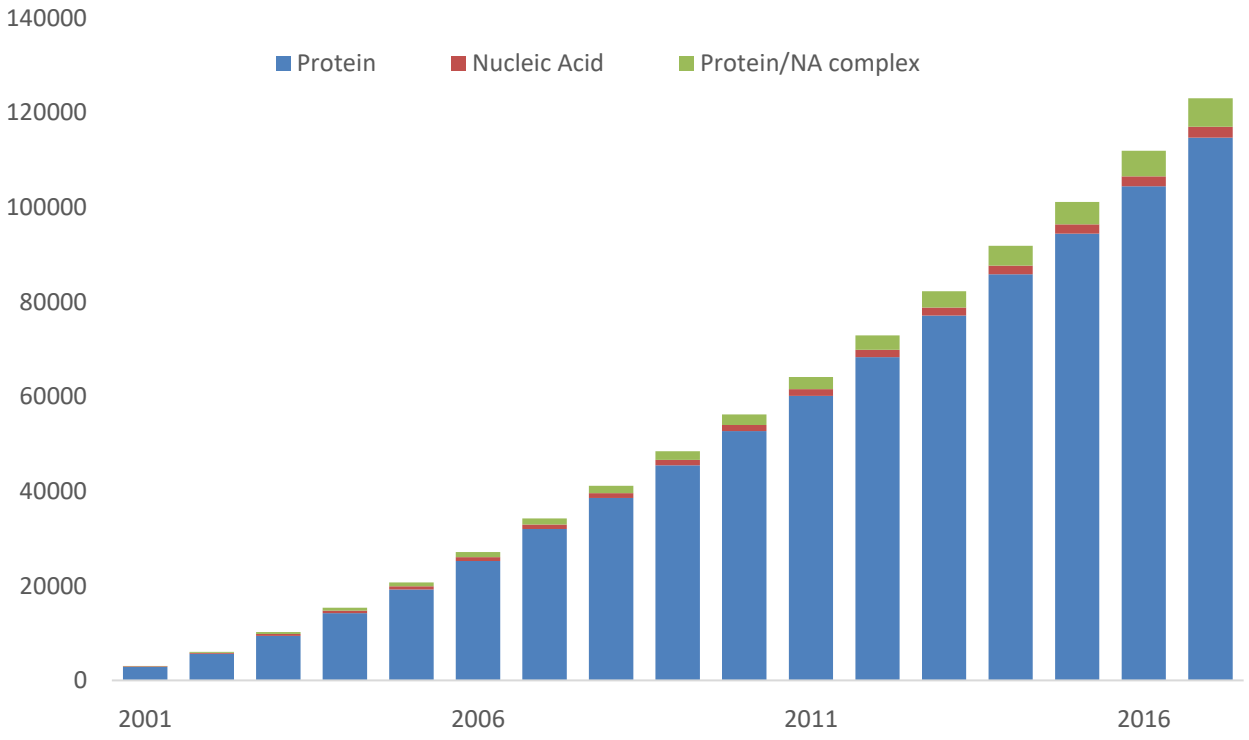


Figure 1.1 Total released structures in PDB each year since 2001

Besides the difficulties of crystal growth, phase problem (phasing or phase determination) is another challenge need to be overcome. During the diffraction data analysis, the application of a Fourier transformation is required to construct the electron density map of the biological macromolecule investigated. In the Fourier transformation, both structure factors amplitude F and phase of the reflection α are needed to calculate the electron density. The former can be measured during the X-ray diffraction experiment. However, the phase information is missing which is critical to producing the correct crystal structure. Several different strategies were applied to address this issue, such as molecular replacement (MR), isomorphism replacement (SIR; MIR) and anomalous dispersion techniques including single- and multi-wavelength anomalous diffraction (SAD and MAD). With an available similar model, MR should be the most convenient and fastest method which don't need to obtain a heavy-atom derivative. However, for novel structures, especially for nucleic acids containing highly repetitive and stacked units[15], molecular replacement could be very difficult. A suitable model can hardly be found. So heavy-atom derivatization method is the only way to solve the phase problem for most nucleic acid molecules. The isomorphous replacement was successfully used in protein crystallography by soaking and co-crystallization with heavy metal cations. However, this strategy can often cause random hydrolysis of the phosphate backbone in nucleic acid crystallization because normally heavy metal cations are Lewis acids. Moreover, this method requires preparing many isomorphous native crystals and heavy atom derivatives to determine the structures. Compared with isomorphous replacement, the SAD and MAD technique could be a better way to solve the phase problem of nucleic acids and nucleic acid-protein complexes, which only requires one single crystal in order to obtain all diffraction information needed for the structure determination.

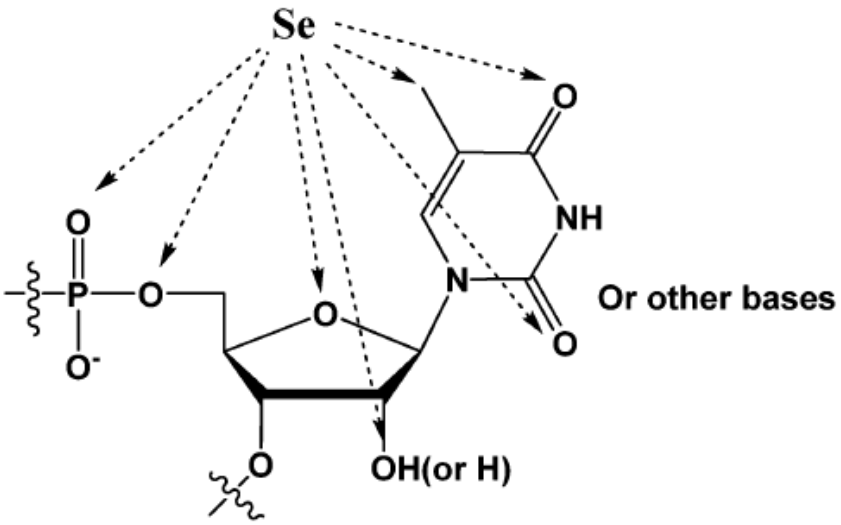


Figure 1.2 Se-atom-specific modifications on nucleic acids

To solve these problems in nucleic acid X-ray crystallography, our research group has pioneered and developed the selenium substituted nucleic acids for structure and function studies (**Figure 1.2**), establishing chemical structural biology of nucleic acids. Based on our previous research, Se-atom-specific modifications will not cause significant structure and function perturbations and could tailor the local structure and molecular dynamics of DNA and RNA and to accelerate the crystallization process of these molecules. During the past several years, the selenium atom, in the same main group of oxygen, has been introduced to many different locations of nucleotides, including phosphate groups, sugar moiety, and nucleobases, to study its unique structural and functional effects on nucleic acids.

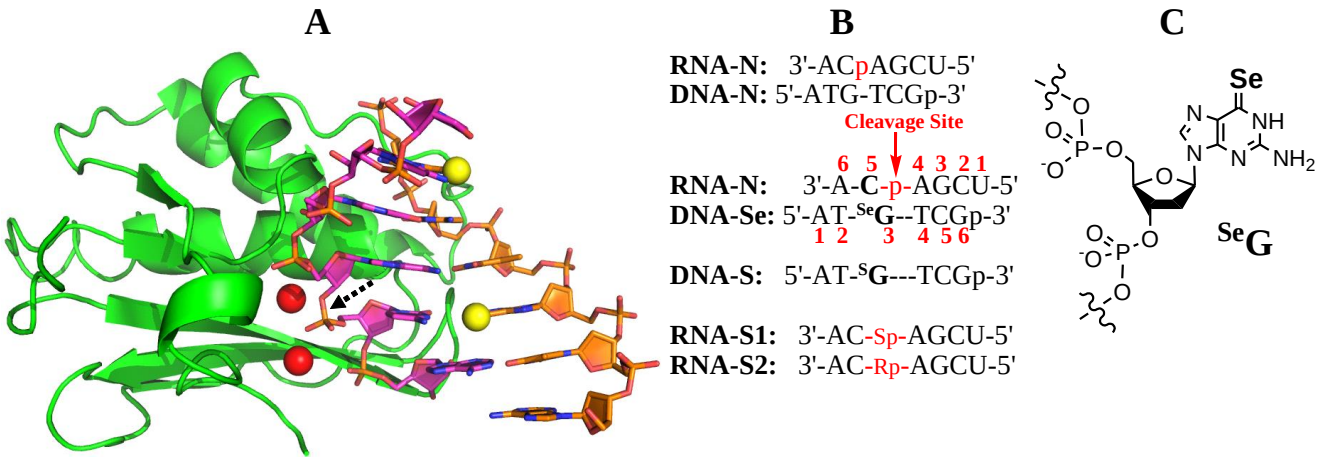


Figure 1.3 Crystal structure of RNase H complexed with the Se-modified DNA and RNA duplex.

(A) The overall structure of Se-DNA/RNA/RNase H ternary complex determined at 1.80 Å resolution (D132N mutant; PDB ID: 3TWH). Protein is in green; RNA and DNA strands are in pink and orange, respectively. The red and yellow spheres are Mg^{2+} ions and Se atoms, respectively. The cleavage site is indicated by the arrow.

(B) Sequences of the native and modified DNAs and RNAs. SeG and SG represent 6-Se-G and 6-S-G, respectively.

(C) The Se-DNA derivatized with the 6-Se-deoxyguanosine.

We have found that selenium modification could help facilitate crystal growth significantly, as well as rationally solving phase problem in nucleic acid X-ray crystallography, where crystallization and phase determination are the two major challenges as we discussed before. Our previous studies on selenium modified nucleic acid X-ray crystallography illustrate that the selenium modification facilitates the crystal growth generally[16-23], while there were no significant structure perturbations determined[24-25]. It also offers a novel method for the structure determination of the protein-nucleic acid complex, which helped to study the relationship between structure and function of enzymes such as RNase-H (**Figure 1.3**)[26-27]. What's more, our previous study has revealed that the 2'-methylseleno modification can facilitate the crystallization of A-form DNA by destabilizing the B-form helix. The reason is the large methyl selenyl group cannot tolerate a B-helix geometry due to the steric hindrance with neighboring residues, but it fits well in the minor

groove of the A-form helix, which helps the transformation from B-form to A-form and further facilitates the crystallization (**Figure 1.4**)[28].

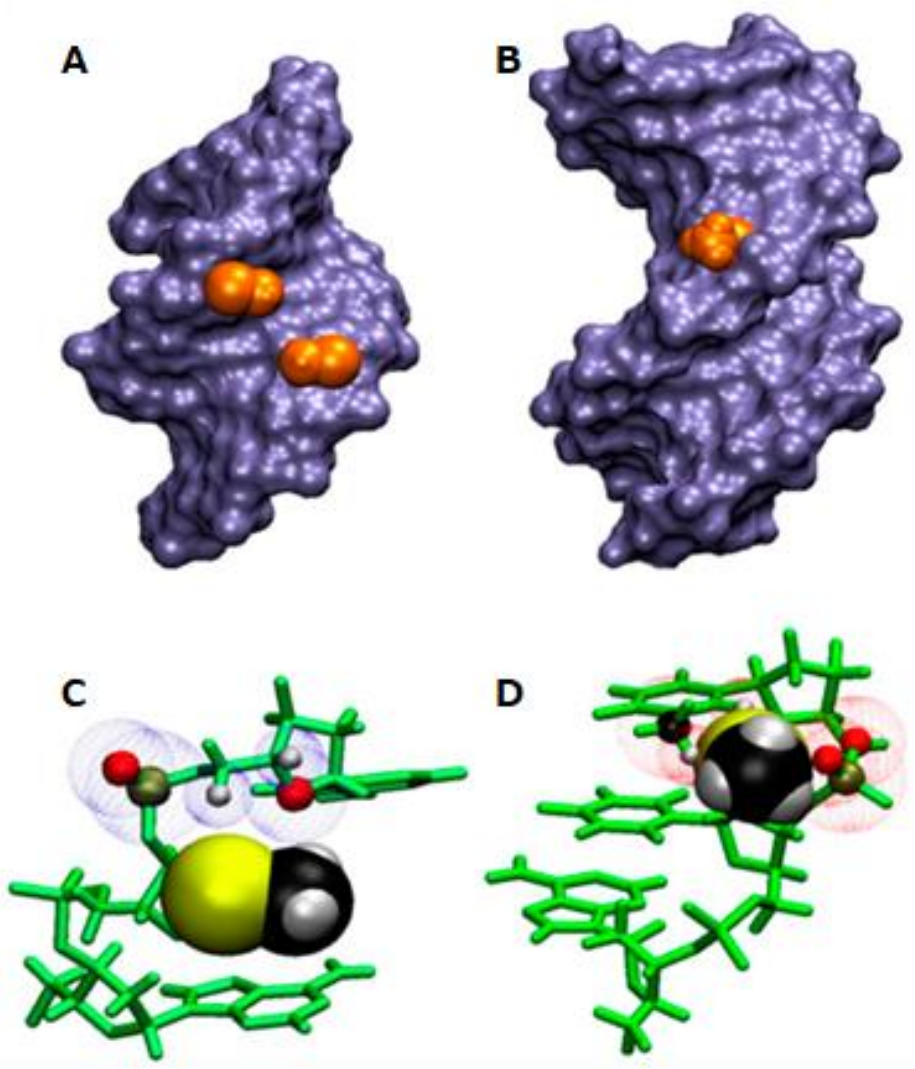


Figure 1.4 The 2'-$SeCH_3$ group implanted in A and B helical structures

1.3 Functional Study of Selenium Substitution on Nucleic Acids

The study of Se-derivatized nucleic acids was started since 2001 and many different kinds of Se-modifications were introduced to DNA and RNA for structural and functional studies. Currently, selenium can almost be introduced at any position in nucleic acids replacing the oxygen atoms or inserting to different places on nucleotides, including sugar moiety, phosphate groups,

and nucleobases. Some examples of each kind of modifications will be introduced to discuss the influences and functions of selenium substitution on nucleic acids.

1.3.1 Se-modifications on sugar moiety

Among all the sites in nucleic acids for selenium derivatization, ribose 2'-position is the most widely explored. This Se-modification could facilitate crystal growth[16-19, 21-23]. Through collaboration among Huang, Egli and other co-workers, first crystal structure of nucleic acids was solved using direct selenium derivatization and MAD phasing (**Figure 1.5A**)[25]. After that, Micura and co-workers have developed the 2'-Se-RNA synthesis and made it possible to synthesize a Se-RNA with up to 100 nucleotides *via* chemical synthesis and T4 RNA ligase ligation[20, 29-30]. As a successful application of 2'-Se-derivatized nucleic acid, a ligated 49 nucleotides ribozyme with 6 selenium atoms modification was crystallized, and the structure was solved in approximately 3 Å resolution (**Figure 1.5B**) using MAD phasing[31]. This work provided new insights into the ribozyme catalytic pocket, substrate recognition, and catalytic mechanism of stereoselective carbon-carbon bond formation. Moreover, Ennifar and co-workers crystallized and solved some structures of complexes of HIV-1 genomic RNA dimerization initiation site (DIS) and antibiotics with 2'-Se-derivatization technique at 1.5 Å to 2.0 Å resolutions (**Figure 1.5C**)[32]. The great quality of the electron density maps of these structures revealed many details of drug–RNA interactions and offered important information for HIV-1 drug discovery.

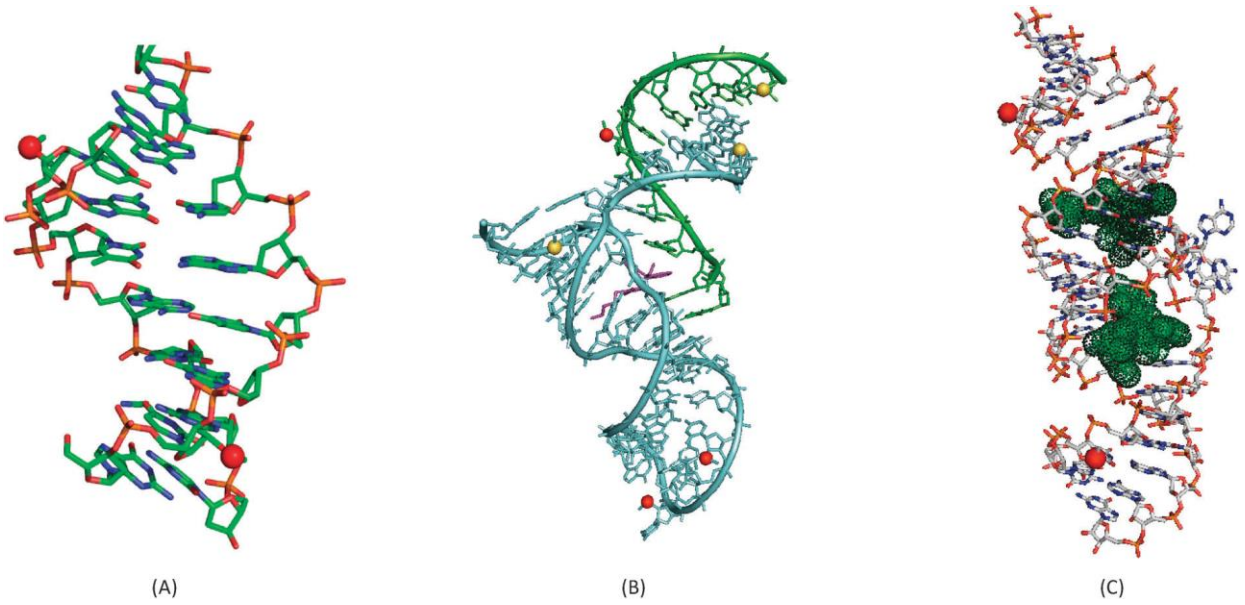


Figure 1.5 The DNA and RNA structures with 2'-Se-modification.
(A) 2'-Se-dU DNA octamer [5'-G(Se-dU)GTACAC-3']$_2$, PDB ID: 1Z7I. The red spheres represent the selenium atoms.
(B) Structure of Diels–Alder ribozyme derivatized with the 2'-Se-functionality (PDB ID:1YLS). The red and yellow spheres represent the selenium atoms of 2'-Se-U (red) and 2'-Se-C (yellow). The compound in purple is the reaction product.
(C) Structure of neomycin bound to HIV-1 genomic RNA dimerization initiation site (DIS) duplex with 2'-methylseleno derivatization. The red balls represent the selenium atoms. Neomycin molecules are in green dots (PDB ID: 3C7R).

When compared with the native structure, both the structures of 2'-Se-DNA and 2'-Se-RNA have well proven that this site-specific 2'-Se-modification does not significantly alter the global and local structures of nucleic acids. The detailed structural and hydration studies in A-form DNAs indicated that the selenium derivatization strategy has better structural isomorphism than the conventional bromine derivatization. Another exciting fact is that this kind of selenium modification at the 2'-position can significantly facilitate the crystal growth and broaden the crystallization buffer conditions for nucleic acids[16-19, 21-23]. The best explanation for this phenomenon is that 2'-selenium modifications could lock the sugar pucker and result in better molecular packing.

Besides 2'-Se-modifications, other modifications, including 4'-Se-modifications[33] and 5'-Se-modification[34], were also reported. Each of these modifications has their specific conformation and potential applications.

1.3.2 Selenium replacement of the phosphate non-bridging oxygen

Egli and co-workers synthesized DNAs with selenium at the phosphate non-bridging position for the crystal structure study[24]. DNA phosphoroselenoates (PSe-DNA) were stable for months under the crystallization conditions. As expected, Se-modified structures are virtually identical with native (**Figure 1.6A**). In addition, the Mg^{2+} ion coordination is all the same in the Se-modified and native DNA structure. It is also worthy to mention that the PSe-DNA synthesized from solid-phase synthesis was a mixture of R and S diastereomers owing to the presence of the phosphorous chiral center. These diastereomers need to be separated by anion exchange chromatography[24]. In order to obtain diastereomerically pure PSe-oligonucleotide without the chromatography separation, our group developed enzymatic methods to synthesize the PSe-DNAs and PSe-RNAs by the enzymatic polymerization using the diastereomerically pure nucleoside 5'-(α-P-seleno)-triphosphates[35-37]. Klenow DNA polymerase is able to recognize both dNTPαSe diastereomers[35], while T7 RNA polymerase specifically recognizes only one NTPαSe diastereomer out of each NTPαSe diastereomer pair[36]. This Se-derivatization strategy on the backbone has been applied to solve some structures that could be extremely difficult for native sequences. The structure of a homo-DNA [$d(CGAATTCG)_2$] (**Figure 1.6B and C**)[38] could be a good example of PSe-DNAs. From this crystal structure, more detailed information about the homo-DNA, especially the role of the backbone-base inclination and the inter-strand stacking in pairing selectivity, was revealed.

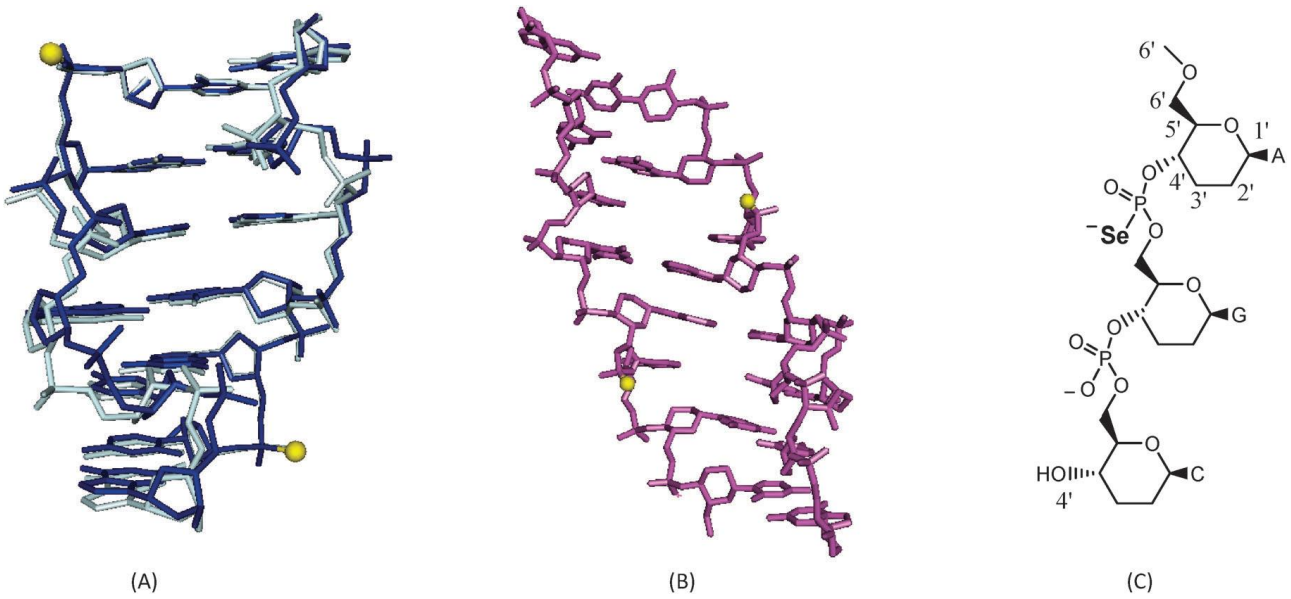


Figure 1.6 Structures of the phosphoroselenoate DNAs (PSe-DNA).
(A) Superimposition of the PSe-DNA (in blue; PDB: 1VRO) and the native DNA with the same sequence [d(CGCGCG)$_2$, in cyan; PDB: 2DCG]. The selenium atoms are shown as the yellow spheres in the PSe-DNA.
(B) Crystal structure (PDB ID: 2H9S) of a homo-DNA [d(CGAATTCG)$_2$]. Yellow spheres indicate the locations of the selenium atoms in the Home-DNA used for MAD phasing.
(C) Chemical structure of the homo-DNA fragment.

1.3.3 Selenium modification on the nucleobases

Our group also led the investigation into the atom-specific modifications with selenium on the nucleobases. The previous research provides novel insights into the base pairing and stacking interactions of nucleic acids. Many different building block phosphoramidites of the 4-Se-T[22], 6-Se-dG[39], 2-Se-T[16] and 5-SeMe-T[18] nucleosides for solid-phase synthesis and the 4-Se-T triphosphate for enzymatic synthesis[40] have been developed and synthesized successfully.

The 4-Se-T was introduced into a self-complementary octamer DNA [5'-G(2'-Se-dU)G(4-Se-T)ACAC-3', **Figure 1.7A**] for the structure study[22]. The crystal structure study indicated that this 4-Se-T modification can shift to accommodate a larger atom, such as selenium, in order to keep the excellent isomorphism in the overall structure with maintaining the similar base-pairing

and duplex stability. These structural data are consistent with the UV-melting study. The selenium-mediated hydrogen bond was observed in this structure with a bond length of 3.35 Å (Se···H–N), which is the first one demonstrated in a biological system. In addition, the 4-Se-T 5'-triphosphate was synthesized and effectively incorporated into DNA oligonucleotides with the catalysis of Klenow fragment of DNA polymerase I[40].

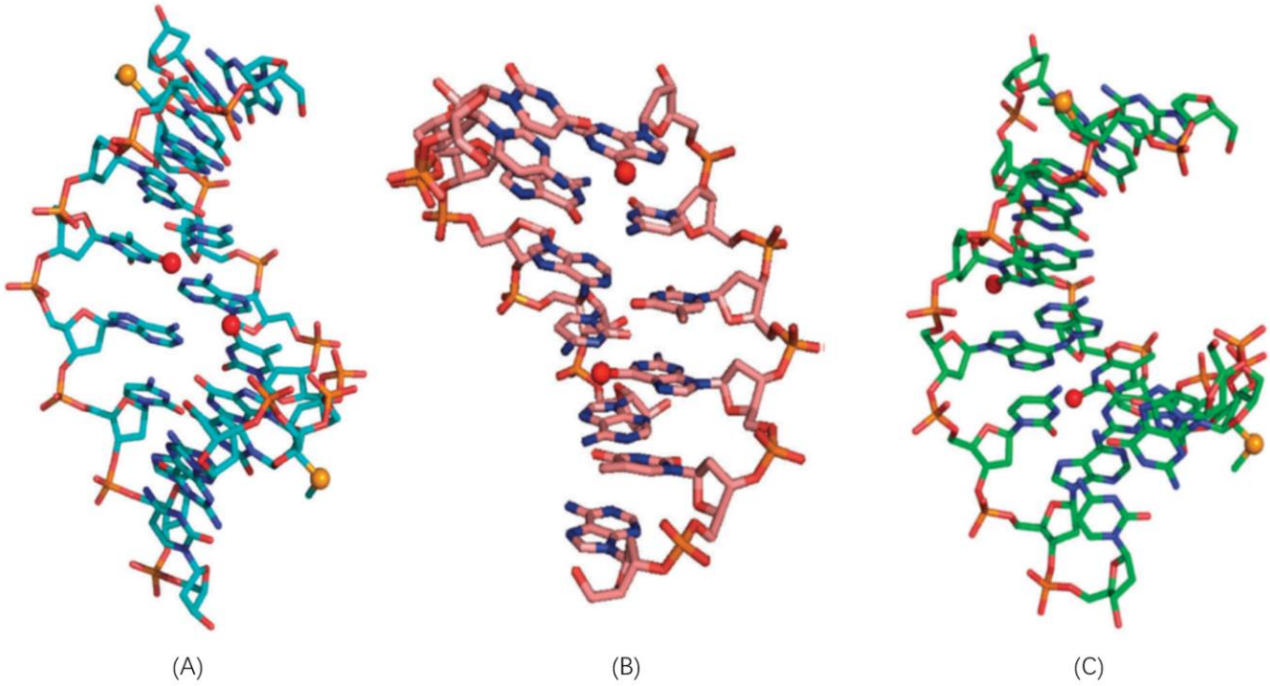


Figure 1.7 Crystal structures of the Se-nucleobase-modified oligonucleotides.
(A) The 4-Se-T-modified DNA octamer [5'-G(2'-Se-dU)G(4-Se-T)-ACAC-3']$_2$, PDB ID: 2NSK. (B) The 2-Se-T-modified DNA octamer [5'-G(2'-Se-dU)G(2-Se-T)ACAC-3']$_2$, PDB ID: 3HGD. The red balls represent the selenium atoms on the nucleobases, and the yellow balls represent the 2'-Se atoms. The 2'-Se-functionality was incorporated in the sequences to facilitate crystallization. (C) The 6-Se-G-DNA/RNA hybrid duplex (5'-AT(6-SeG)TC(6-SeG)-p-3'/5'-UCGACA-3'), PDB ID: 2R7Y.

The oxygen at the 2-position of thymidine has also been successfully replaced by selenium in our lab (**Figure 1.7B**)[16]. Besides the application in crystallography, this 2-Se-T residue has also been tested to probe base pair specificity since the 2-position of T is involved in the T-G wobble pair. Watson–Crick base pairs are the contributors to the sequence-dependent recognition of

nucleic acids, genetic information storage, as well as the high fidelity of DNA polymerase replication and RNA polymerase transcription. However, the wobble base pairing, where T (or U) pairs with G instead of A, is commonly observed in duplexes, especially RNA duplexes, which reduces base-pairing recognition specificity and decreases the high fidelity of enzymatic polymerization. The T-G wobble pairing (or U-G) is caused by the shift of hydrogen bonding between the 2-exo-oxygen on T (or U) and 1-nitrogen on G in the base pairs, while the 2-exo-oxygen atom is not involved in T-A (or U-A) base-pairing. The 2-position is critical for the discrimination of T-A vs. T-G pairs (or U-A vs. U-G). Therefore, we hypothesized that the introduction of a selenium atom at the 2-position can largely increase the electronic and steric effects, resulting in strong discrimination between normal T-A pairing and T-G wobble pairing. To investigate nucleic acid base-pairing fidelity and stacking through the atom-specific mutagenesis and crystallography, we synthesized 2-Se-thymidine phosphoramidite and incorporated it into DNAs via solid phase synthesis. The biophysical results supported our hypothesis and indicated that the 2-Se-T-substitution largely increased the T-A base pair fidelity by discouraging the T-G and T-C mispairs. The crystallographic study further revealed that this selenium-atomic substitution did not significantly change the native T-A base pairing and the overall structure. Thus, this atom-specific selenium substitution of the 2-oxygen of thymidine provides a unique chemical strategy to enhance the base pairing specificity.

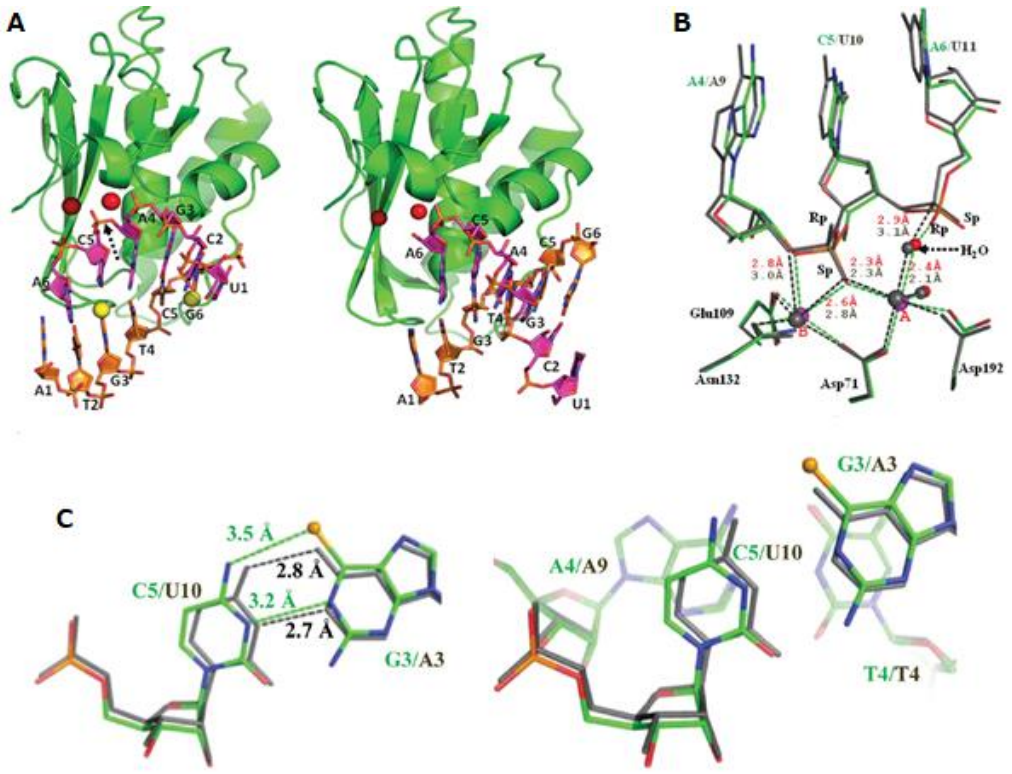


Figure 1.8 Structure of RNase H/Se-DNA/RNA complex with 6-Se-G modification.
(A) Left: The Se-DNA/RNA-RNase H structure (PDB ID: 3TWH); Right: The native DNA/RNA-RNase H structure (PDB ID: 2G8U);
(B) Superimposed Native (in gray) and Se-modified (in color) complex cleavage-site interactions;
(C) Local subtle unwinding of the Se-modified duplex.

Similarly, the 6-Se-dG was incorporated into DNAs (**Figure 1.7C**), and the crystal structure study of the DNA [5'-AT(6-Se-dG)TC(6-Se-dG)-p-3'] was carried out using a ternary complex of the 6-Se-dG-DNA/RNA/RNase H[27]. The structure was reported with the resolution of 1.80 Å (PDB ID: 3TWH). A subtle conformation change was observed based on this high-resolution structure. Comparing to the corresponding native complex with the same sequence (PDB ID: 2G8U), RNase H recognizes the internal position, instead of binding to the junction of the Se-modified DNA/RNA duplex. This result leads to the formation of the substrate-RNase H complex (**Figure 1.8A**). This Se-modified DNA/RNA/RNase H complex was also compared with another native RNase H complex (PDB ID: 1ZBI) which containing a longer substrate sequence.

The Se-dG3/rC5 base pair shows a conformation change by 0.5-0.7 Å which unwinding the duplex and slightly shift the scissile phosphate closer to the RNase H active site by approximately 0.3 Å (**Figure 1.8C**). What's more, the scissile phosphate also positions the water nucleophile in the structure by forming a hydrogen bond (3.10 Å) (**Figure 1.8B**). Further RNA hydrolysis study shows that, in the presence of Se-DNA guide, the rate of hydrolysis is about 62-fold faster, which suggests that this subtle conformation change introduced by Se-modification significantly accelerating the RNA cleavage.

1.4 Potential in clinical therapeutic and drug development

Compared to protein, nucleic acid attracted much less attention as drug targets in the structure-based drug design, partially because of the lack of structural information of nucleic acids with drug candidates. By facilitating the crystallization and improving crystal quality, Se-modified nucleic acids can greatly promote the structure-based drug design of nucleic acids. Those detailed 3D structures can help the researcher understand the structure of drug target or even the interactions between the targets and small molecule ligands.

Moreover, oligonucleotides itself has been studied for over 30 years as potential therapeutics. The major oligonucleotides therapeutics containing antisense oligonucleotides, aptamers, ribozymes, and siRNAs, were studied extensively and lots of progress has been achieved. However, until 2017, there are only six drugs approved by FDA[13]. Many obstacles slow down the development of DNA/RNA-based therapies including the vulnerability to nucleases, off-target effects, poor delivery and low affinity. Lots of chemical modifications of nucleic acids were developed to conquer those challenges such as phosphorothioate (PS), 2'-O-methyl (2'-OMe), locked nucleic acids (LNAs), phosphorodiamidate morpholino oligomer (PMO) and so on[41], which

can increase affinity, nuclease resistance, delivery, and reduce off-target effects by altering the structure and charge of the oligonucleotides.

The previous study on Se-modified nucleic acid shows that replacement of the non-bridging oxygen of DNA and RNA backbone phosphates protect the oligonucleotide from degradation by nucleases[35, 37]. Thus, the Se-modified siRNA may have improved RNase resistance, suggesting a potential therapeutic application on RNAi. Moreover, certain modifications, such as 2'-MeSe, reduce multiple conformations of nucleic acid by destabilizing other unfavorable structure[28] which can potentially increase the affinity and specificity of DNA/RNA duplex by reducing the entropic penalty. The consequent longer half-life allowing effect delivery to tissues of interest. Furthermore, the anticancer activity of selenium-modified nucleotides was also reported[42].

Overall, Se-derivatized nucleic acids (SeNA) possess many unique properties and show great potential in clinical therapeutic as well as drug development. However, the studies on SeNA are still limited and the synthesis is relatively costly, which may prevent broader applications. We expect that further progress and applications of SeNA will be achieved in the future.

2 SELECTIVE GENE SILENCING WITH SE-MODIFIED ANTISENSE DNA

2.1 Introduction

2.1.1 Challenges in antisense therapy

In 1978, Zamecnik and Stephenson reported that an oligonucleotide complementary to the Rous sarcoma virus could block viral replication in chicken fibroblasts.[43] This discovery opened a new world of antisense oligonucleotides. An antisense DNA is a single-strand DNA that is 17-22 nucleotides in length and complementary to a selected gene's mRNA and thereby specifically inhibit expression of that gene.

There are two different strategies of antisense therapeutic, arresting and RNase H cleavage (**Figure 2.1**). Modified antisense DNA could bind to target mRNA through Watson-Crick base pairs. Formation of the RNA/DNA duplex may exert a direct steric effect, blocking the binding of factors required for the initiation of translation or impeding translocation of the ribosome along the mRNA resulting in chain-termination. In this way, the translation was arrested, and the target gene could be downregulated or even silenced. Alternatively, the mRNA may be cleaved by RNase H at the site of the RNA/DNA hybrid. RNase H is a kind of non-sequence-specific endonuclease enzyme that catalyzes the cleavage of RNA on RNA/DNA hybrid. When mRNA was cleaved, antisense DNA will be released and is able to bind to another mRNA and introduce the RNase H cleavage on that mRNA as well. It is obvious that the better strategy is the one involved RNase H because less antisense DNA is needed to have the same effect.

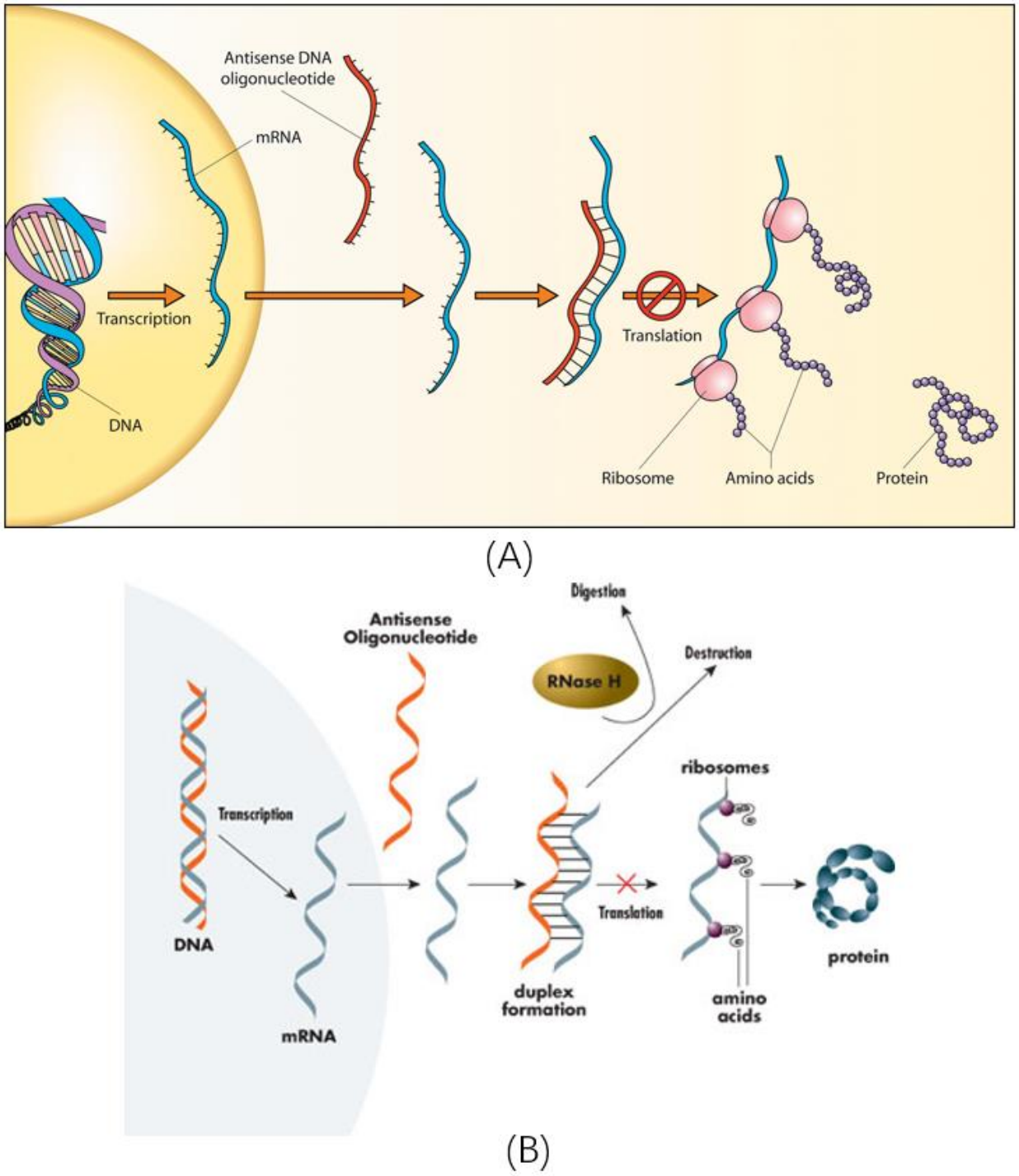


Figure 2.1 Two strategies of antisense therapeutic.
(A) Arresting strategy[44]; (B) RNase H cleavage strategy.

In general, there are several major challenges in DNA antisense strategy: RNase H compatibility, thermo-stability, nuclease resistance, and delivery. Due to its natural compatibility

with RNase H, unmodified native DNA could efficiently assist the enzymatic RNA cleavage. However, it was quickly discovered that unmodified DNA is unstable against nucleases. To enhance the nuclease stability and duplex affinity, the enormous number of DNA chemical modifications, mainly on the backbone and sugar, have been synthesized and applied in antisense strategy, mainly via the mRNA-arresting mechanism. However, heavy modifications may solve the problem of nuclease instability, but they can create some other problems like losing target and RNase H incompatibility. The losing target problem is usually caused by a weaker ability to form stable Watson-Crick base pairs. And the RNase H incompatibility is probably caused by disrupting the natural dynamics of the sugar pucker and backbone and interrupting the necessary interactions. Though gene silencing without RNase H activity can still be achieved via the mRNA-arresting mechanism, the catalytic power of RNase H in RNA inactivation is not harnessed. This is a significant loss in gene silencing efficiency.

2.1.2 Current modification of antisense DNA

Currently, the chemical modification is one of the effective strategies to optimize antisense therapies. In the past decades, numerous efforts have been devoted to achieving chemically modified nucleic acid drug molecules, which improve the nucleic acids properties tremendously *in vivo*. Since the first oligonucleotide drug, Fomivirsen, got approved by FDA in 1998[45], there are several oligonucleotide drugs have been approved by FDA[13]. And early this year 2018 in August, Patisiran became the first FDA approved siRNA drug[46]. All of these oligonucleotide drugs have one or several different kinds of modifications to stabilize them and avoid hydrolysis by nucleases. These modifications involve all three types of internucleoside linkage or backbone modifications, sugar modifications and nucleobase modifications[41].

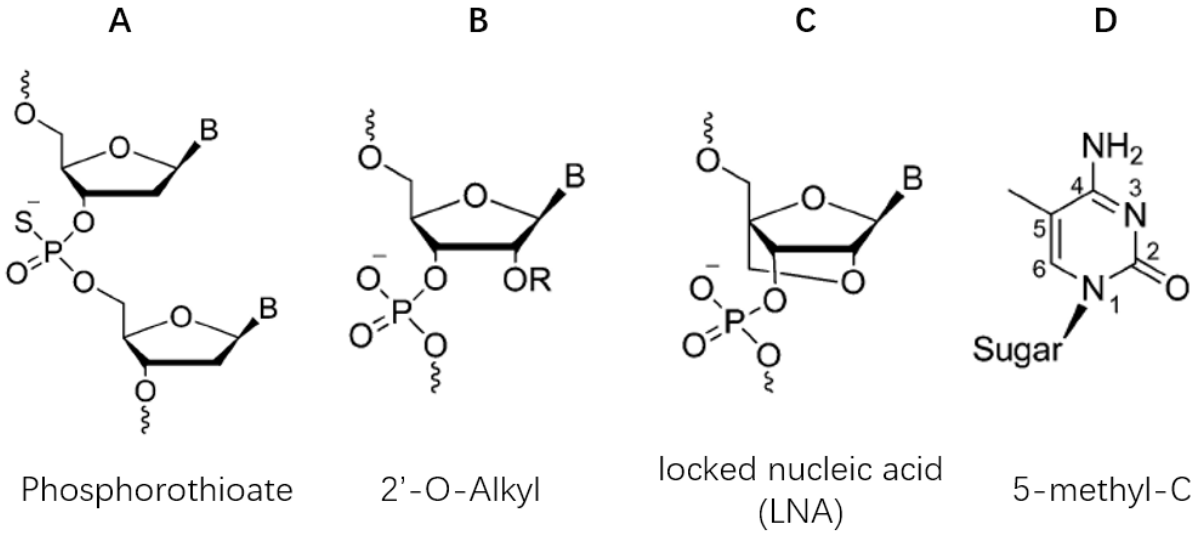


Figure 2.2 Some modifications that exist in FDA approved antisense drugs.
(A) Phosphorothioate; (B) 2'-O-Alkyl; (C) LNA; (D) 5-methyl-C.

2.1.2.1 Internucleoside linkage or backbone modifications

These kind modifications are also referred to as the first generation of chemically modified antisense agents, including phosphorothioate, methylphosphonate, N-3'-phosphoramidate and many other modifications.

Among them, phosphorothioate oligonucleotides are the most successful in gene silencing (**Figure 2.2A**). Among the FDA approved drugs, Fomivirsen, Mipomersen, and Nusinersen involve this kind of modification. Introducing phosphorothioate linkage to oligonucleotides could significantly increase the resistance to nuclease degradation. In addition, phosphorothioate oligonucleotides could also permit RNase H, form regular Watson-Crick base pairs, carry negative charges for cell delivery as compared to natural phosphodiesters[47-49]. However, their binding affinity to target mRNA sequences and specificity are less satisfactory[50]. Despite the disadvantages, phosphorothioate oligonucleotides are the most widely used and successful modification for oligonucleotide drugs.

2.1.2.2 Sugar modifications

In recent years, there are great progress in the synthesis of conformationally constrained nucleoside analogues by modifying the sugar moiety in various ways. These modifications include synthesis of nucleoside analogs containing an electronegative atom or substituent at the 2'-position of sugar[51-52], synthesis of bicyclic nucleoside analogs having an extra ring fused to the sugar moiety[53-58] and some other methods. 2'-O-Methyl (2'-OMe) (**Figure 2.2B**) is a good example for the former while locked nucleic acid (LNA) (**Figure 2.2C**) is a widely used example for the latter. These secondary generation oligonucleotides could solve some problem of poor binding affinity to the target RNA, lack of specificity and low cellular uptake that phosphorothioate could have.

The major structural difference between DNA and RNA is RNA has a 2'-OH group on sugar ring, which leads to C3'-endo conformation. The RNA binding affinity of antisense oligonucleotides could be improved by mimicking RNA structures with 2'-modified nucleosides. Electronegative substituents like fluorine and oxygen influence the furanose sugar C3'-endo conformation[59] and make it easier to form A-type duplexes with target RNA strand[52]. Various reported 2'-substitutions have shown excellent results in antisense therapeutics as they provide high metabolic stability and high affinity to target the complementary mRNA, such as 2'-OMe and LNA. Mipomersen and Nusinersen are good examples of FDA approved drugs which contain this kind of modifications.

2.1.2.3 Nucleobase modifications

Normally most scientists modify the nucleobases to increase the binding affinity of oligonucleotides instead of nuclease resistance because nucleobases provide the prime recognition

site for Watson–Crick base pairs[60]. Thus, nucleobase modifications are less popular than those backbone and sugar modifications. However, some kinds of modifications also have great applications in antisense oligonucleotide drugs. The most attractive sites for substitution of the nucleoside bases are those positions in the major grooves because modifications on these positions neither interfere with base pairing nor induce steric hindrance and influence the general geometry of the double helix[61-63].

The stacking interactions between the planar heterocycles of nucleic acids are largely responsible for the stability of DNA and RNA duplexes. Maximizing stacking interactions through chemical modification provides a means of creating duplex helices of greater stability. However, so far there are no FDA drugs with modifications in this way, despite the fact that Mipomersen and Nusinersen have 5-methyl-C modifications (**Figure 2.2D**).

2.1.3 The advancement of selenium modifications for antisense drugs

As we discussed before, even though a lot of various modification methods have been developed, none of them can solve all the problems. The basic conflict is that atom specific modifications may not be able to confer sufficient resistance to nucleases or decrease the specificity and binding affinity, while heavy modifications could lead to loss of RNase H compatibility.

However, we think selenium atom specific modification could be a solution that can solve all those problems. First of all, selenium is a kind of essential micronutrients in the human body, so it has low toxicity. And We spent many years on Se-modified nucleic acids research and made a lot of progress in the past 15 years. We found that selenium substitution of the non-bridging oxygen of DNA and RNA backbone phosphates protect the oligonucleotide from degradation by

nucleases[35, 37]. On the other hand, the 2-Se-T modification could significantly increase the specificity and binding affinity of Watson-Crick base pairs and avoid wobble pairs[64-65]. What's more, the most exciting result is that we found 6-Se-dG modified oligonucleotides could not only maintain but even increase the RNase H activity[27].

Although these modifications are at different positions, we believe Se-modified nucleic acids have the potential to become a novel successful antisense oligonucleotide drug. We would like to further study whether there is a possibility to find a single atom specific selenium substituted modification that could solve all the problems that we mentioned before. In this way, we synthesized 5-methylselenium-thymidine (**Figure 2.3**). 5-methylselenium-thymidine contains a selenium atom at 5 position on the pyrimidine ring on thymidine. The 5 position does not involve in base pair formation, so we predict this modification will not affect the ability forming stable base pairs with adenosine on target mRNA. And the modification is in the major groove while when RNase H is binding to the minor groove of DNA/RNA duplex. Therefore, we believe this modification will not lose the compatibility with RNase H. However, we believe base modifications can also stabilize the antisense from nucleases. To prove our hypothesis, we insert it to oligonucleotides to investigate how it performs as antisense oligonucleotide and whether it has therapeutic potential as an antisense drug.

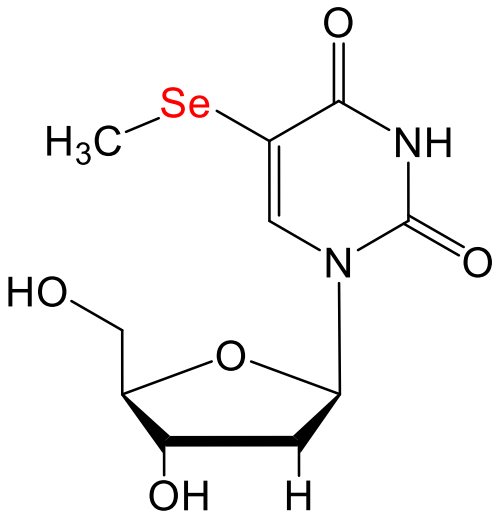


Figure 2.3 Structure of 5-methyselenium-thymidine.

2.2 Material and Methods

2.2.1 Se-modified antisense DNA resistance to serum

Kinase reaction for DNA substrate: Set up 20 μL reaction with DNA substrate (2 μL, 50 μM), T4 polynucleotide kinase (1 μL, Biolab), [γ-^{32}P]-ATP (1 μL, PerkinElmer) and 10× kinase reaction buffer (2 μL, Biolab) at 37 °C for 1 h. Stop the reaction by adding high concentration NaCl (1 μL, 3M). Add ethanol (11 μL) and keep at -20 °C for 15 min. Centrifuging at 14,800 rpm for 20 min. Remove the supernatant carefully. Let it dry and redissolve the labeled DNA in water (10 μL).

Digestion of labeled DNA substrate by Serum: Set up 5 μL reaction with labeled DNA substrate (1 μL, 5 μM), and Fetal Bovine Serum (final 4% v/v, Thermofisher) at 37 °C. Stop the reaction by adding dye (10 μL).

Polyacrylamide gel electrophoresis (PAGE) detection: Prepare 19% polyacryl-amide gel (16.8 g urea dissolve in 20 mL 40% acrylamide and 8 mL 5× TBE buffer, add 100 μL 10% APS and 20 μL TEMED). Load samples and start electrophoresis in 1× TBE buffer (21.6 g Tris

base, 11 g Boric Acid, 1.5 g EDTA add water to 2 L and adjust pH to 8.3) at 800 V for 1 h. Place the gel on dryer for 2 h. Expose on film for 1 h and develop. Gel image was quantified by ImageQuant TL software.

2.2.2 RNase H activity with Se-modified antisense DNA

2.2.2.1 RNase H expression and purification

Transformation of RNase H plasmid to BL21: Place human RNase H plasmids (From Dr. Wei Yang's lab) and competent BL21 cell on ice bath. Add plasmid (1 μL, 200 ng/μL) to the competent cell and keep on ice bath for 30 min. Keep it at 42 °C water bath for 90 s. Then put it back on ice bath for 3 min. Add S.O.C medium (200 μL, Invitrogen) and shake at 37 °C 225 rpm for 45 min. Then pour 100 μL culture to the LB plate with 100 μg/mL ampicillin. Incubate at 37 °C overnight.

Expression of RNase H: Pick single colony from the LB plate into 30 mL autoclaved LB broth with 100 μg/mL ampicillin. Shake at 37 °C 225 rpm overnight. Then add 10 mL overnight incubated BL21 to each litter of 2×1 L LB broth with 100 μg/mL ampicillin and shake at 37 °C 225 rpm for 3 h. Add IPTG (1 mL, 1 mol /L) to each litter of the broth to induce the expression of RNase H and continue shaking overnight at 20 °C. Harvest the cell by centrifuging at 4 °C 5,000 rpm for 30 min. Keep the precipitation at -80 °C. Dissolve the cell in His-tag column buffer A (20 mM Tris pH 7.4, 1 M NaCl, 10 mM Imidazole, 5% Glycerol). Sonicate the cells with 60 × 15 s bursts. After each 15 s sonication wait 20 s to let the suspension cool down. Centrifuge at 4 °C 10,000 rpm for 30 min. Keep the supernatant at -80 °C.

Purification by His-tag column: Wash the column by buffer B (20 mM Tris-HCl pH 7.4, 1 M NaCl, 300 mM Imidazole, 5% Glycerol) and followed by buffer A each for 5-column-volume.

Load sample. Wash the column by buffer A until the UV value is flat. Then wash the column by 10% buffer B until the UV value almost go back to the low value. Elute sample by gradient from buffer A to buffer B (up to 100% in 60 min). Dialyze the protein against 2 L Phenyl column buffer B (20 mM Tris-HCl pH 7.4, 100 mM NaCl, 5% Glycerol) overnight. Detect the concentration of sample by UV analysis. Use thrombin to cleavage His-tag by the concentration of 5 U/mg. Add same volume of buffer C [20 mM Tris-HCl pH 7.4, 100 mM NaCl, 4 M $(NH_4)_2SO_4$, 5% Glycerol].

Purification by Phenyl column: Wash the Phenyl column by buffer B (20 mM Tris-HCl pH 7.4, 100 mM NaCl, 5% Glycerol) and then buffer A [20 mM Tris-HCl pH 7.4, 100 mM NaCl, 2 M $(NH_4)_2SO_4$, 5% Glycerol], each for 5-column-volume. Load sample. Wash the column by buffer A until the UV value is flat. Elute sample by program increased buffer B (up to 100% in 60 min). Dialyze the protein against 2 L storage buffer (20 mM Tris-HCl pH 7.4, 2.8 mM β-mercaptoethanol, 100 mM NaCl, 5% Glycerol) overnight. Concentrate the sample to 8 mg/mL (0.5 mmol/L) and store the enzyme at -80 °C.

SDS-PAGE: Prepare the gel-loading sample with protein sample (2 μL) in mixture of water (6 μL) and 5×loading buffer (2 μL, 250 mM Tris-HCl pH 6.8, 10% SDS, 0.02% bromophenol blue, 5% β-mercaptoethanol, 30% Glycerol). Load samples on 4%-20% SDS gel (Bio-rad) and start electrophoresis in 1 × Running buffer (3.02 g Tris base, 14.4 g glycine, 1 g SDS, add water to 1 L and adjust pH to 8.3) at 150 V for 40 min. Take out the gel and put in Coomassie Brilliant Blue Staining Solution (Bio-Rad) in microwave for 1 min. Then put in destaining buffer (10% ethanol and 10% acetic acid in water) and shake at 37 °C for 1 h.

2.2.2.2 RNase H reaction

Digestion of biotin labeled RNA substrate by RNase H: Set up 10 μL reaction with biotin labeled RNA substrate (1 μL, 1 μM), DNA template (1 μL, 1 μM), 10× RNase H reaction buffer (1 μL, Biolab) and RNase H (1 μL, 100 nM) at 37 °C. Stop the reaction by adding dye (10 μL).

2.2.2.3 Biotin labeled RNA substrate detection

Polyacrylamide gel electrophoresis (PAGE) and transfer membrane: Prepare 19% polyacrylamide gel (16.8 g urea dissolve in 20 mL 40% acrylamide and 8 mL 5× TBE buffer, add 100 μL 10% APS and 20 μL TEMED). Load samples and start electrophoresis in 1× TBE buffer (21.6 g Tris base, 11 g Boric Acid, 1.5 g EDTA add water to 2 L and adjust pH to 8.3) at 150 V for 1 h. Soak the biodyne B nylon membrane (Thermofisher), gel and filter papers in 1× TBE buffer. Assemble the transfer sandwich and make sure no air inside. Transfer with Bio-rad Trans-Blot Turbo at 1.3 mA for 7 min.

Incubation and imaging: Block the membrane with warmed blocking buffer (20 mL, Thermofisher) for 30 min. Prepare conjugate/blocking buffer solution by adding the stabilized streptavidin-horseradish peroxidase conjugate (70 μL) to warmed blocking buffer (20 mL, Thermofisher). Incubate the membrane in conjugate/blocking buffer for 20 min. Wash membrane four times for 5 minutes each in 1× wash solution (20 mL) with gentle shaking. Incubate membrane in substrate equilibration buffer (30 mL) for 5 minutes with gentle shaking. Prepare chemiluminescent substrate working solution by adding luminol/ enhancer solution (6 mL, Thermofisher) to stable peroxide solution (6 mL, Thermofisher). Pour the substrate working solution onto the membrane and incubate membrane in the substrate solution for 3 minutes.

Remove and wrap the moist membrane in plastic wrap. Place the membrane in a film cassette and expose to film. Develop the film. Gel image was quantified by ImageQuant TL software.

2.2.2.4 UV-thermal denaturation

UV-thermal denaturation: The thermal denaturation experiments were performed on a Varian CARY-300 spectrophotometer equipped with a Peltier thermoprogrammer for data collection and analysis. Stoppered quartz cuvettes of 10 mm optical path length and 1 ml volume were used for the measurements. The cell holder was thermostatted with circulating water. The DNA/RNA samples were prepared by diluting to 4 μM each strand in buffer (350 mM NaCl, 10 mM sodium phosphate pH 7.0, 0.1 mM EDTA) and then annealing by heating up to 90 °C for 2 min and then slowly cool down to 4 °C. The temperature dependence of the absorption value of the DNA was monitored at 260 nm. The temperature of the cell holder was increased from 4 °C to 60 °C at a rate of 0.5 °C/min. A temperature probe, immersed directly in a control cuvette, measured the sample temperature. The thermal melting temperature (T_m) was determined from the peak of the computer-generated first derivative of the absorbance versus temperature profile.

2.2.3 Structure study on DNA/RNA/RNase H complex

Crystallization of DNA/RNA/RNase H complex: Prior to co-crystallization with RNase H, the purified native or Se-modified DNA (5'-ATGTCG-3') and RNA (5'-UCGACA-3') were annealed at 1:1 molar ratio by first heating the mixture to 90 °C for 1 min, and then allowing it to cool slowly down to 25 °C. The resulting DNA/RNA duplex was mixed with the protein (final concentration: 2 mg/mL) at 1:1 molar ratio in the presence of 2 mM $MgCl_2$. Co-crystallization of Se-DNA/RNA hybrid with RNase H was achieved by screening with the QIAGEN Classics Suite

Kit (www.qiagen.com). By using the sitting-drop vapor diffusion method at 25 °C, the crystals were readily obtained from the mixture #96 of the crystallization screen [Buffer: 0.1 M MES, pH 6.5; precipitant: 12% (w/v), PEG 20000].

Data collection and structure determination: Crystal diffraction data of the Se-DNA/RNA/RNase H complex were collected at beamline 8.2.1 and 8.2.2 in the Advanced Light Source (ALS). 25% glycerol was used as cryoprotectant while X-ray data were collected under the liquid nitrogen stream at 100 °K. Crystal of Se-DNA/RNA/RNase H complex diffracted X-rays to 1.80. The data were processed and scaled using HKL2000, and the structures were solved by molecular replacement using Phasing in CCP4i with Dr. Wei Yang solved structure (PDB# 2g8u) as model. The resulted model was refined using Refmac5 within CCP4i. The modified DNA was modeled into the structure using Coot. Metal ions and water molecules were added either automatically or manually using Coot.

2.2.4 Cellular level gene silencing with Se-modified antisense DNA

2.2.4.1 HeLa cell culture

HeLa Cell Recovery: Take the cell out from liquid nitrogen and keep it in 37 °C water bath. Centrifuge cells for 4 min at 1,000 rpm as soon as the media melt. Remove the supernatant. Resuspend cells in about 1 mL media. Transfer the cells to a cell culture flask with about 10 mL media. Continue culture the cell in incubator and change the media in 24 h.

HeLa cell passaging and splitting: Check cells in flask under microscope to confirm that the cells are 80% - 100% confluent. Use the aspirator, empty liquid media covering cells. Use PBS (5 mL, Thermofisher) to wash the cells for two times. Add trypsin (1 mL, Thermofisher) to flask and keep in incubator at 37 °C for about 3 min. Check cells in flask under microscope to confirm

that the cells are detached. Add DMEM media (5 mL, Thermofisher) to the cells and then transfer to a 15 mL tube. Centrifuge cells for 4 min at 1,000 rpm. Remove the supernatant. Cell pellet should remain at base of tube. Resuspend cells in about 10 mL and separate them to 3 or 4 flasks. Continue culture the cell in incubator.

HeLa cell cryopreservation: Check cells in flask under microscope to confirm that the cells are 80% - 100% confluent. Use the aspirator, empty liquid media covering cells. Use PBS (5 mL, Thermofisher) to wash the cells for two times. Add trypsin (1 mL, Thermofisher) to flask and keep in incubator at 37 °C for about 3 min. Check cells in flask under microscope to confirm that the cells are detached. Add DMEM media (5 mL, Thermofisher) to the cells and then transfer to a 15 mL tube. Centrifuge cells for 4 min at 1,000 rpm. Remove the supernatant. Cell pellet should remain at base of tube. Resuspend cells in about 1 mL media with DMSO (100 μL). Transfer the cells to a 1.5 mL Eppendorf tube. Keep the cell at 4 °C for 30 min. Then move to -20 °C for 1 h. Keep the cells at -80 °C overnight. Finally keep the cells in liquid nitrogen.

2.2.4.2 Transfection and fluorescence detection

Transfection: Check cells in flask under microscope to confirm that the cells are 70% - 90% confluent. Dilute enhanced green fluorescence protein (eGFP) plasmid (0.5 μL, 1 μg/μL) and antisense DNA (1 μL, 10 μM) in Opti-MEM medium (200 μL, Thermofisher). Dilute Lipofectamine Reagent (1.5 μL, Thermofisher) in Opti-MEM medium (200 μL, Thermofisher). Add diluted DNA to diluted Lipofectamine 2000 reagent and then incubate for 5 min at room temperature. Add DNA-lipid complex to cells. Detect the cells under fluorescence microscope after 24 h.

Fluorescence detection with plate reader: Grow cells on 12-well-plate. Allow cell to continue grow for 24 h or 48 h after transfection. Set excitation wavelength as 488 nm and emission wavelength as 509 nm. Fluorescence of the cell plates was quantified by plate reader.

2.2.4.3 Western blotting

Cell Collection: Use the aspirator, empty liquid media covering cells. Use PBS (5 mL, Thermofisher) to wash the cells for two times. Add trypsin (1 mL, Thermofisher) to flask and keep in incubator at 37 °C for about 3 min. Check cells in flask under microscope to confirm that the cells are detached. Add DMEM media (5 mL, Thermofisher) to the cells and then transfer to a 15 mL tube. Centrifuge cells for 4 min at 1,000 rpm. Remove the supernatant. Cell pellet should remain at base of tube. Add Ripa Lysis buffer (200 μL, Thermofisher) and keep on ice for 30 min. Then centrifuge the samples for 30 min at 10,000 rpm. Add 5× loading buffer (40 μL, 2 μL, 250 mM Tris-HCl pH 6.8, 10% SDS, 0.02% bromophenol blue, 5% β-mercaptoethanol, 30% Glycerol). Boil the sample at 100 °C for 5 min.

SDS-PAGE: Load samples on 4%-20% SDS gel (Bio-rad) and start electrophoresis in 1× Running buffer (3.02 g Tris base, 14.4 g glycine, 1 g SDS, add water to 1 L and adjust pH to 8.3) at 150 V for 40 min.

Transferring membrane: Prepare the PVDF membrane (Bio-rad) soaking in methanol for about 1 min. Soak the membrane, gel and papers in the transfer buffer. Assemble the transfer sandwich and make sure there is no air inside. Transfer with Bio-rad Trans-Blot Turbo at 1.3 mA for 7 min.

Antibody incubation and imaging: Wash the membrane with TBST (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1% Tween 20). Block the membrane with 5% milk in TBST at room

temperature for 1 h. Incubate the membrane with primary antibody overnight at 4 °C. Wash the membrane with TBST for 3 times. Incubate the membrane with secondary antibody at temperature for 1 h. Wash the membrane with TBST for 3 times. Then incubate the membrane with AP substrate for 3 min. Remove and wrap the moist membrane in plastic wrap. Place the membrane in a film cassette and expose to film. Develop the film. Gel image was quantified by ImageQuant TL software.

2.3 Results and Discussion

2.3.1 Se-modified antisense DNA resistance to serum

Five DNAs shared the same sequence were used to test the resistance to serum. One of them is native and other four have 5-MeSe-T modified at different positions. Their sequences and MALDI-MS analysis information are shown in **Table 2.1**.

Table 2.1 MALSI-TOF MASS analysis of 5-SeMe-T containing DNA for resistance test

#	DNA Sequence	Calc. $[M+H^+]^+$	Observed
2-1	5'- GTG CAC TGA TCA ATT AAT GTC GAC-3'	7352.9	7353.1
2-2	5'- GTG CAC TGA TCA ATT AAT GTC GAC-3'	7432.8	7433.2
2-3	5'- GTG CAC TGA TCA ATT AAT GTC GAC-3'	7432.8	7439.9
2-4	5'- GTG CAC TGA TCA ATT AAT GTC GAC-3'	7432.8	7437.1
2-5	5'- GTG CAC TGA TCA ATT AAT GTC GAC-3'	7432.8	7438.3

We set up the reactions with all five DNAs in 4% serum for 20 minutes and 60 minutes to compare their resistance to serum. The results are displayed in **Figure 2.4**. It is obvious that DNA 2-5 has the best resistance among all the sequences. And from the statistical result, we can also find all modified DNA has a better resistance to serum compared with native DNA.

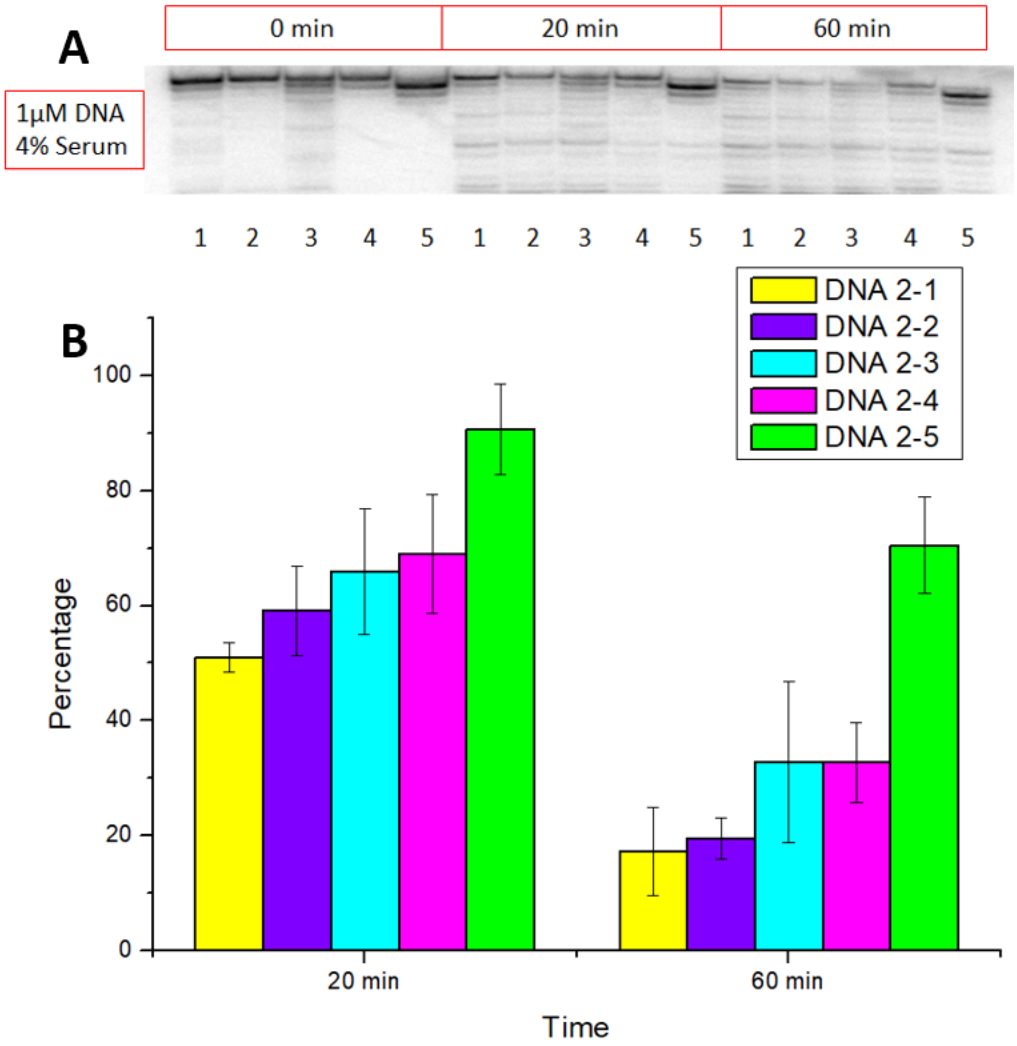


Figure 2.4 Resistance to serum of all five DNAs.
(A) Gel image of all 5 DNA resistance test to serum for 20 and 60 minutes.
(B) Statistical result of several repeats of the experiment.

We have further compared the time course digestion curve of DNA 2-1 and 2-5 (shown in **Figure 2.5**). It is very interesting that native DNA 2-1 was digested immediately after mixed with serum and continuously digested until 60 minutes later when there was almost no full-length DNA remaining. On the other hand, DNA 2-5 was also digested when mixed with serum. However, after about 20 minutes digestion, the full-length DNA kept about 60% remaining and no longer digested. We think the DNases in the serum were able to bind to the Se-modified DNA, but it might have

difficulties to cleave the substrate and finish the reaction. That's why the reaction rate was getting lower and lower and the full-length DNA finally kept about 60% remaining. Another possible explanation is that our modified antisense DNA make the RNase H enzyme lose its activity. No matter which is the reason that made this result happened, it proved that our 5-Se-T-modified DNA could increase the resistance against blood serum.

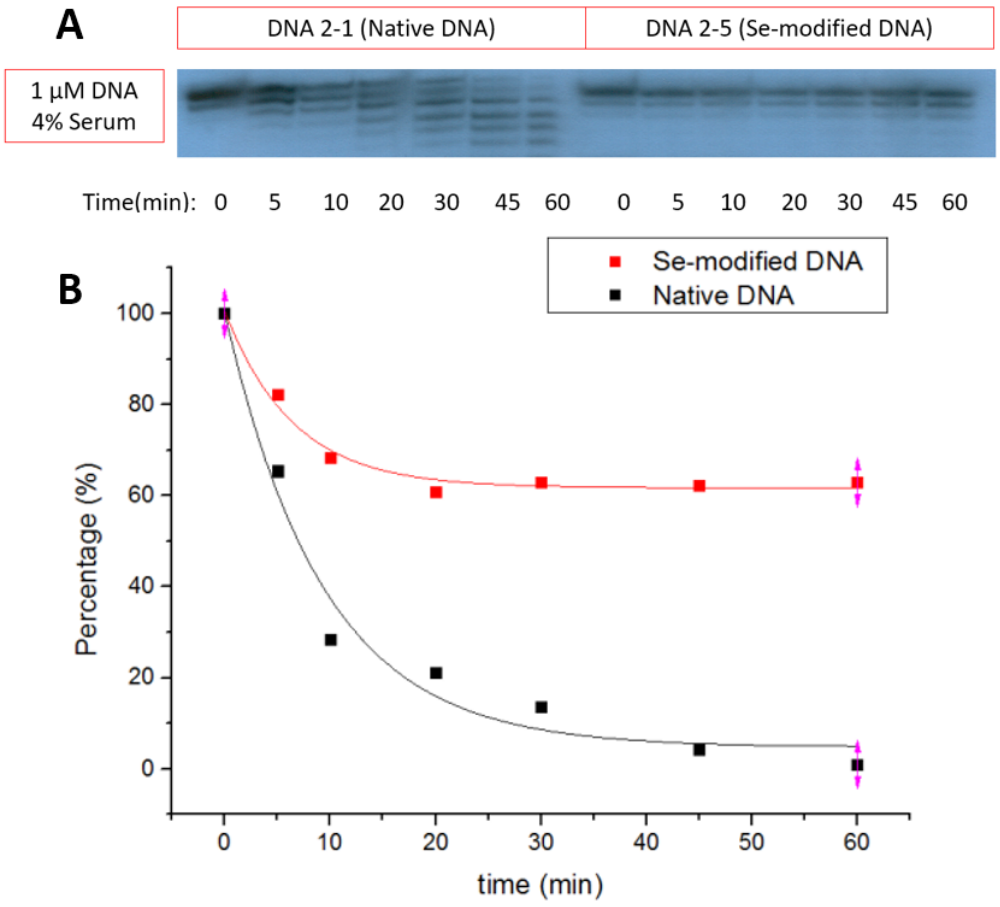


Figure 2.5 Time course experiment of resistance to serum of DNA 1 and 5.
(A) Gel image of time course serum resistance test for DNA 2-1 and 2-5.
(B) Digestion curve of DNA 2-1 and 2-5 during the time course experiment.

2.3.2 RNase H activity with Se-modified antisense DNA

RNase H is a kind of sequence non-specific nuclease that can digest the RNA when complementary DNA is hybridized, therefore, the RNase H mediated degradation of target mRNA is one of the basic mechanisms for antisense drug working. Some of the modifications in synthetic oligonucleotides could cause the loss of RNase H activity, probably due to the restriction in DNA backbone dynamic. Thus, although it was proved that our nucleobase modification could dramatically improve the stability against serum, whether it could still keep RNase H digestion on complementary mRNA is another concern.

2.3.2.1 RNase H expression and purification

RNase H (*Bacillus halodurans*) truncated (TR), and truncated mutant (D132N) constructs (pET 42 and pET13, respectively) were kindly given by Dr. Wei Yang lab at NIH as a gift. The truncated RNase H (59-196 aa) was created and retains the RNA hydrolytic activity on RNA/DNA duplex. Truncated inactive RNase H mutant was created, for crystal structure study, by mutating the aspartate residue (D132) of the truncated RNase H to asparagine (D132N). This mutation abolishes the metal ion coordination (B site) capability of the enzyme and makes the enzyme inactive catalytically.

Protein expressions were carried out in BL21 (DE3; pLys E. coli; purchased from Invitrogen). Then after two columns purification, we obtained pure truncated RNase H. We kept the samples in each step of purification running PAGE gel as in **Figure 2.6**.

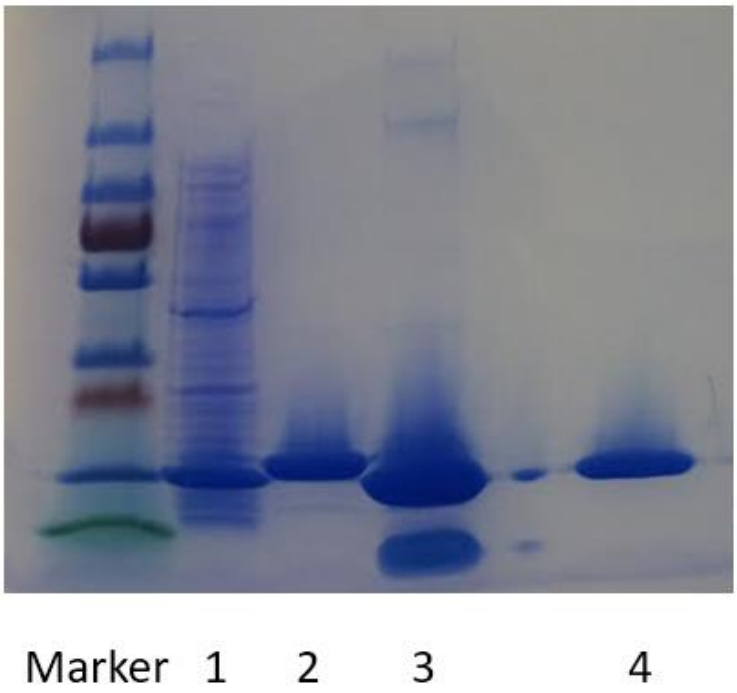


Figure 2.6 RNase H purification.
1: Supernatant after cell lysis; **2:** Purified by Ni column; **3:** After thrombin digestion and buffer change; **4:** Purified by phenyl column.

2.3.2.2 Se-modified antisense DNA's impact on RNase H cleavage

We designed hexamer DNA and RNA for RNase H digestion test. Four DNAs shared the same sequence were used as DNA guide in RNase H cleavage. One of them is native, two have single modification on different positions while the last one is double modified on both positions. The sequence and MALDI-MS analysis information are in **Table 2.2**.

Table 2.2 MALDI-TOF MASS analysis of DNA and RNA for digestion test

#	DNA Sequence	Calc. $[M+H^+]^+$	Observed
2-6	5'- ATG TCG-3'	1808.3	1807.3
2-7	5'- ATG TCG-3'	1887.3	1886.5
2-8	5'- ATG TCG-3'	1887.3	1885.7
2-9	5'- ATG TCG-3'	1966.2	1965.2
Biotin-RNA	5'-Biotin- UCG ACA-3'	-	-

We set up reactions with RNase H, DNA template and biotin-labeled RNA substrate (purchased from Integrated DNA Technologies) and build up time course curve of RNA digestion with different DNA templates. The results are displayed in **Figure 2.7**.

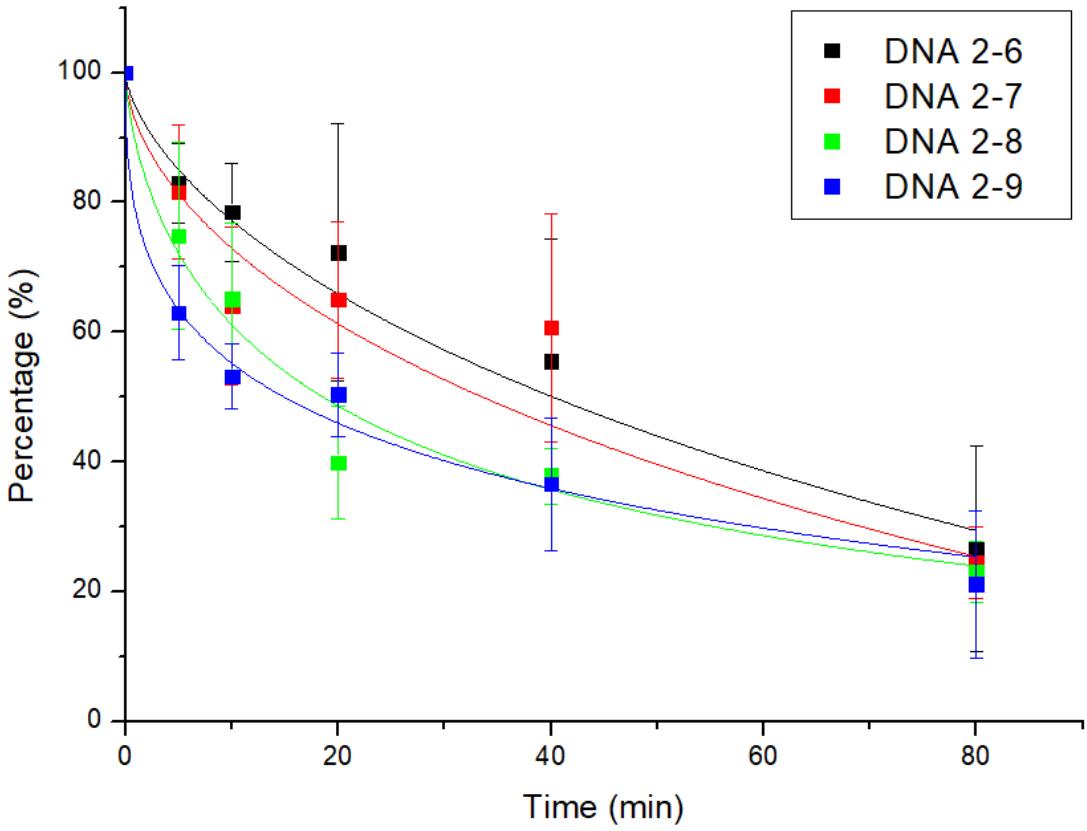


Figure 2.7 RNase H digestion curve of hexamer RNA binding to different DNA template.

We surprised to find all the modified DNA guides keep the compatibility with RNase H and all of them make RNase H have even higher activity. We calculated the initial rate of the curves and found that all the modified DNAs lead to faster initial reaction rate and double modified DNA 2-9 has the highest value. This result indicates our Se-modified DNA has good RNase H compatibility and even makes RNase H has higher activity.

However, because the sequence for our DNA/RNA duplex is very short, most of them in solution at room temperature may not appear to be duplexes. Since RNase H can only bind to DNA/ RNA duplexes to cleavage the RNA strand, that could be a question that how the Se-modification can improve the activity of RNase H. A possible explanation is that the Se-modification may stabilize the duplex and increase the ratio of DNA/RNA duplexes in aqueous. However, the previous research based on 5-Se-T modified DNA duplexes does not support this result. It is reported that 5-Se-T modified DNA duplexes have almost the same T_m comparing with native DNA with the same sequence[66]. DNA/RNA duplex may have the different performance with DNA duplex, so we also measured the T_m for all four duplexes. The results are shown in **Table 2.3**.

Table 2.3 Relative initail catalysis rate and T_m for different DNA/RNA duplexes

DNA #	DNA/RNA duplex sequences	T_m
2-6	5'- ATG TCG-3' / 5'- UCG ACA-3'	13.5 ± 0.5 ℃
2-7	5'- ATG TCG-3' / 5'- UCG ACA-3'	13.1 ± 0.7 ℃
2-8	5'- ATG TCG-3' / 5'- UCG ACA-3'	13.2 ± 0.5 ℃
2-9	5'- ATG TCG-3' / 5'- UCG ACA-3'	13.4 ± 0.4 ℃

It is obvious that Se-modifications do not make significant differences on the T_m of DNA/ RNA duplexes. So the thermos-stability of duplexes is not the critical reason of RNase H activity increase. We need to carry out a different explanation that why Se-modification can activate RNase H. In this way, we need to look in the structure of RNase H/DNA/RNA complex and try to draw out a better explanation on interactions between RNase H and DNA/RNA duplexes.

2.3.3 Structure study on RNase H/DNA/RNA complex

We already proved that our modified DNA have almost the same or sometimes even better ability to work with RNase H and explained that the reason has nothing to do with the thermos-stability. In order to understand how this happened through structural biology, we did crystal growth and X-ray structure determination to study how selenium affects the RNase H complex. The structure of *Bacillus halodurans* RNase H complexed with DNA/RNA duplex was well determined before by Dr. Wei Yang's lab. Here we use the same sequences of DNAs and native RNA (5'-UCGACA-3') for catalytic experiments (DNA sequence information is displayed in **Table 2.2**).

2.3.3.1 Crystallization of RNase H/DNA/RNA complex

We mixed 0.5 mM DNA and complimentary 0.5 mM RNA at 1:1 ratio. Then the DNA/RNA duplex was annealed by heating up to 80°C for 2 min, and slowly cooling down to room temperature. Then the duplex mixture was added to 3 mg/mL RNase H at 1:1 ratio in the presence of 2 mM $MgCl_2$. Crystallization screening was performed with QIAGEN Classics Suite Kit with 0.2 μL sample and 0.2 μL crystallization buffer by sitting drop vapor diffusion method at 25 °C. The crystals were readily obtained overnight from the mixture #96 of the crystallization screen [Buffer: 0.1 M MES, pH 6.5; precipitant: 12% (w/v), PEG 20000]. Then the crystallization was repeated manually with 1 μL sample and 1 μL crystallization buffer equilibrate against 500 μL of 15% (v/v) MPD using hanging drop vapor diffusion method at 25 °C to get more crystals. The crystal was mounted, shock-frozen in liquid nitrogen with 25% glycerol as cryo-protectant. The crystals are displayed in **Figure 2.8**.

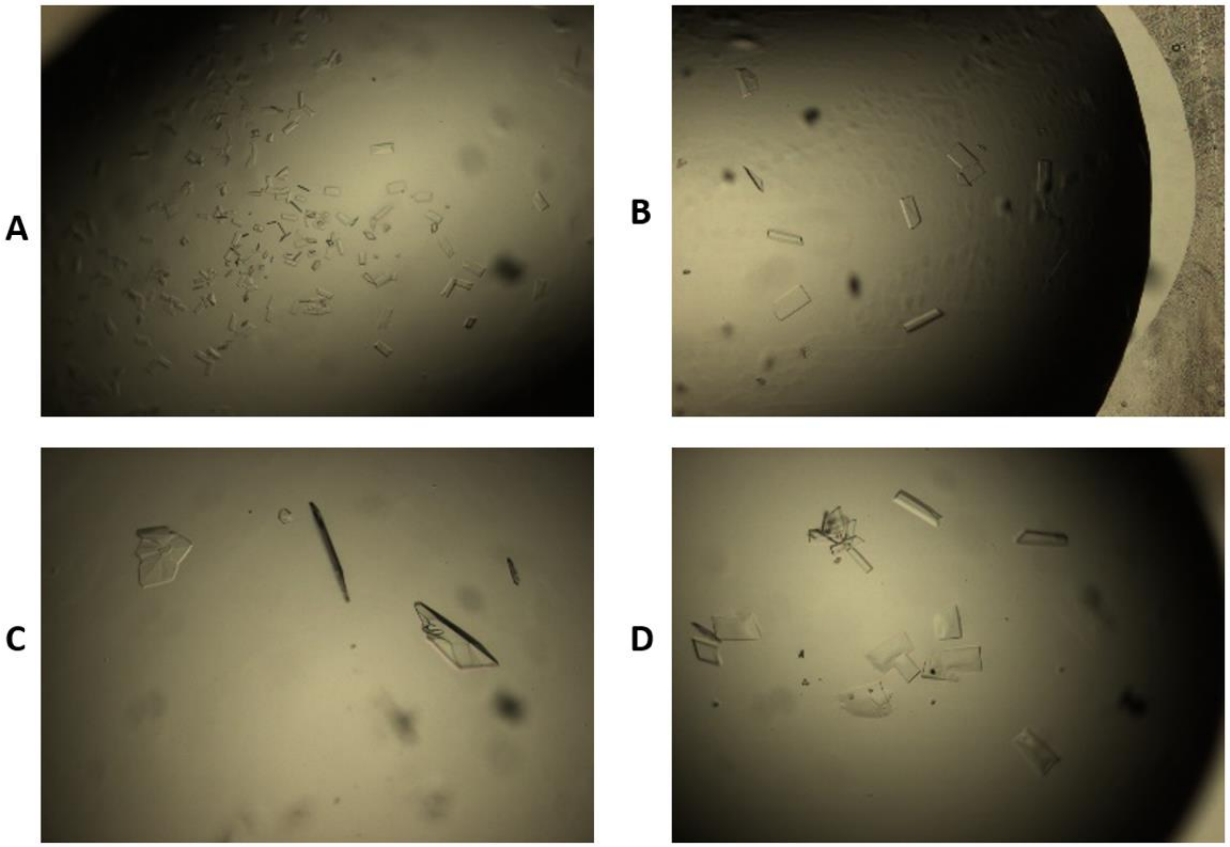


Figure 2.8 RNase H/DNA/RNA complex crystallization.
(A) DNA 2-6; (B) DNA 2-7; (C) DNA 2-8; (D) DNA 2-9.

2.3.3.2 Structure determination of RNase H/DNA/RNA complex

The diffraction data of RNase H/DNA/RNA complex crystal was collected at beamline 8.2.1 and 8.2.2 of the ALS (Advanced Light Source) at the Lawrence Berkeley National Laboratory. Several crystals were screened to find the crystal with the best diffraction quality. The diffraction data was collected at 1 Å wavelength for 1 second per frame with 1° oscillation angle. The best resolution for all 4 different DNA template was about 1.7 - 1.8 Å. The data were integrated and scaled by HKL2000.

Table 2.4 Data collection and refinement statistics of RNase H/DNA/RNA complex

Complex structure	RNase H/ RNA/ DNA 2-6	RNase H/ RNA/ DNA 2-7	RNase H/ RNA/ DNA 2-8	RNase H/ RNA/ DNA 2-9
Wavelength (Å)	1			
Resolution range (Å)	50.0-1.70 (1.76-1.70)	50.0-1.80 (1.86-1.80)	50.0-1.80 (1.79-1.73)	50.0-1.70 (1.76-1.70)
Space group	C2			
Unit-cell a, b, c (Å)	80.2, 37.6, 62.1	81.2, 37.7, 62.2	80.9, 37.6, 62.2	81.0, 37.8, 62.0
α, β, γ (°)	90, 96.0, 90	90, 96.4, 90	90, 96.1, 90	90, 96.2, 90
Unique reflections	18866 (1081)	17472 (1744)	19601 (1917)	20662 (2028)
Completeness (%)	92.1 (53.7)	100 (100)	99.8 (99.5)	100 (100)
Rmerge (%)	7.7 (23.5)	6.4 (40.8)	8.9 (22.5)	5.9 (41.4)
I/σ(I)	21 (3.5)	29 (3.2)	54 (6.7)	31 (3.2)
R value (%)	18.9	18.9	19.2	17.8
Rfree (%)	23.4	21.8	23.4	22.1
Average B value ($Å^2$)	15.6	24.4	27.8	22.1
R.m.s.d. bond length (Å)	0.019	0.020	0.021	0.021
R.m.s.d. bond angle (°)	1.85	1.98	2.04	2.13

The structure of RNase H/DNA/RNA complex was solved by molecular replacement using Phaser in CCP4. The corresponding native structure obtained by Dr. Wei Yang's lab (PDB entry 2g8u; 2.70 Å resolution) with the same sequences was used as model[67]. The resulting model was refined using REFMAC in CCP4 and Coot. After several cycles of refinement, water molecules were added either automatically or manually using Coot. Data collection and structure refinement statistics are given in **Table 2.5**. Structure factor and final coordinate have been deposited in the Protein Data Bank (PDB ID: 5WJR, 5USA, 5USE, and 5USG).

2.3.3.3 Structure study of selenium atom impact on RNase H active center

As shown in **Figure 2.9A**, the overall structure and binding site of RNase H/DNA/RNA complexes are all the same despite modification number and location. In this case, RNase H is binding to the gap of RNA where the red scissors are in **Figure 2.9B**. And all four structures are superimposed very well.

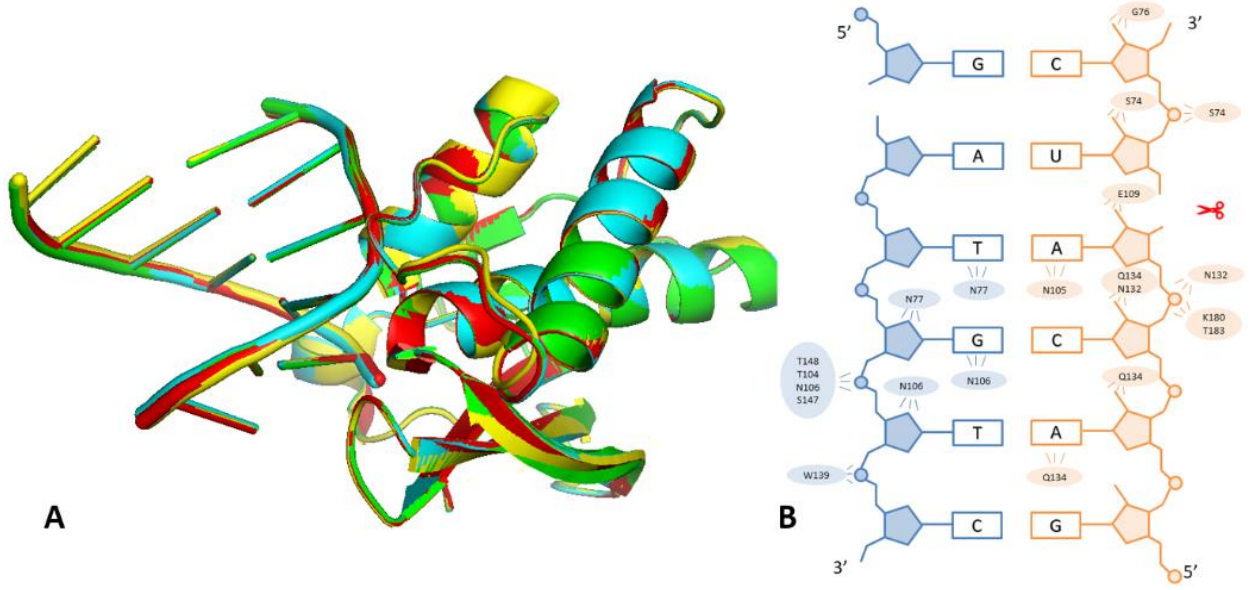


Figure 2.9 Overall structure and binding site of RNase H/DNA/RNA complexes.
(A) Overall structure of RNase H/DNA/RNA complexes. Complex with DNA 2-6 is green, DNA 2-7 is yellow, DNA 2-8 is cyan, DNA 2-9 is red.
(B) Binding site of RNase H/DNA/RNA complexes.

We used RNase H complex with native RNA and DNA 2-9, which has most selenium modifications, as an example to illustrate why Se-modifications will keep the overall structure in **Figure 2.10**. We can find that the RNase H enzyme is binding to the minor groove of DNA/RNA duplex while selenium modifications are in the major groove because it is modified on 5 position of thymidine. We believe this is the reason why the RNase H complexes display consistency on the structure and RNase H enzyme keeps working with modified DNA templates.

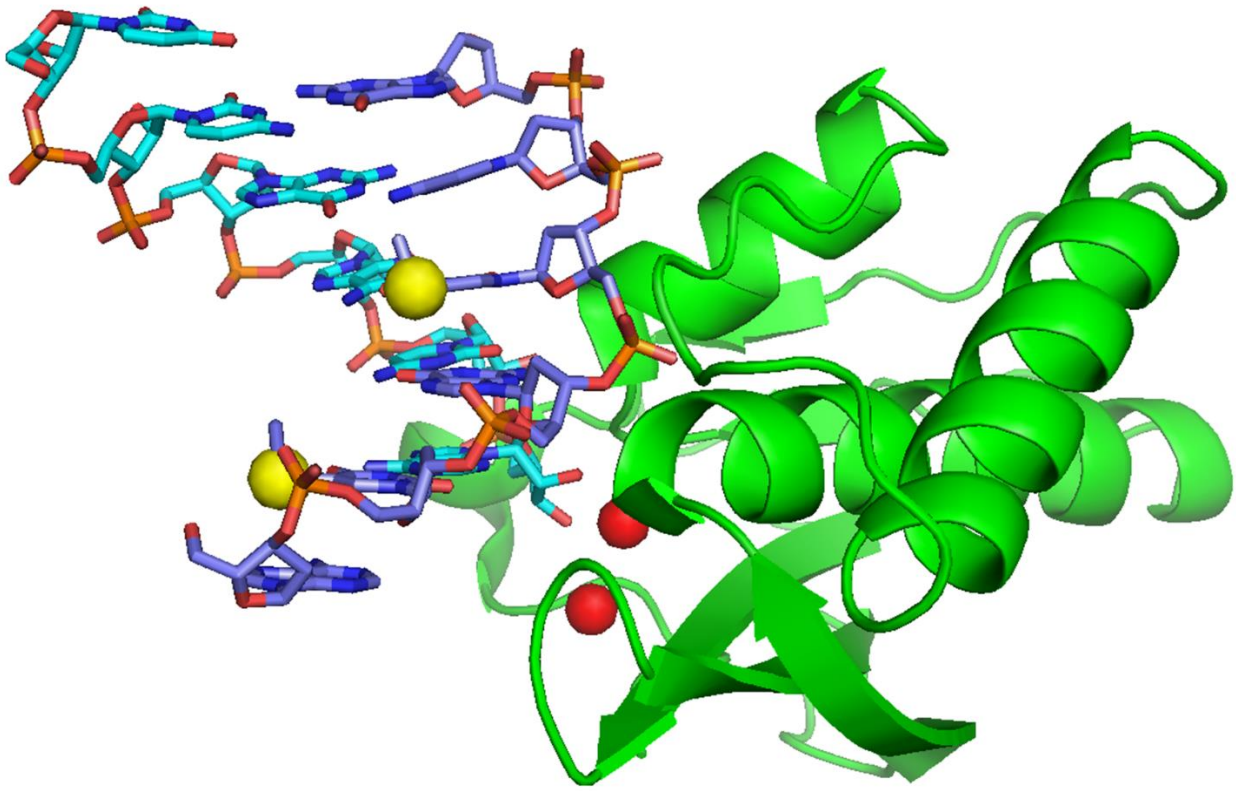


Figure 2.10 RNase H complex structure with native RNA and DNA 2-9
DNA strand is displayed in purple and RNA strand is displayed in cyan.
Yellow spheres represent selenium atoms and red spheres represent magnesium ion.

However, as we investigate the active center of RNase H, we can find some interesting results. **Figure 2.11** display the structure of base pair at the active site of RNase H for all 4 complexes. Because there is no phosphate group at the active center, we consider the distance between the magnesium ion and 3'-O on the sugar moiety of adenosine onsite as the most important value that how easy the reaction could happen. We believe the reaction is easier to happen when this distance is closer. We find if there is a modification on the deoxythymidine at the complementary place of adenosine onsite (DNA 2-7 and 2-9), the distance between magnesium ion and 3'-O on the sugar pucker of adenosine get closer. And this may be caused by changes on base pairs. We collected the related distance data and displayed those in **Table 2.5**.

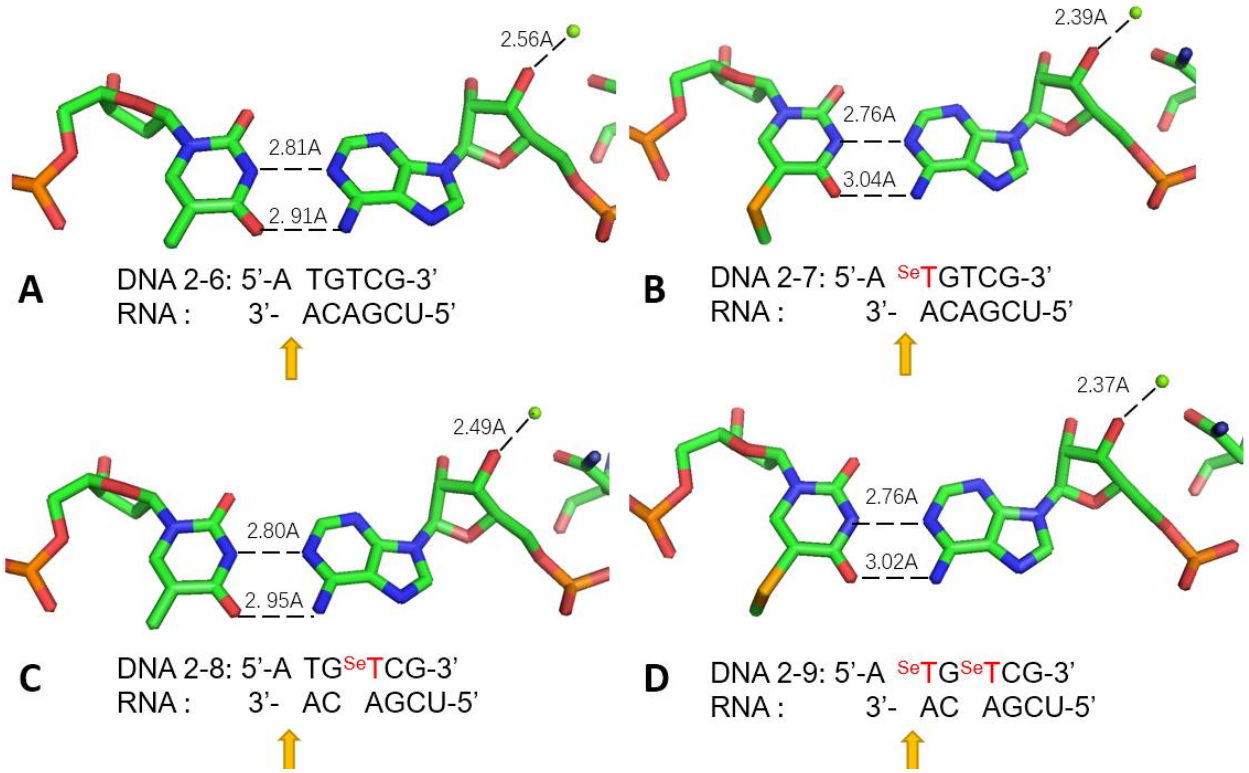


Figure 2.11 Active site structure of RNase H/DNA/RNA complex.
(A) DNA 2-6; (B) DNA 2-7; (C) DNA 2-8; (D) DNA 2-9.

Table 2.5 Important Distance between in RNase H/DNA/RNA complex

Distance (A) Between		DNA 2-6	DNA 2-7	DNA 2-8	DNA 2-9
Mg	O3'/A6	2.56	2.39	2.49	2.37
N3/T2	N1/A6	2.81	2.76	2.80	2.76
O4/T2	N6/A6	2.91	3.04	2.95	3.02

It seems that when there is a 5-Se-T on-site modification, the distance between N3 on dT and N1 on A get slightly closer, while the hydrogen bond between O4 on dT and N6 on A get longer. We think the large electron density of Se make the base pairs a little bit unwinding and push the Adenosine on RNA a little bit into the active core of RNase H and get closer to magnesium ion there. So in this way, the on-site modification could enhance RNase H activity. Off-site

modification may also have some indirect impact, and this is why DNA 2-8 also has a shorter distance between magnesium ion and 3'-O compare with DNA 2-6.

As we accept this hypothesis, we can try to explain the result for RNase H activity research. Although we are using the same sequences of DNA and RNA in RNase H activity study and structure determination, the cleavage site and binding site in the crystal are different. In crystallization, the RNase H binds to the gap between two RNAs. However, if RNase H binds to the same place in catalysis, RNA could not be cleaved. Only when RNase H active center binds to the middle of the sequence and there are phosphodiester bonds existing, the cleavage could happen, and we can observe RNA is cleaved. By investing our previous published paper, we think the cleavage site on RNA is between A4 and C5 (displayed in **Figure 2.12**). In this way, DNA 2-8 and 2-9 have the on-site modification while DNA 2-7 and 2-9 have the off-site modification. As we talked before, the on-site modification could lead to a significant increase in RNase H activity. This is constant with the catalysis results that DNA 2-9 has the highest RNase H activity and DNA 2-8 comes later.

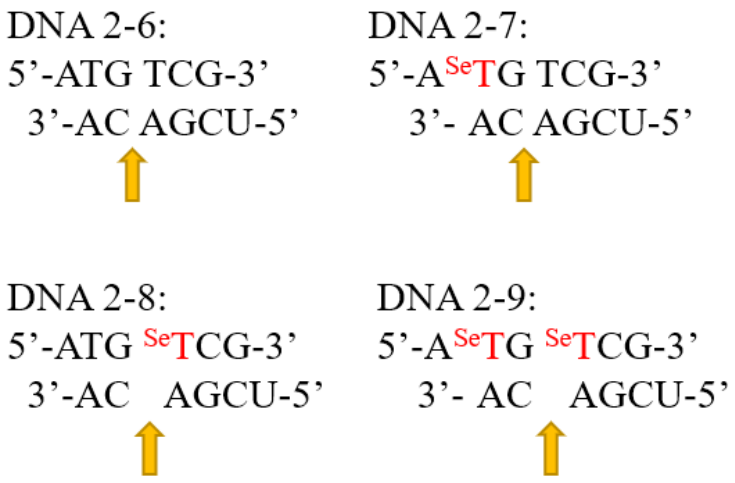


Figure 2.12 Cleavage site of RNase H catalysis research.

2.3.4 Cellular level gene silencing with Se-modified antisense DNA

It has been shown that our selenium modification has great potential as antisense drugs, due to both nuclease stability and RNase H compatibility. To investigate the gene silencing ability of this novel modified DNA targeting some specific gene in live cells, we performed the transfection experiment in HeLa cells using some CH_3Se-containing DNAs. To visualize the gene silencing effect, we initially transfected the cells with the eGFP-containing plasmid. The plasmid was a generous gift from Professor Jenny Yang's lab, GSU, and the green fluorescence protein gene was inserted in the pcDNA 3.1+ vector. The antisense oligonucleotide we chose was 18 nucleotides long and targeted to the position 1496 of the eGFP gene. The sequence of antisense DNA was indicated in **Table 2.6**, and the CH_3Se modification was labeled on those red T. 100 μM DNA solutions were prepared in water. The HeLa cells were grown in Dulbecco' MEF containing 10% heat-inactivated fetal bovine serum, 200 μg/mL glutamine, 100 u/mL penicillin, and 100 μg/mL streptomycin at 37 °C, 5% CO_2 in a moist atmosphere.

Table 2.6 Sequence information of 5-SeMe-T DNA for cellular experiments

#	DNA Sequence
2-10	5'-GAG CTG CAC GCT GCC GTC-3'
2-11	5'-GAG CTG CAC GCT GCC GTC-3'
2-12	5'-GAG CTG CAC GCT GCC GTC-3'
2-13	5'-GAG CTG CAC GCT GCC GTC-3'
2-14	5'-GAG CTG CAC GCT GCC GTC-3'

2.3.4.1 Fluorescence determination results

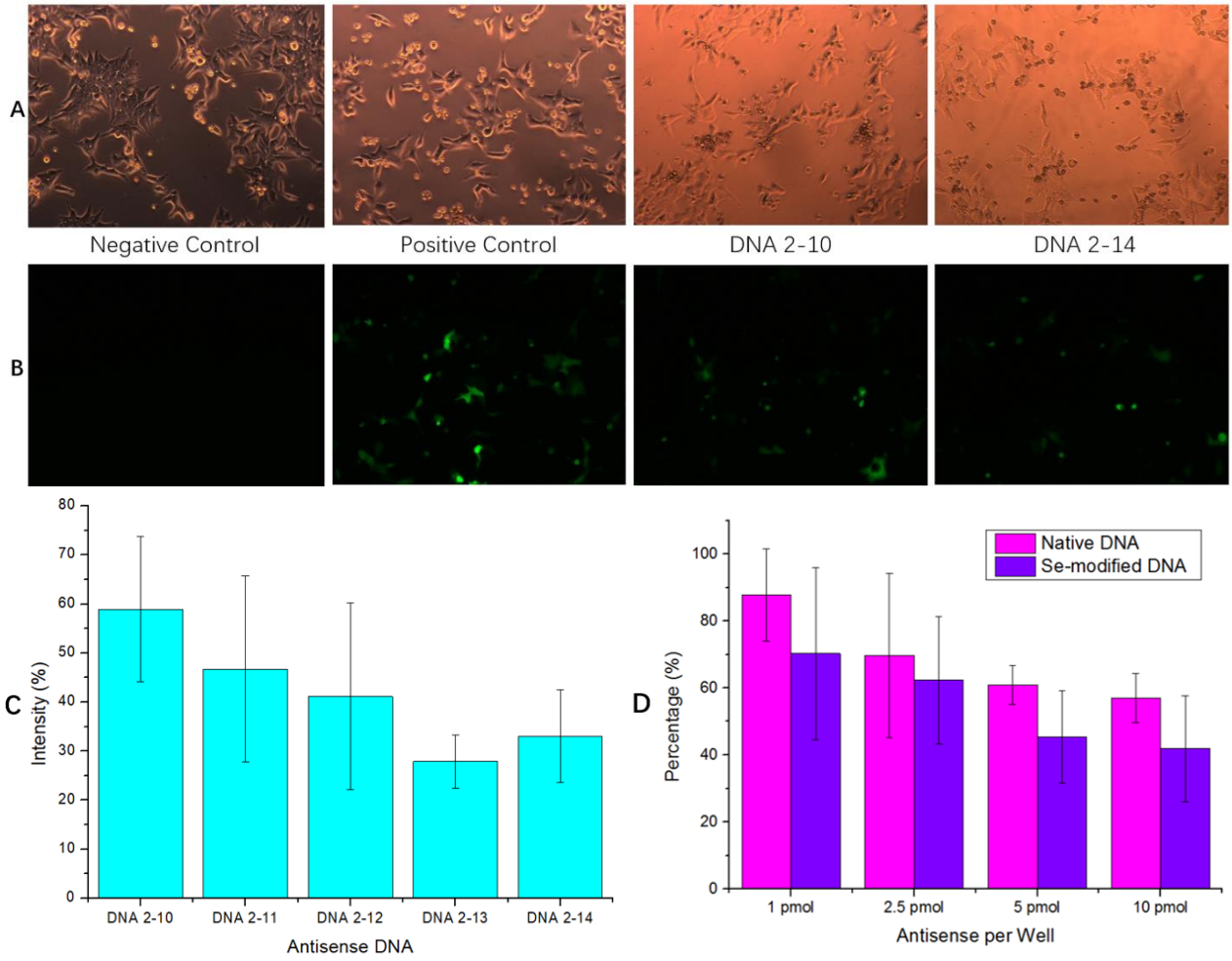


Figure 2.13 Cell images under fluorescence microscope.

(A) Cell images under white light.

(B) Cell images excited by blue light.

(C) Statistic result of fluorescence intensity for all 5 DNAs by plate reader.

(D) Statistic result of fluorescence intensity for different amount of DNA 2-10 and 2-14 by plate reader.

Cells were seeded in 6-well plates 24 h prior to transfection at a density of 6 x 105 cells per well and incubated in 3 mL medium. For the transfection, the medium was exchanged with fresh medium. The eGFP gene-containing plasmid and antisense oligonucleotides were added simultaneously to the cells with lipofectamine 2000 reagent assisting and incubated for 24 h. After 24 h incubation, the cells were visualized under Axiovert 200 fluorescence microscope excited by

blue light. The fluorescent intensity was measured by Plate Reader to have an idea about gene expression level. The results were demonstrated in **Figure 2.13**. Without transfecting plasmid, no eGPF was expressed at all (negative control). While without introduction of antisense oligonucleotides, the eGFP gene was actively expressed (positive control). When Cells were transfected with both eGFP plasmid and antisense DNAs, the gene expression was significantly suppressed. And Se-modified antisense DNA made the GFP expression even lower than with native antisense DNA.

2.3.4.2 Western blotting results

We performed Western Blot experiment to systematically quantify the different oligonucleotides' gene silencing activity. After gene expression, HeLa cells were washed three times with phosphate saline buffer (PBS) and then scraped into 0.1 mL of RIPA lysis buffer containing the protease inhibitor. After incubation for 10 min on ice, the lysate was transferred to new tubes, followed by centrifuging at 14,000 g for 15 min at 4°C. After analyzing lysates by SDS-PAGE gel, the analysis verified the downregulation of EGFP expression in cells treated with our antisense DNAs. Moreover, our Se-oligonucleotide has about 20% more gene silencing activity than the native one. Finally, the relative eGFP expression was calculated as the ratio of the fluorescence intensity on the control of beta-actin expression. Our quantitative result proved that DNA 4 and 5 with CH_3Se-modifications, silenced the eGFP gene expression more effectively. The gel image and quantitative result are displayed in **Figure 2.14**.

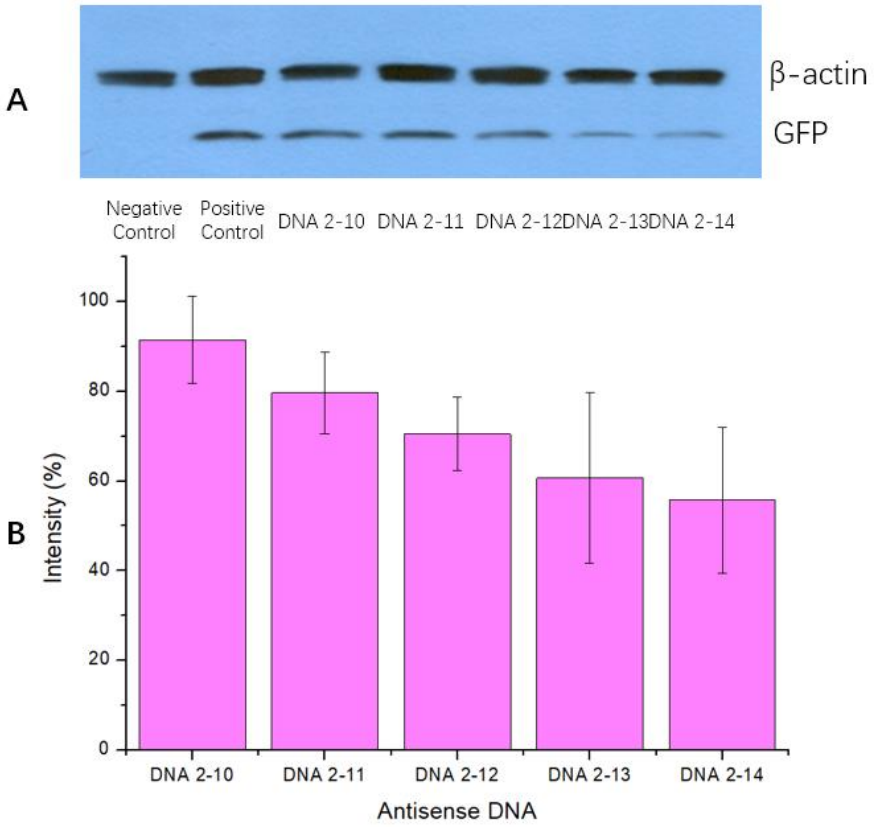


Figure 2.14 Western blotting results.
(A) Gel image for Western blotting.
(B) Quantitative result for Western blotting.

2.4 Conclusion

The derivatization of DNA with selenium can give us a new strategy in the antisense drug application. The 5-Se-T modification on oligonucleotides provided better resistance to blood serum. A reasonable hypothesis is that the selenium atoms at 5-position has steric hindrance with some enzyme residue so that the protein structure slightly shifts and cleavage is inhibited. Moreover, even though the methyl selenyl group has an impact on nuclease catalysis, it maintains the RNase H activity when the modified DNA binds to complementary RNA because it does not cause significant structure changes. And for our case, the modified DNA even lead to higher initial catalysis rate of RNase H. We think this is because introducing selenium atom to on-site position

lead to base pair unwinding and the complementary RNA strand is pushed a little bit into the RNase H activity center. This generates an important pathway for antisense oligonucleotide working. The gene silencing experiments in HeLa cells also demonstrated that our novel modified antisense DNA has a brilliant prospective. We anticipate the nucleobase derivatization with selenium element can provide novel efficient chemical modification strategy in gene therapy design drug market.

3 SE-ANTISENSE DNA X-RAY FLUORESCENCE IMAGING

3.1 Introduction

Massager RNA (mRNA) is the center of the central dogma. It is the bridge between genetic DNA and functional protein. Translation is the process that ribosomes synthesize protein based on the information in mRNA. It is very important to understand where and how this process happens. Most scientists detect this process through newly generated fluorescence protein labeling[68]. However, during the translation process, we may not be able to visualize the protein before it is well folded and modified. When start visualizing the functional protein, it is already translated and well folded, however, it may not keep the same location where it was synthesized. If we can detect mRNA and protein at the same time, then we could judge that the place where the signal of mRNA and protein overlapped is the location that translation is taking place. In this way, not only to determine the protein but to visualize the localization and transportation of mRNA is very important to.

Currently, most scientists are using fluorescent labeling for intracellular nucleic acids and oligonucleotides detection[69-72]. As a mains research method developed for decades, fluorescent labeling has many advantages. Since many of the fluorescent molecules with different excitation and emission wavelengths are developed[73], it is possible for us to detect several different targets with different labeling at the same time. Real-time intracellular monitoring inside living cells is also accomplished with fluorescent labeling[74-75]. However, there are still some problems for nucleic acid detecting with fluorescent labeling. First of all, fluorescent labels are usually some small molecules attached to 5’ or 3’ of the nucleic acid[73]. These fluorescent molecules usually have aromatic rings, which may affect in base stacking and change the structure of nucleic acid. In this way, they may cause the off-target problem during the experiment. Quenching is another

possible problem, which will cause low intensity in long-term experiment[76]. In addition, there may be also some other molecules that could generate fluorescence on the same excitation wavelength and cause high background[77].

Compared with normal fluorescence detection with small molecule labeling, micro-X-ray fluorescence (microXRF) detection with specific element could be more specific with higher intensity and lower background[78-79]. In this way, microXRF could be a better method for intracellular large biomolecule detection. Some metal ions have been used as the biosensor to detect intracellular biological processes involved different enzymes[80-85]. However, there is no effective sensor for nucleic acid so far. We developed Se-modified nucleic acids (SeNAs) for microXRF. SeNAs are single atom modification on nucleic acids. In our previous work SeNAs show the ability to remain the same structure and properties as native in many cases[24-25]. And most of the SeNAs are stable under normal experimental condition. Moreover, the intracellular Se concentration is very low[86], so there could be almost no background for microXRF. All above conclude that SeNA could be an ideal method for intracellular specific nucleic acid study.

In this way, we designed Se-modified antisense DNA with 5-Se-thymidine and the sequence targets to green fluorescence protein (GFP) mRNA. When we transfect this antisense DNA into the cell, it will bond to GFP mRNA (**Figure 3.1**). We can use microXRF to get the selenium signal which indicates where mRNA is. Meanwhile, we can also detect new generated GFP by fluorescence microscope at the same time. After comparing these two signals, we could tell where translation is taking place.

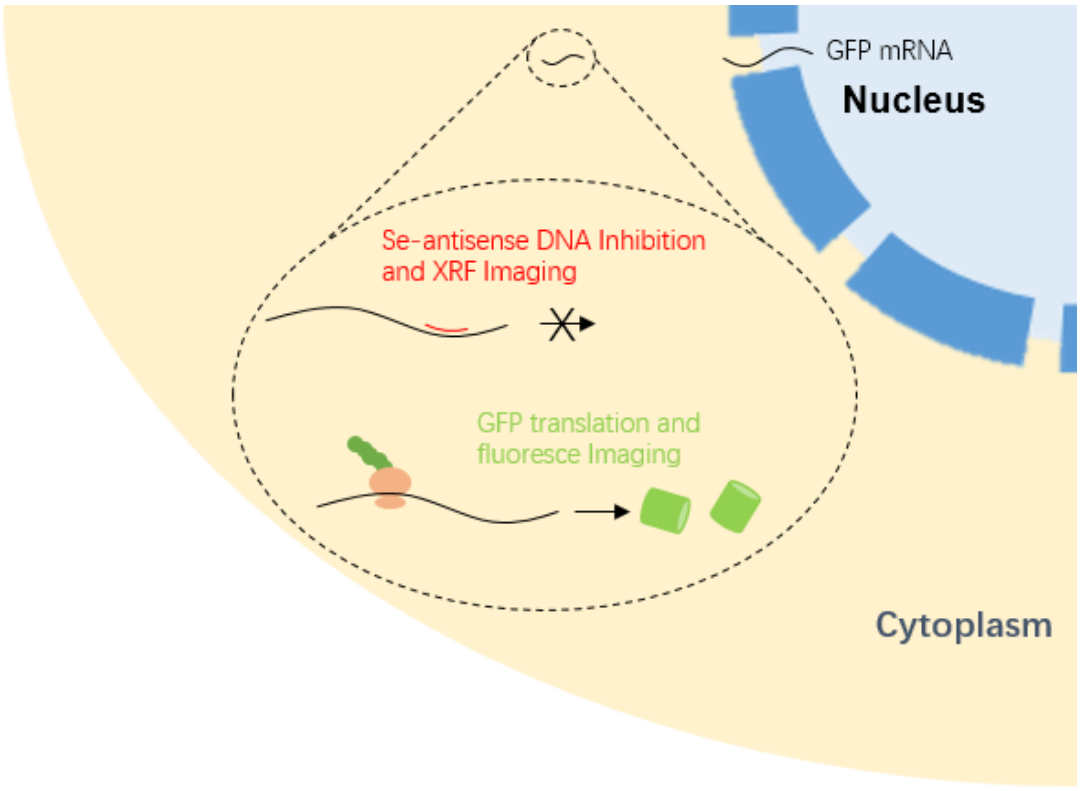


Figure 3.1 Processing of the experiment.
After transcription, mRNA come out from nucleus to cytoplasm. Protein is synthesized on ribosome which binds to mRNA. This processing can be blocked by antisense DNA that has Se on it in our design.

3.2 Material and Methods

HeLa cell culture on microchips: Check cells in flask under microscope to confirm that the cells are 80% - 100% confluent. Use the aspirator, empty liquid media covering cells. Use PBS (5 mL, Thermofisher) to wash the cells for two times. Add trypsin (1 mL, Thermofisher) to flask and keep in incubator at 37 °C for about 3 min. Check cells in flask under microscope to confirm that the cells are detached. Add DMEM media (5 mL, Thermofisher) to the cells and then transfer to a 15 mL tube. Centrifuge cells for 4 min at 1,000 rpm. Remove the supernatant. Cell pellet should remain at base of tube. Resuspend cells in about DMEM media (10 mL, Thermofisher).

Place 4-5 microchips in each well of a 6-well plate. Place the resuspended cells in 6-well plates with microchips.

HeLa cell transfection with microchips: Check cells in flask under microscope to confirm that the cells are 70% - 90% confluent. Dilute eGFP plasmid (2 μL, 100 ng/μL) and antisense DNA (0.5 μL, 10 μM) in Opti-MEM medium (200 μL, Thermofisher). Dilute Lipofectamine Reagent (0.6 μL, Thermofisher) in Opti-MEM medium (200 μL, Thermofisher). Add diluted DNA to diluted Lipofectamine 2000 reagent and then incubate for 5 min at room temperature. Add DNA-lipid complex to cells. Let the cells continue growing for 24 h.

Cell Fixation: After 24 h, remove the medium. Wash cells twice with PBS buffer (200 μL each, Thermofisher) pre-warmed at 37 °C. Fix cells with 3.7% formaldehyde. Incubate at room temperature for 10 min. Remove 3.7% formaldehyde. Wash the cells with PBS (200 μL each, Thermofisher) for 3 times. Wash the cells 3 times with isotonic ammonium acetate (0.1 M, 200 μL each) in sterile distilled water. Air-dried overnight, and then take out the chips from wells, store and ship.

MicroXRF determination: We used XRF microprobe analysis to examine element distributions and concentrations in cells on microchips at beamline 2-ID-E of the Advanced Photon Source. An Si(III) single-bounce monochromator was used to tune the undulator synchrotron X-ray beam to 12.9 keV. The beam was focused into a 0.5 μm spot using Fresnel zone plate optics. For element distribution maps, the sample was rastered through the focal spot with a step of 1-2 μm and dwell time of 4 sec. A full energy-dispersive XRF spectrum was collected at each step using a Canberra three-element UltraLEGe detector (Canberra Instruments, Meriden, CT). Elemental maps were created by fitting the spectrum at each scan position with modified Gaussians using MAPS software (Vogt 2003). Area concentrations for elements were calculated by

comparing integrated intensities of fluorescence peaks with those from thin film standards (NBS 1832 and 1833; National Bureau of Standards, Gaithersburg, MD).

3.3 Results and Discussion

3.3.1 Selenium aggregation of different DNA sequences

Four different sequences of Se-modified DNA are used in the experiment and the sequences are displayed in **Table 3.1**. DNA 3-1 (perfect matched) is antisense DNA exactly target to GFP mRNA. DNA 3-2 is 4 bases mismatched and DNA 3-3 is 6 bases mismatched. DNA 3-4 (control DNA) is a random sequence that is designed not to target any mRNA in HeLa cells. DNA 3-1, 3-2 and 3-3 have 3 modifications on the sequences while DNA 3-4 have only 2 modifications. In order to make the selenium amount be the same level in cells, the DNA 3-4 was used to be 1.5-fold concentration comparing to the other three sequences. We co-transfected the Se-DNAs with GFP plasmid to HeLa cells at the same time. The X-ray fluorescence microscope images for zinc and selenium of cell transfected with perfect matched DNA 3-1, mismatched DNA 3-2 and 3-3, as well as control DNA 3-4 are shown in **Figure 3.2**. There is a very low level of natural selenium in human cells, so the selenium background is close to zero. Selenium signal of perfect matched DNA is focus and appears as condensed spot. For mismatched DNA, we can also determine the condensed spot signal. However, the intensity of the signal is only about 50% comparing with the perfect matched group. And there is no condensed spot at all in those cells transfected with control DNA.

Table 3.1 Sequence information of 5-SeMe-T DNA for Imaging experiments

DNA sensor	DNA Sequence
DNA 3-1 (perfect matched)	5'-GAG CTG CAC GCT GCC GTC-3'
DNA 3-2 (4 bp mismatched)	5'-TAT CTG AAC TCT GCA GTA-3'
DNA 3-3 (6 bp mismatched)	5'-TAT CTG AAC TCT GCA GTA-3'
DNA 3-4 (control DNA)	5'-CTC CCA TCC-3'

It is obvious that antisense DNA is going to bind and block GFP mRNA at certain area in the cells, so selenium aggregated at the certain places and appeared to be condensed spots in those cells transfected with perfect matched and mismatched Se-modified antisense DNA. What's more, the cells transfected with mismatched DNA has weaker Se-aggregation phenomenon and the more bases mismatched, the weaker Se aggregated. We believe the reason for this is that mismatched DNA has a weaker ability to bind to the target mRNA. Meanwhile, control DNA do not bind to GFP mRNA at all, so it would be spread to everywhere inside and outside cells, which means no 'hot spots' could be found in the images. In the way, the signal of Se also suggests where the target mRNA is.

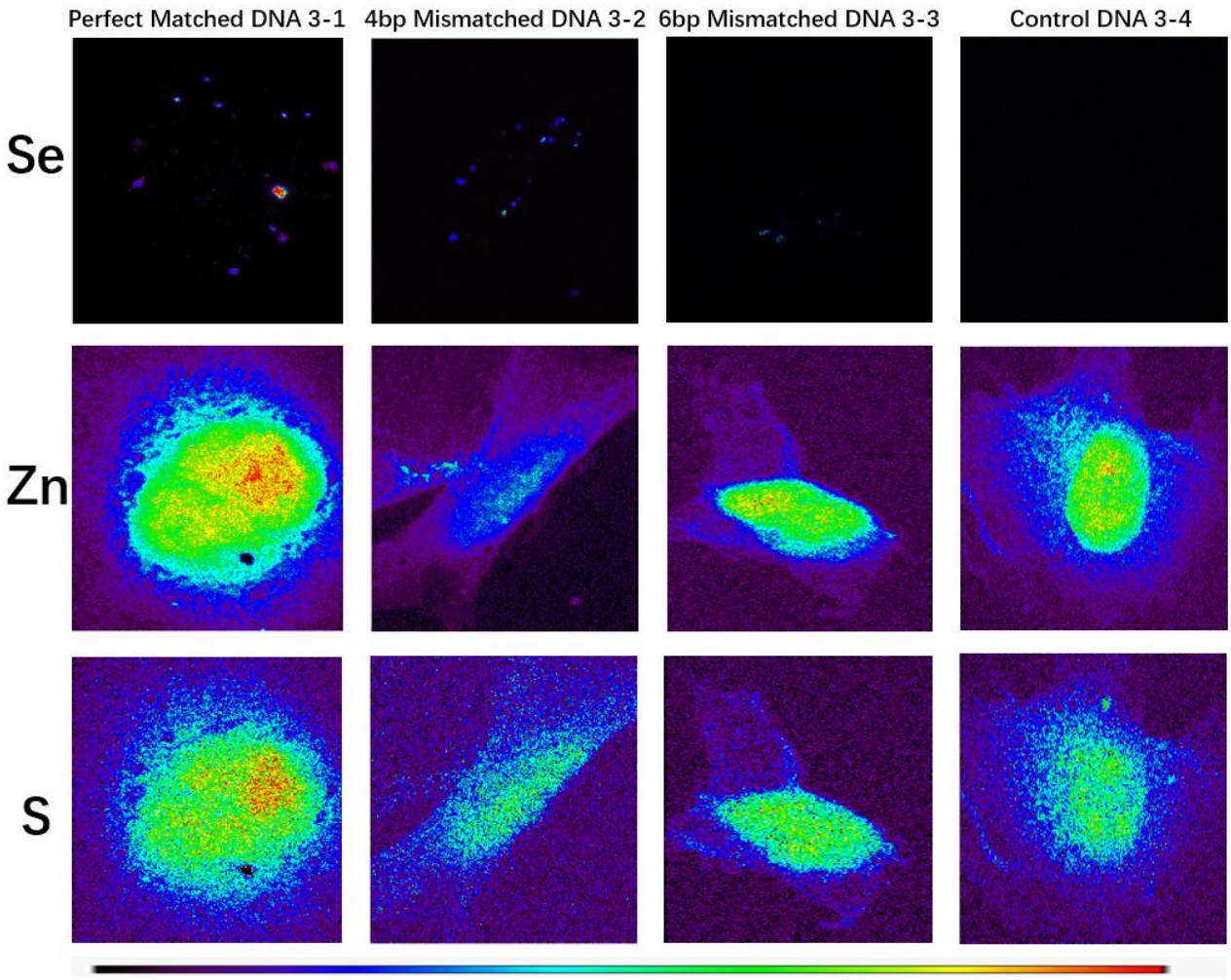


Figure 3.2 Images of Se-oligonucleotides distribution in cells by X-ray fluorescence visualization.
(A) DNA 3-1 is perfect matched with mRNA.
(B) DNA 3-2 has 4 base pairs mismatched.
(C) DNA 3-3 has 6 base pairs mismatched.
(D) DNA 3-4 is control sequence.

To compare the aggregation of Se more directly, we measured the strongest intensity of the signal in each picture and the average is displayed in **Figure 3.3**. The result agrees with our previous analysis that perfect matched DNA 3-1 has the highest average intensity, mismatched DNA 3-2 and 3-3 are weaker, and control sequence DNA 3-4 have the weakest signal and it is close to zero.

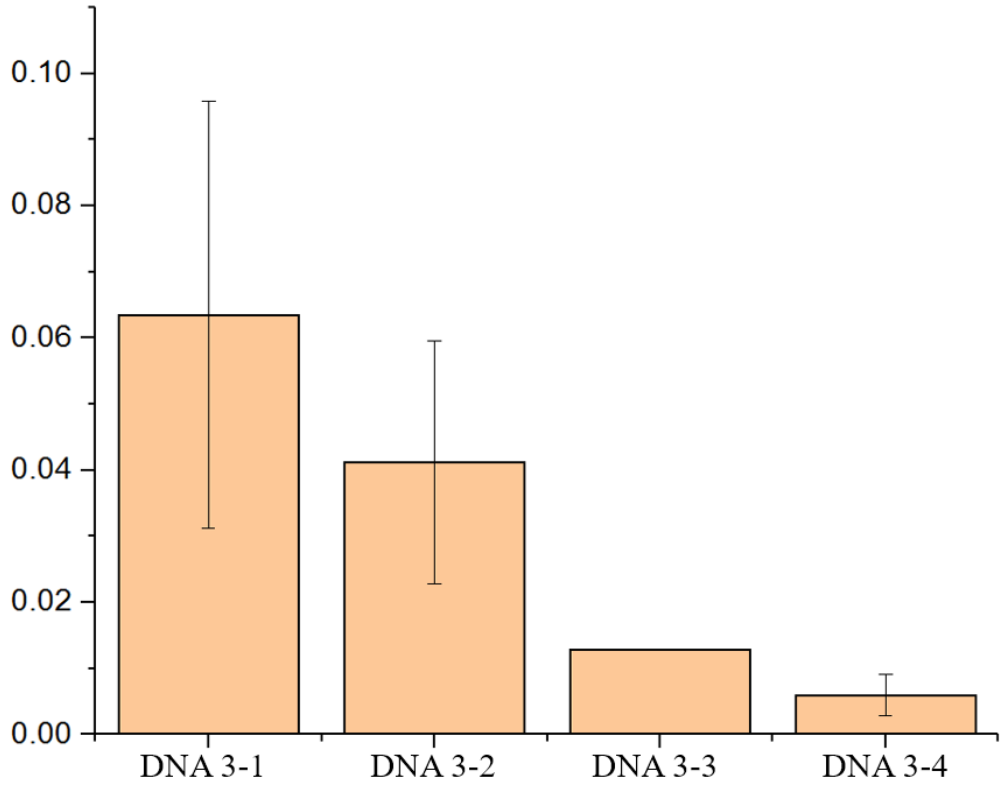


Figure 3.3 Average strongest intensity of Se for the pictures of each DNA.

3.3.2 Certain mRNA distribution in cells

We obtain even more images of Se and Zn for those cells transfected with DNA 1 and they are displayed in **Figure 3.4**. As we known, zinc is normally condensed in cell nucleus because of zinc figure structures[87]. We consider zinc signal as the location of the nucleus in the cell. Through the overlap pictures of zinc and selenium signal in **Figure 3.2 and 3.4**, we can conclude that Se-antisense DNAs are outside the nucleus.

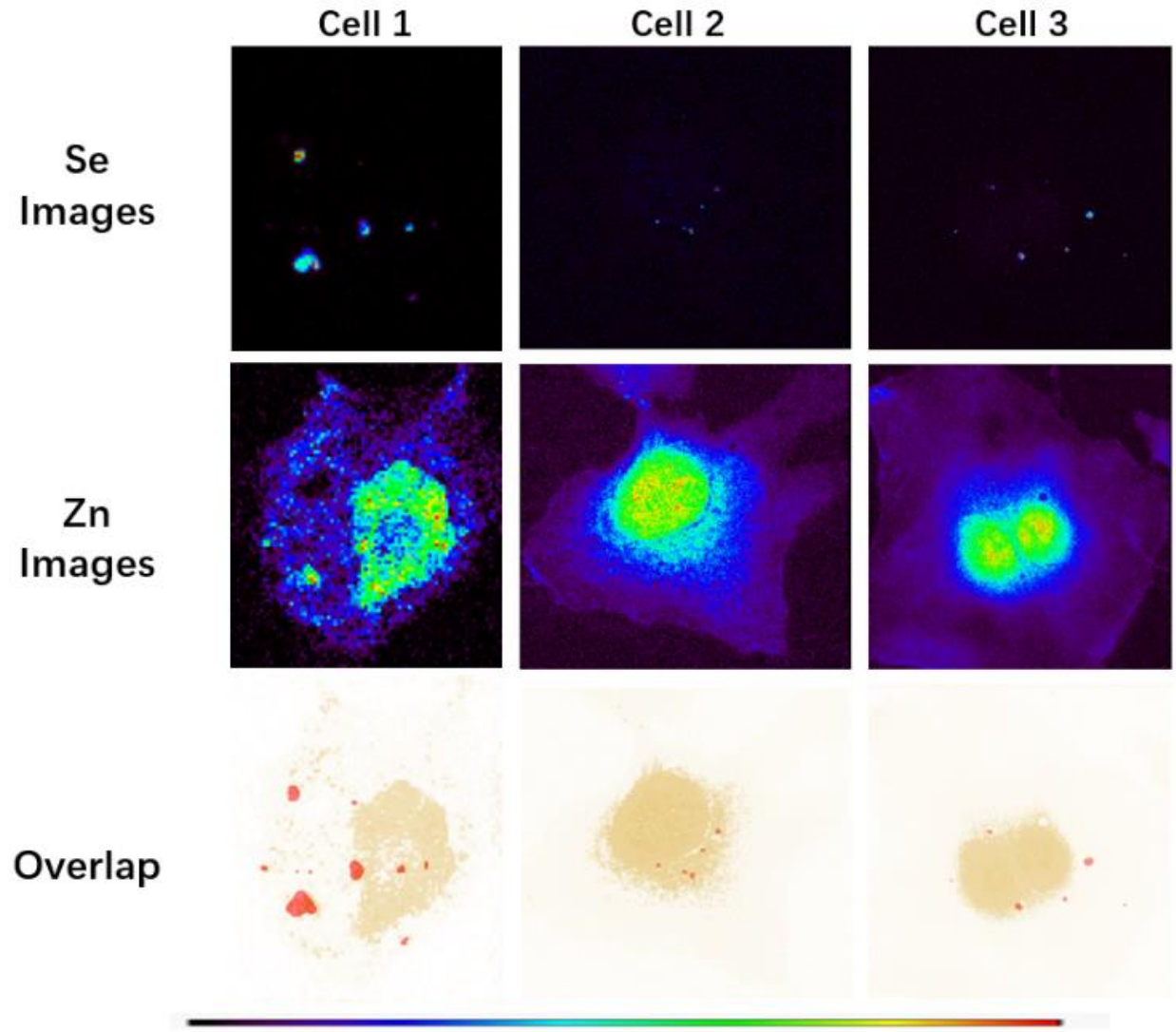


Figure 3.4 More images of Se and Zn distribution in cells transfected with perfect matched DNA 3-1 by X-ray fluorescence visualization.

It seems that the distribution of Se-antisense DNA is not even in cells. This indicates that mRNA of GFP is located at some certain places in the cytoplasm. In most of the cells, the signals are close to the nucleus. However, there are still some cells with selenium signal keeping away from nucleus. This may indicate that these cells are in different stages of the cell cycle. And we could speculate the routes of mRNA transportation in this way.

3.3.3 Co-location of Selenium and GFP

Figure 3.5 displays co-aggregation image of selenium and GFP in the cell transfected with perfect matched DNA. GFP signal was obtained by fluorescence detection. It seems that the aggregation spots of selenium and GFP overlap well in **Figure 3.5C**, which means Se-modified antisense DNA is located at the position where GFP was synthesized. As we mentioned the Se signal suggest the location of GFP mRNA, we think the spots that signal of Se and GFP overlapped are the place that GFP translation is taking place. In this way, we can confirm our hypothesis that the condensed spots of selenium signal indicate the places GFP is translating.

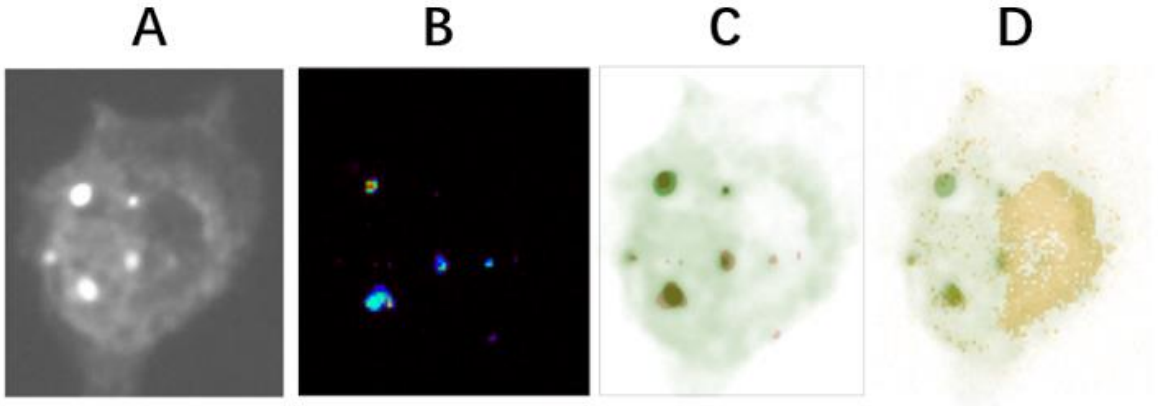


Figure 3.5 X-ray fluorescence microscope images of cell transfected with DNA 3-1.
(A) GFP fluorescence image;
(B) Selenium X-ray fluorescence image;
(C) Overlap image of Se (red color) and GFP (Green color);
(D) Overlap image of Zn (yellow color) and GFP (Green color).

3.4 Conclusion

We observed that Se-modified perfect matched antisense DNA aggregated surrounding nucleus in the cytoplasm when transfected with GFP plasmid together. This phenomenon was not found in those cells transfected with the Se-modified control sequence. Therefore, the Se-modified antisense DNA is binding the target mRNA at some certain area in cells. By analyzing many cell images in the different stage of cell cycle, we found that condensed selenium signal spots are

always away from zinc signal. As we considering zinc signal as the location of nucleus in the cell, we think Se-modified antisense DNA blocked mRNA outside the nucleus in the cytoplasm. Meanwhile, the aggregating spot is also the place where GFP is condensed. This result is consistent with our hypothesis that the condensed spot of Se suggest the position where translation is taking place. In this way, we can conclude that translation is taking place at certain position in cells and we can use Se-modified antisense DNA as a biosensor to detect mRNA during this processing.

4 DNA/DNA POLMERASE COMPLEX STRUCTURE STUDY

4.1 Introduction

DNA polymerases exist in all organisms and replicate genome. Most research on DNA polymerases is based on DNA Polymerase I (Pol I) from *Escherichia coli*[88]. Pol I has three discrete domains and each of them has different enzymatic activity: 5'-3' exonuclease, 3'-5' proofreading exonuclease, and polymerase activity[89]. The large C-terminal fragment, which is known as Klenow fragment (KF), includes both the 3'-5' exonuclease and polymerase domains[90]. The structure of KF was well determined last century[91-92]. When Pol I is binding to the DNA double strands, the conformation of the nucleotide is approximately that found in B-form DNA, having a C3' endo sugar pucker[92].

Many other DNA polymerases from different families were well studied after Pol I delineated[93]. One of the most important family is thermostable DNA polymerase, including *Thermus aquaticus* DNA polymerase I (Taq)[94] and *Bacillus stearothermophilus* DNA polymerase I (Bst)[95]. Those enzymes are very useful and are already widely used to amplify DNA sequences in polymerase chain reaction (PCR). The structure of both Taq[96-97] and Bst[98-99] were well determined.

The large carboxy-terminal fragment of Bst maintains the polymerase catalytic activity, which allows direct observation of the products of several rounds of nucleotide incorporation. And Bst just like other Pol I, when binding to the DNA duplex, the sugar pucker of nucleotide was found a C3' endo conformation, which is close to B-form DNA.

A big challenge for protein-nucleic acid complex structure determination is that we can hardly get good crystals for these complexes. So we designed Se-modified-DNA. We would like to use well studied Bst as an example and test if our Se-modified DNA could make some changes

when binding to Bst DNA polymerase and whether our Se-modified DNA could facilitate the complex crystal growth process. This could be a big discovery that Se-modifications on nucleic acids could not only facilitate the crystal growth for nucleic acids themselves but also works for protein-nucleic acid complexes. Meanwhile, if the structure of complexes does not have significant changes, this could give us a different method to grow the protein-nucleic acid complex crystals and investigate the structure of complexes. In this way, we may solve the problem of complexes crystallization and determine many more structures of protein-nucleic acid complexes.

4.2 Material and Method

4.2.1 Bst DNA polymerase expression and purification

Transformation of Bst DNA polymerase plasmid to BL21: Place Bst DNA polymerase plasmids and competent BL21 cell on ice bath. Add plasmid (1 μL, 200 ng/μL) to the competent cell and keep on ice bath for 30 min. Keep it at 42 °C water bath for 90 s. Then put it back on ice bath for 3 min. Add S.O.C medium (200 μL, Invitrogen) and shake at 37 °C 225 rpm for 45 min. Then pour 100 μL to the LB plate with ampicillin (100 μg/mL). Incubate at 37 °C overnight.

Expression of Bst DNA polymerase: Pick single colony from the LB plate into 30 mL autoclaved LB broth with ampicillin (100 μg/mL). Shake at 37 °C 225 rpm overnight. Then add 10 mL overnight incubated BL21 to each litter of 2×1 L LB broth with ampicillin (100 μg/mL) and shake at 37 °C 225 rpm for 3 h. Add IPTG (1 mL, 1 mol /L) to each litter of the broth to induce the expression of RNase H and continue shaking overnight at 20 °C. Harvest the cell by centrifuging at 4 °C 5,000 rpm for 30 min. Keep the precipitation at -80 °C. Dissolve the cell in His-tag column buffer A (20 mM Tris pH 7.4, 1 M NaCl, 10 mM Imidazole, 5% Glycerol). Sonicate the cells with 60 × 15 s bursts. After each 15 s sonication wait 20 s to let the suspension

cool down. Centrifuge at 4 °C 10,000 rpm for 30 min. Heat up the supernatant to 60-70 °C for about 10 min. Centrifuge again at 4 °C 10,000 rpm for 30 min. Keep the supernatant at -80 °C.

Purification by His-tag column: Wash the column by buffer B (20 mM Tris-HCl pH 7.4, 1 M NaCl, 300 mM Imidazole, 5% Glycerol) and followed by buffer A each for 5-column-volume. Load sample. Wash the column by buffer A until the UV value is flat. Then wash the column by 10% buffer B until the UV value almost go back to the low value. Elute sample by gradient from buffer A to buffer B (up to 100% in 60 min). Dialyze the protein against 2 L Heparin column buffer A (20 mM Tris-HCl pH 7.4, 100 mM NaCl, 2.8 mM β- Mercaptoethanol, 5% Glycerol) overnight. Detect the concentration of sample by UV analysis. Use ULP1 to cleavage His-tag by the concentration of 5 U/mg.

Purification by Heparin column: Wash the Heparin column by buffer B (20 mM Tris-HCl pH 7.4, 100 mM NaCl, 2.8 mM β- Mercaptoethanol, 5% Glycerol) and then buffer A (20 mM Tris-HCl pH 7.4, 1 M NaCl, 2.8 mM β- Mercaptoethanol, 5% Glycerol), each for 5-column-volume. Load sample. Wash the column by buffer A until the UV value is flat. Elute sample by program increased buffer B (up to 100% in 60 min). Dialyze the protein against 2 L storage buffer (20 mM Tris pH 7.4, 2.8 mM β-mercaptoethanol, 100 mM NaCl, 5% Glycerol) overnight. Concentrate the sample to 24 mg/mL (0.4 mmol/L) and store the enzyme at -80 °C.

SDS-PAGE: Prepare the gel-loading sample with protein sample (2 μL) in mixture of water (6 μL) and 5×loading buffer (2 μL, 250 mM Tris-HCl pH 6.8, 10% SDS, 0.02% bromophenol blue, 5% β-mercaptoethanol, 30% Glycerol). Load samples on 4%-15% SDS gel (Bio-rad) and start electrophoresis in 1 × Running buffer (3.02 g Tris base, 14.4 g glycine, 1 g SDS, add water to 1 L and adjust pH to 8.3) at 150 V for 40 min. Take out the gel and put in

Coomassie Brilliant Blue Staining Solution (Bio-Rad) in microwave for 1 min. Then put in destaining buffer (10% ethanol and 10% acetic acid in water) and shake at 37 °C for 1 h.

4.2.2 Structure study on DNA/Bst DNA polymerase complex

Crystallization of DNA/Bst DNA polymerase complex: 1 mM self-complementary DNA were annealed by first heating to 90°C for 1 min, and then allowing it to cool slowly down to 25 °C. The resulting DNA solution was mixed with the protein (final concentration: 24 mg/mL) at 1:1 molar ratio. Co-crystallization of Se-DNA/RNA hybrid with RNase H was achieved by screening with the QIAGEN Classics Suite Kit (www.qiagen.com). By using the sitting-drop vapor diffusion method at 25°C, the crystals were readily obtained from the mixture #67 of the crystallization screen [Salt: 0.2 M Calcium acetate; Buffer: 0.1 M sodium cacodylate, pH 6.5; precipitant: 18% (w/v), PEG 8000].

Data collection and structure determination: Crystal diffraction data of the Se-DNA/RNA/RNase H complex were collected at beamline 8.2.1 and 8.2.2 in the Advanced Light Source (ALS). 25% glycerol was used as cryoprotectant while X-ray data were collected under the liquid nitrogen stream at 100 °K. Crystal of DNA/Bst DNA polymerase complex diffracted X-rays to 1.80. The data were processed and scaled using HKL2000, and the structures were solved by molecular replacement using Phasing in CCP4i with Dr. Wei Yang solved structure (PDB# 2g8u) as model. The resulted model was refined using Refmac5 within CCP4i. The modified DNA was modeled into the structure using Coot. Metal ions and water molecules were added either automatically or manually using Coot.

4.2.3 Circular dichroism (CD)

CD spectra: The DNA solution was prepared by diluting to 4 μM in buffer (100 mM NaCl, 10 mM sodium phosphate pH 7.0, 0.1 mM EDTA) and then annealing by heating up to 90 °C for 2 min and then slowly cool down to room temperature. CD spectra were recorded on a JASCO J-1500 spectropolarimeter. Four scans of the spectrum were collected over the wavelength range 200 - 350 nm at a scanning rate of 20 nm/min. The average of multiple scans was used for analysis. The scan of the buffer alone recorded at room temperature was subtracted from the average scans for each DNA duplex. Data were collected in units of millidegrees versus wavelength and normalized to total DNA concentration.

4.3 Results and discussions

4.3.1 Bst DNA polymerase expression and purification

The large fragment of DNA polymerase I from Bacillus stearothermophilus (Klenow Fragment) was expressed and purified in our lab. Protein expressions were carried out in BL21 (DE3; pLys E. coli; purchased from Invitrogen). Because the target protein is thermostable, we heated the sample to high temperature as a step to purify it. Under higher temperature, while many other proteins could be denatured and aggregated to precipitate, Bst DNA polymerase could keep stable and soluble in buffer solution. In this way, we can get purer protein before loading it to the affinity column for further purifications.

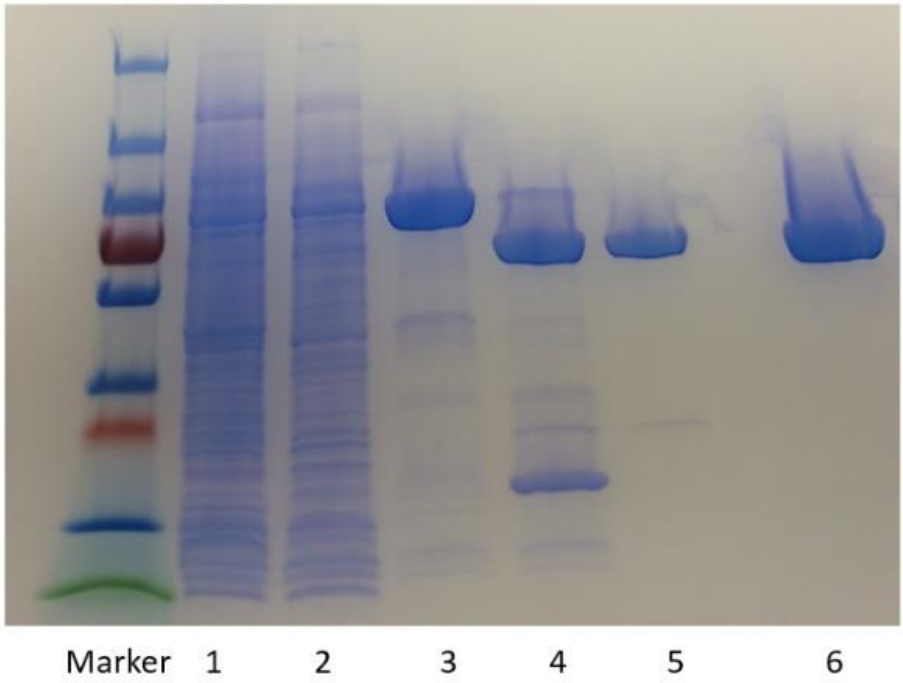


Figure 4.1 Bst DNA Polymerase Purification
1: Supernatant; 2: After heating 10 min at 60 °C; 3: 1st time for Ni column; 4: After cleavage; 5: 2nd time for Ni column (flow through); 6: Final sample.

After heating and two columns purification, we got very pure Bst DNA polymerase protein **(Figure 4.1**). And it is ready for crystallization and structure studies.

4.3.2 DNA/Bst DNA polymerase complex crystallization

Four DNAs shared with same sequence (5'-CGCGAATTCGCG-3') but modified on different positions were used for the comparison of DNA/Bst DNA polymerase complex crystallization. This sequence is well known as Dickerson dodecamer and is normally B form conformation, which is good for DNA polymerase binding. The sequence and MALDI-MS analysis information are shown in **Table 4.1**.

Table 4.1 MALDI-TOF MASS analysis of Dickerson dodecamer DNA

#	DNA Sequence	Calc. $[M-H^+]^-$	Observed
4-1	5'- CGC GAA TTC GCG -3'	3645.4	3645.4
4-2	5'- CGC GAA TTC GCG -3'	3724.4	3723.3
4-3	5'- CGC GAA TTC GCG -3'	3724.4	3725.2
4-4	5'- CGC GAA TTC GCG -3'	3803.3	3802.0

The 0.5 mM DNA was annealed by heating up to 80°C for 2 min, and slowly cooling down to room temperature. Then the DNA was added to 12 mg/mL Bst DNA polymerase at 1:1 ratio. Crystallization screening was performed with QIAGEN Classics Suite Kit with 0.2 µL sample and 0.2 µL crystallization buffer by sitting drop vapor diffusion method at 25 °C. The crystals were readily obtained overnight from the mixture #67 of the crystallization screen [Salt: 0.2 M Calcium acetate; Buffer: 0.1 M sodium cacodylate, pH 6.5; precipitant: 18% (w/v), PEG 8000]. Then the crystallization was repeated manually with 1 µL sample and 1 µL crystallization buffer equilibrate against 500 µL of 15% (v/v) MPD using hanging drop vapor diffusion method at 25 °C to get more crystals. The crystal was mounted, shock-frozen in liquid nitrogen with 25% glycerol as cryo-protectant. The crystals are displayed in **Figure 4.2**.

There were no significant differences found between native and Se-modified sequences. The shape and the speed of crystallization are almost the same.

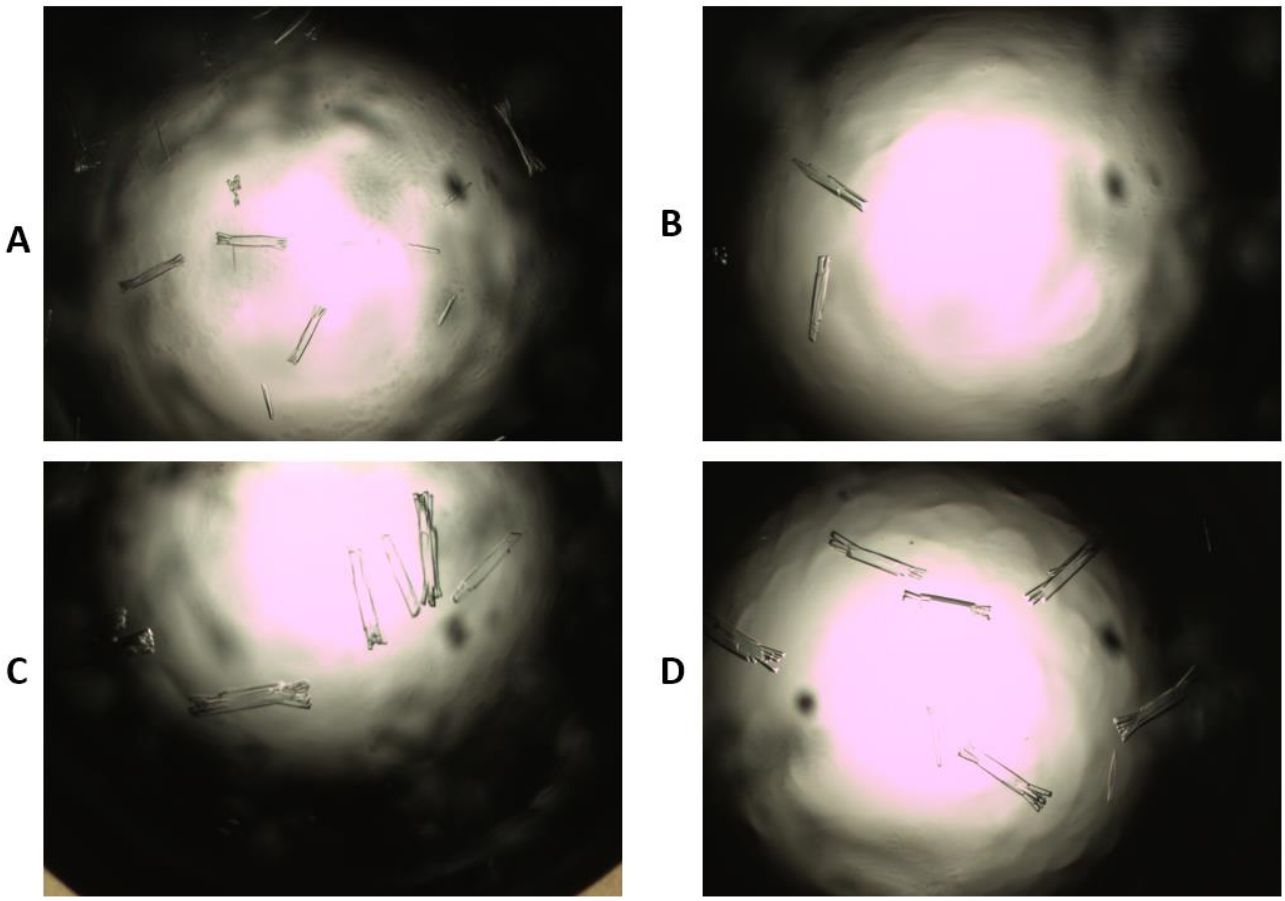


Figure 4.2 Dickerson dodecamer DNA/Bst DNA polymerase complex crystallization
(A) DNA 4-1; (B) DNA 4-2; (C) DNA 4-3; (D) DNA 4-4.

4.3.3 DNA/Bst DNA polymerase structure study

4.3.3.1 Dickerson dodecamer DNA/Bst DNA polymerase structure study

The diffraction data of Dickerson dodecamer DNA/Bst DNA polymerase complex crystal was collected at beamline 8.2.1 and 8.2.2 of the ALS (Advanced Light Source) at the Lawrence Berkeley National Laboratory. Several crystals were screened to find the crystal with the best diffraction quality. The diffraction data were collected at 1 Å wavelength for 1 second per frame with 1° oscillation angle. The best resolution for all 4 different DNA template was about 1.8-2.4 Å. The data were integrated and scaled by HKL2000.

The structure of Dickerson dodecamer DNA/Bst DNA polymerase complex was solved by molecular replacement using Phaser in CCP4. The corresponding native structure obtained by our lab previous work (PDB entry 4dsl; 2.45 Å resolution) with the same protein and different DNA sequence was used as a model. The resulting model was refined using REFMAC in CCP4 and Coot. After several cycles of refinement, water molecules were added either automatically or manually using Coot. Data collection and structure refinement statistics are given in **Table 4.2**.

Table 4.2 Data collection and refinement statistics of Dickerson dodecamer DNA/Bst DNA polymerase complex

Complex structure	DNA4−1/Bst	DNA4−2/Bst	DNA4−3/Bst	DNA4−4/Bst
Wavelength (Å)	1			
Resolution range (Å)	50-1.85 (1.92-1.85)	50.0-2.13 (2.21-2.13)	50-2.40 (2.49-2.40)	50-2.18 (2.26-2.18)
Space group	$P2_12_12_1$			
Unit-cell a, b, c (Å)	86.7, 94.1, 105.0	87.3, 94.2, 105.8	86.8, 94.3, 105.8	86.7, 94.2, 105.8
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Unique reflections	73185(6683)	47767(3919)	34483(3357)	45948(4496)
Completeness (%)	99.1(91.5)	96.2 (80)	99.7 (99.1)	100 (100)
Rmerge (%)	7.4 (43.1)	13.8 (48.3)	12.9 (50.7)	10.6 (48.7)
I/σ(I)	32.7 (3.6)	18.5 (2.8)	20.9 (5.4)	29.5 (5.7)
R value (%)	18.72	17.91	20.41	17.53
Rfree (%)	22.36	22.43	25.65	22.49
Average B value ($Å^2$)	22.65	30.18	26.99	33.84
R.m.s.d. bond length (Å)	0.0189	0.0185	0.0160	0.0192
R.m.s.d. bond angle (°)	1.938	2.027	1.857	2.062

As the structures shown of different angles in **Figure 4.3**, the overall structure of Dickerson dodecamer DNA/Bst DNA polymerase complexes are quite similar to each other despite the

different modifications. All four structures are superimposed very well. We believe in this case, the Se atom specific modifications do not change the structure of the complex as we found on many other different structures before.

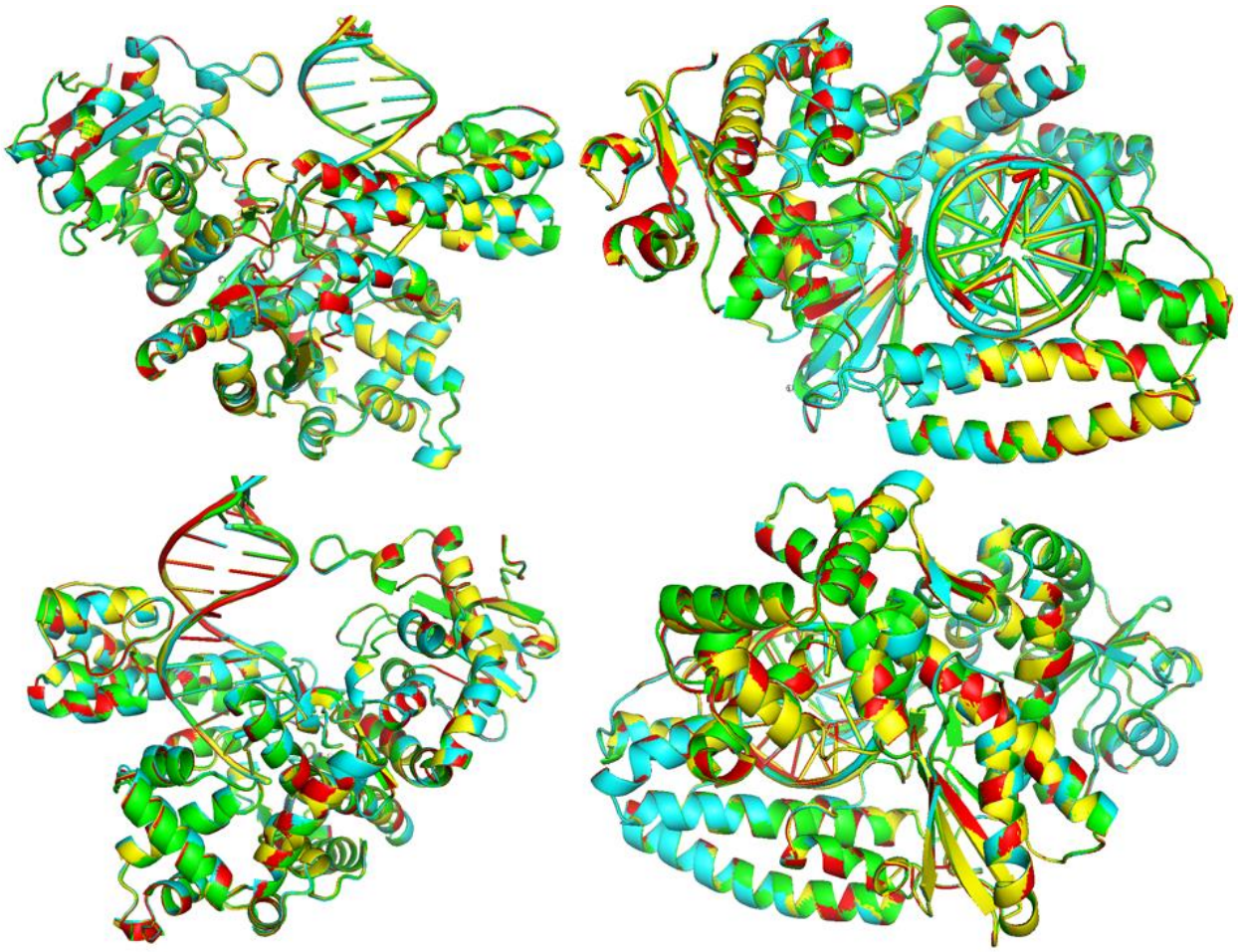


Figure 4.3 Different angle of overall structure of Dickerson dodecamer DNA/Bst DNA polymerase complex
Complex with DNA 4-1 is green, DNA 4-2 is yellow, DNA 4-3 is cyan, DNA 4-4 is red.

However, we would like to know why these modified DNA duplex will not change the structure of Bst complexes. We use Bst complex with DNA 4-4, which has the most selenium atoms, as an example (**Figure 4.4**) to display the interaction between Bst DNA polymerase and DNA duplex. We found that the Bst DNA polymerase is binding to the minor groove of DNA duplex. However, our modification is on 5 position on thymidine which is in the major groove.

And we notice that there is no interaction between the yellow sphere which represent selenium atoms and protein in **Figure 4.4**. Therefore, this modification will not perturb any interaction with Bst binding, so this modification will not change the overall structure of Bst complex at all.

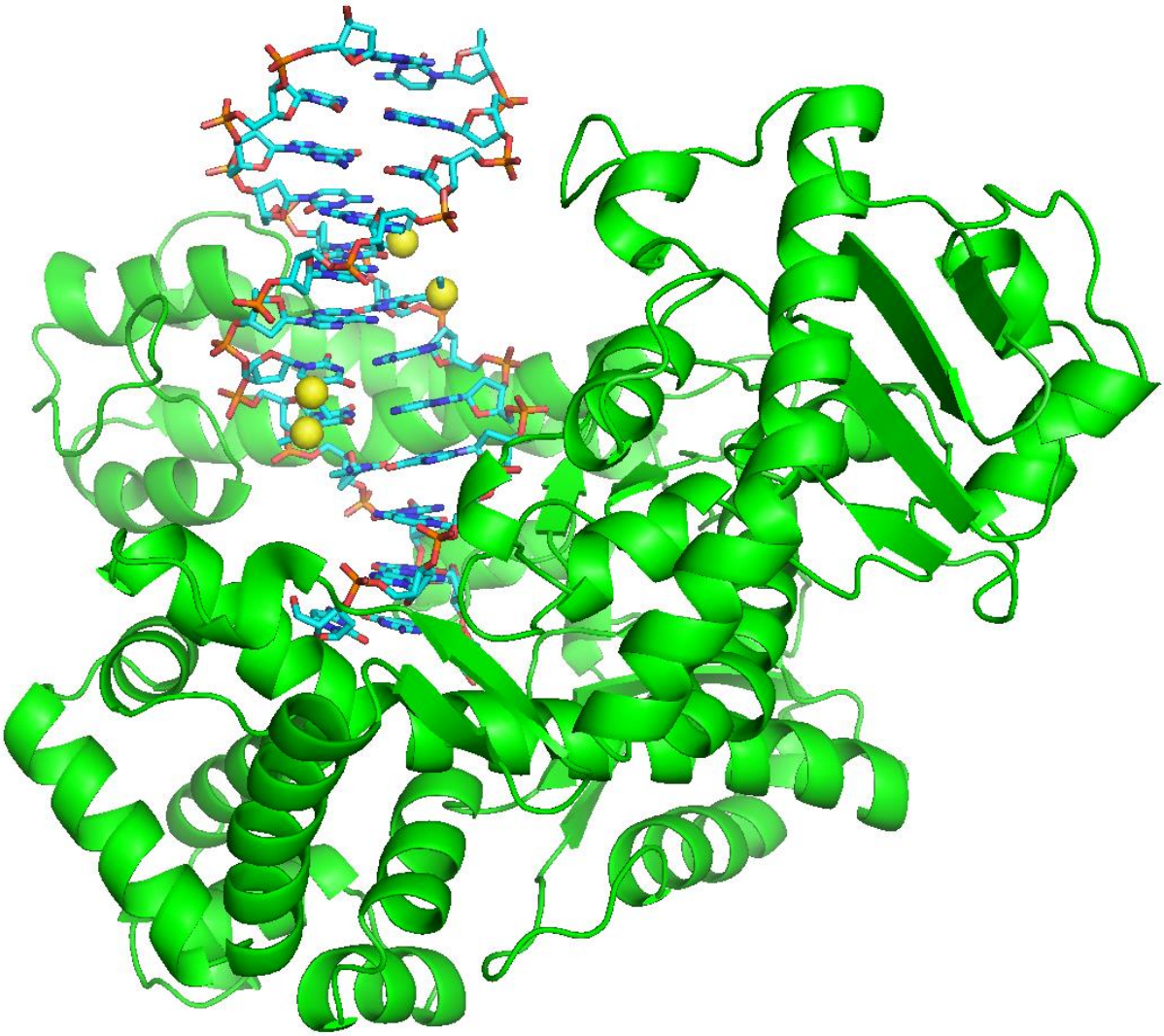


Figure 4.4 Structure of Bst complex with DNA 4-4
Yellow spheres represent selenium atoms.

Even though the Se-modifications have no impact on the structure of Bst DNA polymerase, we would like to compare the DNA duplexes in the structure of Bst complexes. The structures are displayed in **Figure 4.5**. Though the terminal base pairs show some slight disorder because of its

flexibility at that position, the main body parts of DNA duplexes display amazing consistency. The modifications on DNA duplexes do not only keep the overall structure of Bst DNA polymerase but also have no significant perturbation on DNA duplexes themselves.

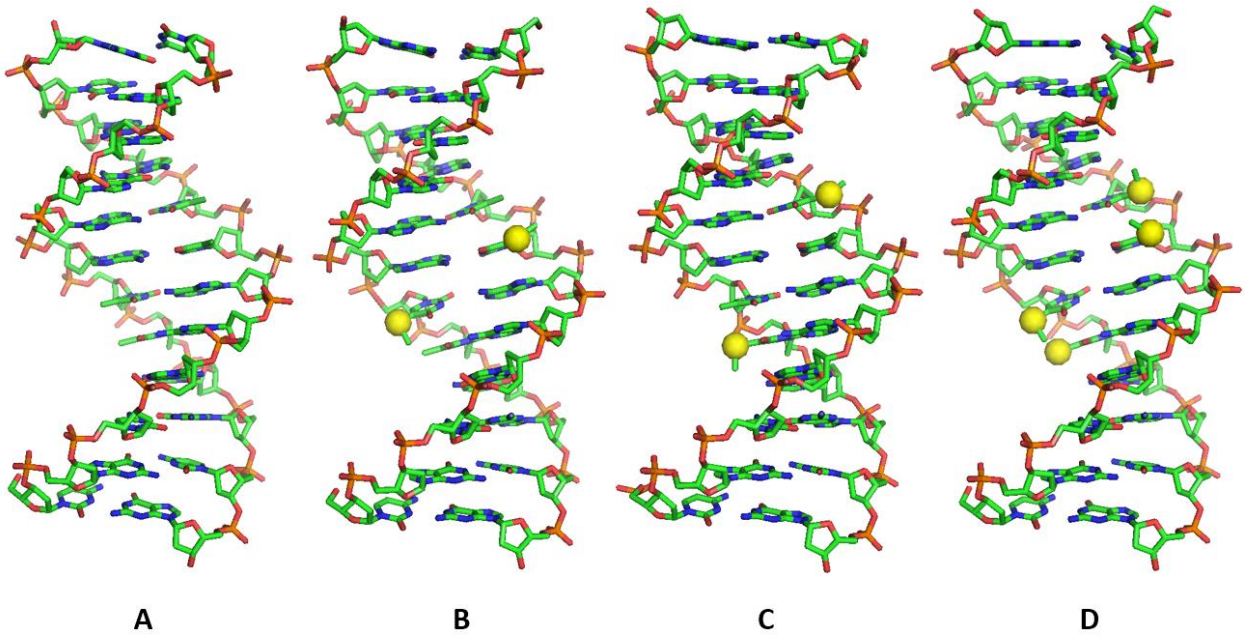


Figure 4.5 The structures of DNA duplexes in Bst complexes with Dickerson dodecamer DNA (A) DNA 4-1; (B) DNA 4-2; (C) DNA 4-3; (D) DNA 4-4.

Another interesting comparison we could do is to compare the structure of Dickerson dodecamer DNA when it crystallized by itself (PDB entry 3u2n; 1.25 Å resolution) and the structure of the same sequence in DNA/Bst DNA polymerase complex (**Figure 4.6**). Again, we did not find any significant differences between these two structures. We also enlarge the nucleotide in both structures (**Figure 4.6B and D**) and find that both sugar puckers are 2'-endo. This result indicates that Bst DNA polymerase binding will not change the conformation of Dickerson dodecamer DNA duplex.

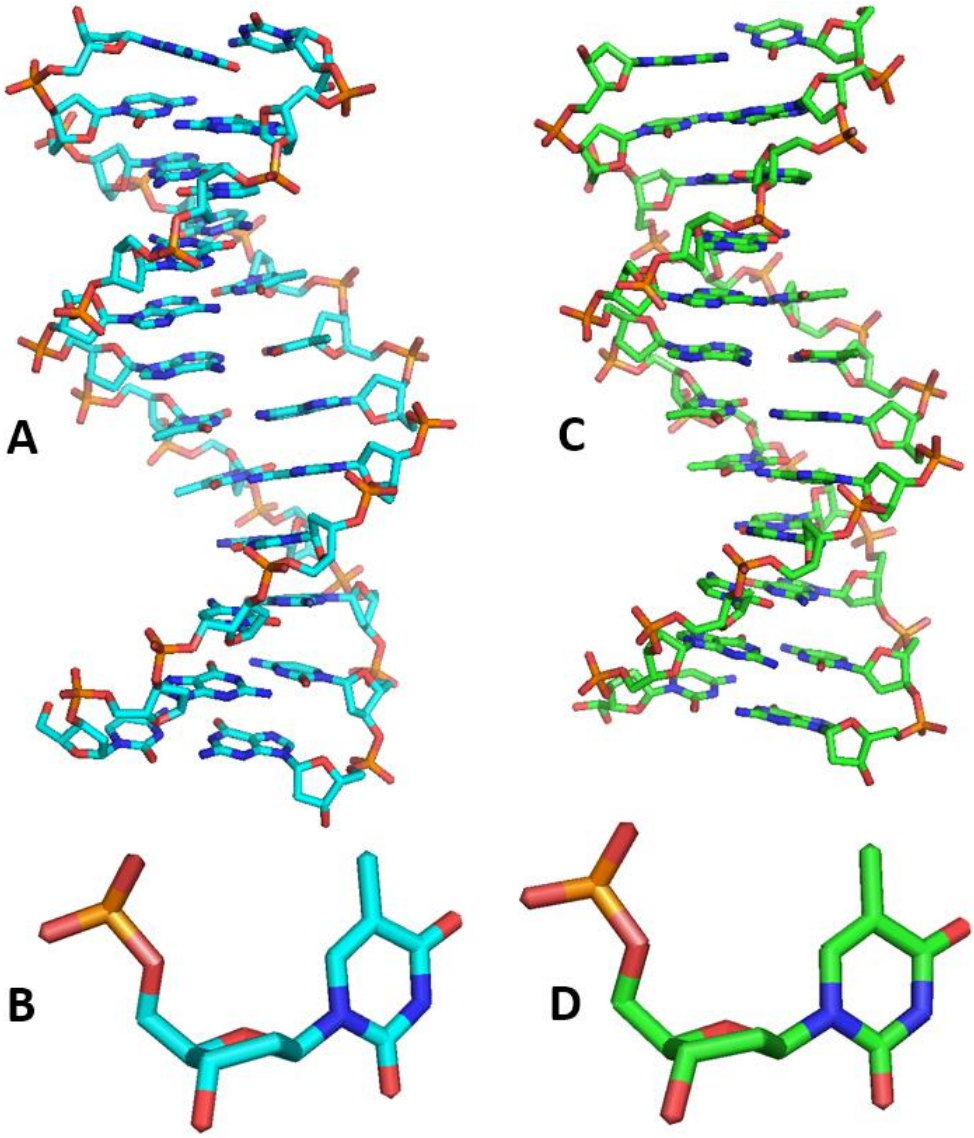


Figure 4.6 Comparation of the structure of Dickerson dodecamer DNA in Bst complex and crystallized by itself

(A) Structure of Dickerson dodecamer DNA in DNA/Bst DNA polymerase complex;
(B) Nucleotide conformation in Dickerson dodecamer DNA in Bst complex;
(C) Structure of Dickerson dodecamer DNA crystallized by itself;
(D) Nucleotide conformation in Dickerson dodecamer DNA crystallized by itself.

4.3.3.2 (5'-GTGTACAC-3')$_2$ DNA/Bst DNA polymerase complex structure study

When Bst DNA polymerase binding to the DNA duplex, the sugar pucker of nucleotide was found a C3' endo conformation, which is close to B-form DNA. However, some DNA

duplexes may prefer A-form conformation by its nature. We would like to know when Bst binding to this kind of DNA sequences, what conformation would the sugar pucker be and whether it will make any change if we introduce selenium atom to the A-form preferred DNA.

For this part, four DNAs shared with the same sequence (5'-GTGTACAC-3') but modified on different positions were used. This sequence was found A form conformation in crystals. The sequence and mass information are shown in **Table 4.3**.

Table 4.3 MALDI-TOF MASS analysis of (5'-GTGTACAC-3')$_2$ DNA

#	DNA Sequence	Calc. $[M-H^+]^-$	Observed
4-5	5'- GTGTACAC -3'	2410.6	-
4-6	5'- GTGTACAC -3'	2489.6	2489.5
4-7	5'- GTGTACAC -3'	2489.6	2489.6
4-8	5'- GTGTACAC -3'	2568.5	2569.5

Same crystallization condition was used for this (5'-GTGTACAC-3')$_2$ DNA/Bst DNA polymerase complex crystallization. The diffraction data of (5'-GTGTACAC-3')$_2$ DNA/Bst DNA polymerase complex crystal was collected at beamline 5.0.2 of the ALS (Advanced Light Source) at the Lawrence Berkeley National Laboratory. Several crystals were screened to find the crystal with the best diffraction quality. The diffraction data was collected at 1 Å wavelength for 1 second per frame with 0.25° oscillation angle. Data processing and structure refinement are all in the same way comparing with Dickerson dodecamer DNA/Bst DNA polymerase complex. Data collection and structure refinement statistics are given in **Table 4.4**.

Table 4.4 Data collection and refinement statistics of (5'-GTGTACAC-3')$_2$ DNA/Bst DNA polymerase complex

Complex structure	DNA4−5/Bst	DNA4−6/Bst	DNA4−7/Bst	DNA4−8/Bst
Wavelength (Å)	1			
Resolution range (Å)	50-2.26 (2.34-2.26)	50.0-2.64 (2.73-2.64)	50-2.38 (2.47-2.38)	50-2.32 (2.40-2.32)
Space group	$P2_12_12_1$			
Unit-cell a, b, c (Å)	87.7, 93.7, 106.1	87.7, 93.8, 106.0	86.9, 93.8, 106.8	87.5, 93.8, 106.0
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Unique reflections	41490(4064)	26585(2627)	35893(3514)	25165(2572)
Completeness (%)	99.7(99.1)	100 (100)	100 (100)	100 (100)
Rmerge (%)	10.8 (52.9)	13.2 (55.8)	10.6 (53.0)	10.8 (52.3)
I/σ(I)	27.4 (6.4)	21.8 (6.6)	31.4 (9.0)	27.7 (6.9)
R value (%)	18.31	18.17	18.24	17.87
Rfree (%)	22.87	23.91	23.60	23.10
Average B value (Å^2)	34.57	39.34	37.30	32.95
R.m.s.d. bond length (Å)	0.0185	0.0153	0.0173	0.0182
R.m.s.d. bond angle (°)	1.944	1.863	1.920	2.004

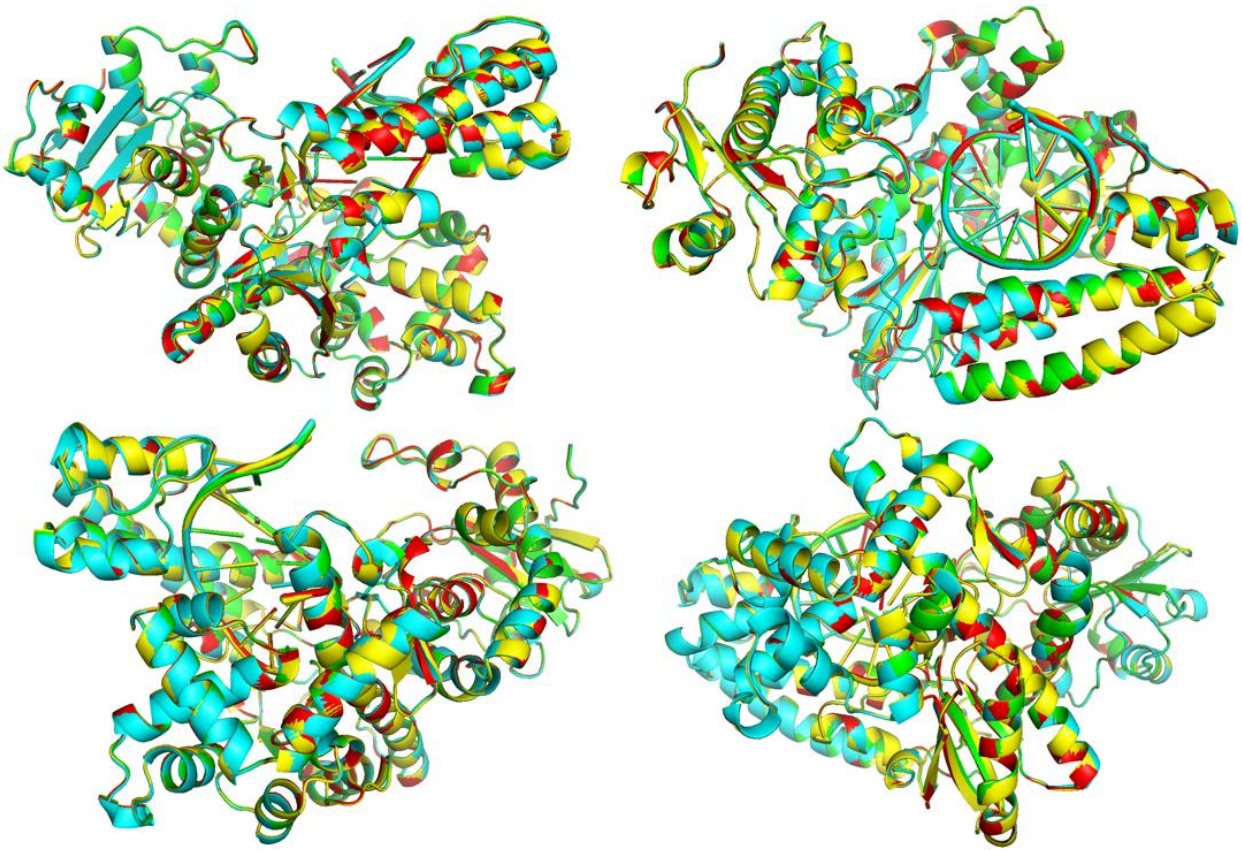


Figure 4.7 Different angle of overall structure of (5'-GTGTACAC-3')$_2$ DNA/Bst DNA polymerase complex
Complex with DNA 4-5 is green, DNA 4-6 is yellow, DNA 4-7 is cyan, DNA 4-8 is red.

The overall structures are shown in different angles in **Figure 4.7**. Similar to the Dickerson dodecamer sequences, no significant difference was found since all four structures are superimposed very well. And we think the reason for this consistency of those structures is the same, that the modifications are in the major groove while Bst DNA polymerase is binding to the minor groove of DNA duplexes. We use the Bst complex with most modified DNA 4-8 as an example to illustrate the relationship between the selenium modifications and DNA/Bst DNA polymerase binding interactions in **Figure 4.8**.

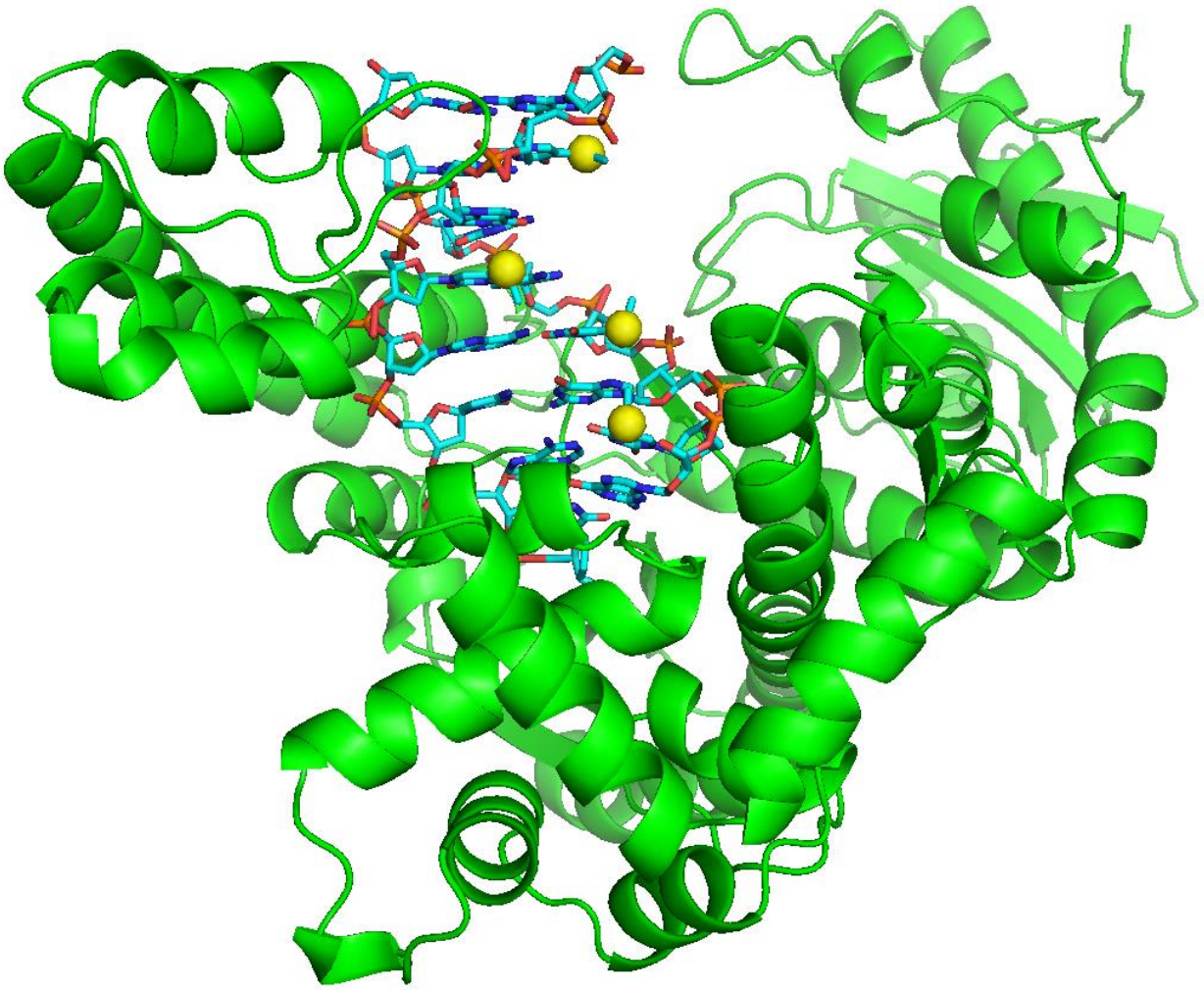


Figure 4.8 Structure of Bst complex with DNA 4-8
Yellow spheres represent selenium atoms.

Again, we would like to compare the (5'-GTGTACAC-3')$_2$ DNA duplexes in the structure of Bst complexes. The structures are displayed in **Figure 4.9**. This time the structures of DNA duplexes are even more similar to each other as this sequence is shorter than the sequence of Dickerson dodecamer and have less flexibility. As a conclusion based on this case and Dickerson dodecamer DNA/Bst complexes, we believe the 5-Se-T modifications will not perturb the structure of DNA duplexes on Bst DNA polymerase complexes regardless the sequence of the DNA.

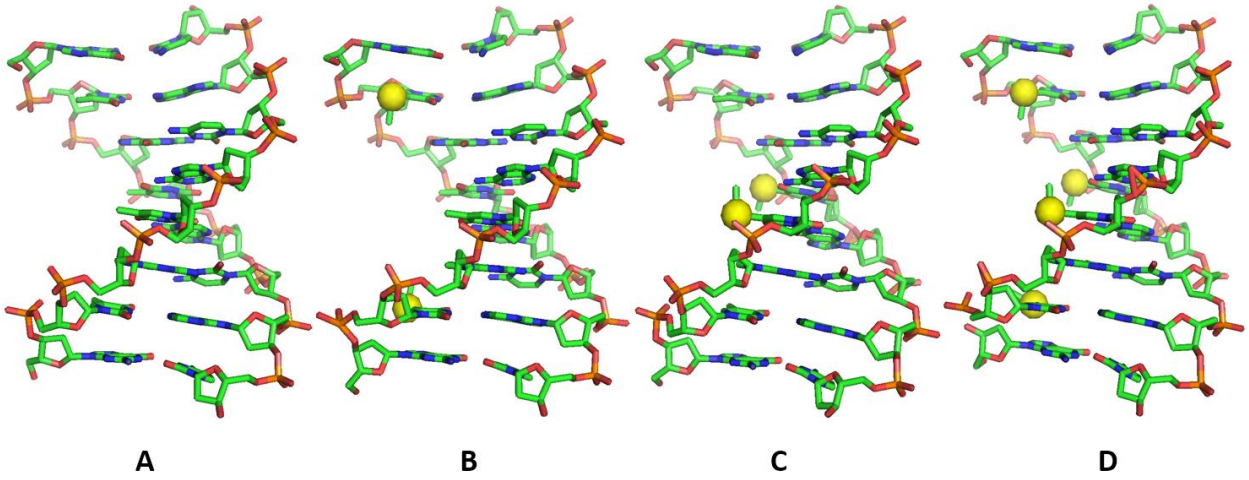


Figure 4.9 The structures of DNA duplexes in Bst complexes with (5'-GTGTACAC-3')$_2$ DNA
(A) DNA 4-1; (B) DNA 4-2; (C) DNA 4-3; (D) DNA 4-4.

We are also interested in whether the (5'-GTGTACAC-3')$_2$ DNA duplex structure changes when binding to Bst DNA polymerase. Therefore, we compare the DNA duplex in the complex with the structure of the same sequence by itself (PDB entry 1d79; 1.45 Å resolution) in **Figure 4.10**. It is obvious the structure of DNA duplex is completely different. We enlarged one of the nucleotides in the structure and found some interesting structure changes. Unlike the (5'-GTGTACAC-3')$_2$ DNA duplex crystalized by itself, the sugar puckers of nucleotides in Bst complex are 2'-endo instead of 3'-endo, which makes the structure is similar to B-form conformation. What's more, we find this conformation change in all the Bst DNA polymerase complexes with Se modified (5'-GTGTACAC-3')$_2$ DNA, which indicates that the Se-modifications do not even have impacts on the conformation shifting of the nucleic acids in nucleic acid-protein complex in this case.

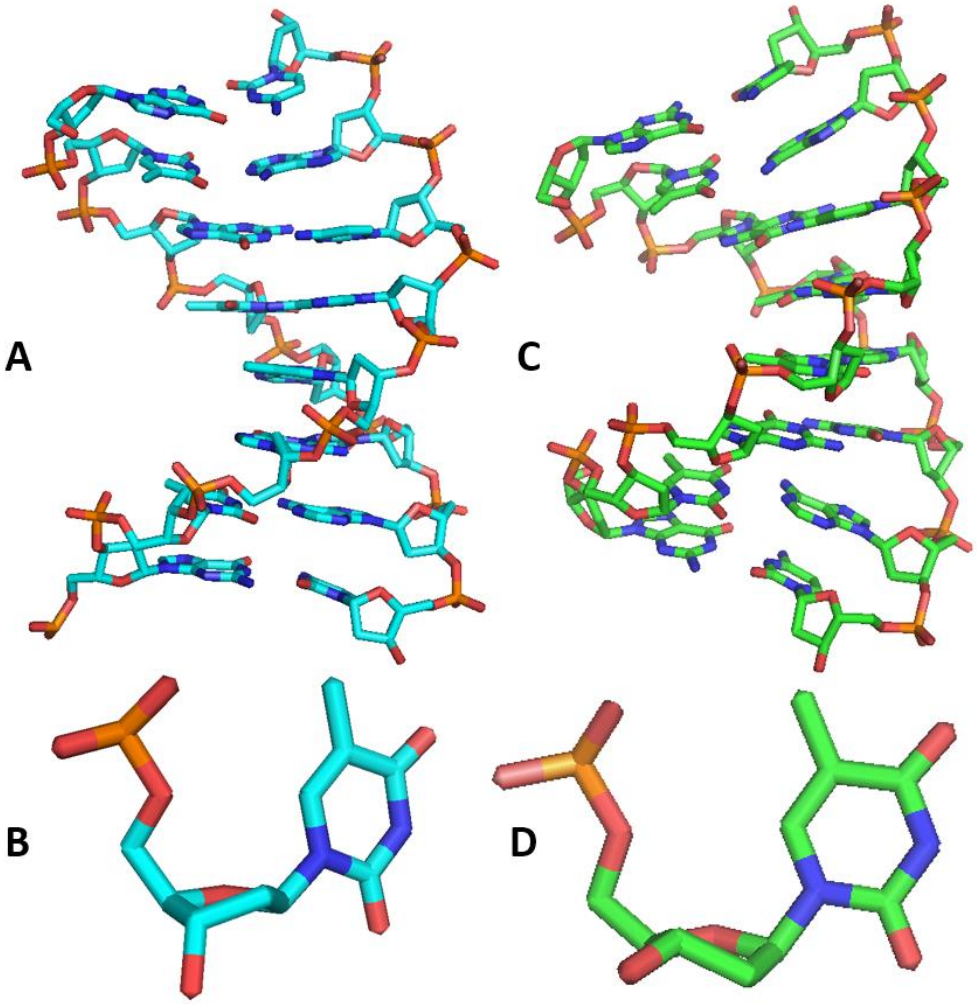


Figure 4.10 Comparation of (5'-GTGTACAC-3')$_2$ DNA in Bst complex and crystallized by itself

(A) Structure of (5'-GTGTACAC-3')$_2$ DNA in DNA/Bst DNA polymerase complex;
(B) Nucleotide conformation in (5'-GTGTACAC-3')$_2$ DNA in Bst complex;
(C) Structure of (5'-GTGTACAC-3')$_2$ DNA crystallized by itself;
(D) Nucleotide conformation in (5'-GTGTACAC-3')$_2$ DNA crystallized by itself.

However, the conformation of DNA duplex in crystals may not be exactly the same as it is in aqueous. Although we determine the DNA duplex conformation changes for (5'-GTGTACAC-3')$_2$ DNA in DNA/Bst DNA polymerase complex compared with the same sequence DNA crystallized by itself, it does not mean that Bst changes the conformation of DNA duplex. In this

way, we carried out CD spectra results to analysis the conformation of DNA duplexes in water solution. The results are displayed in **Figure 4.11**.

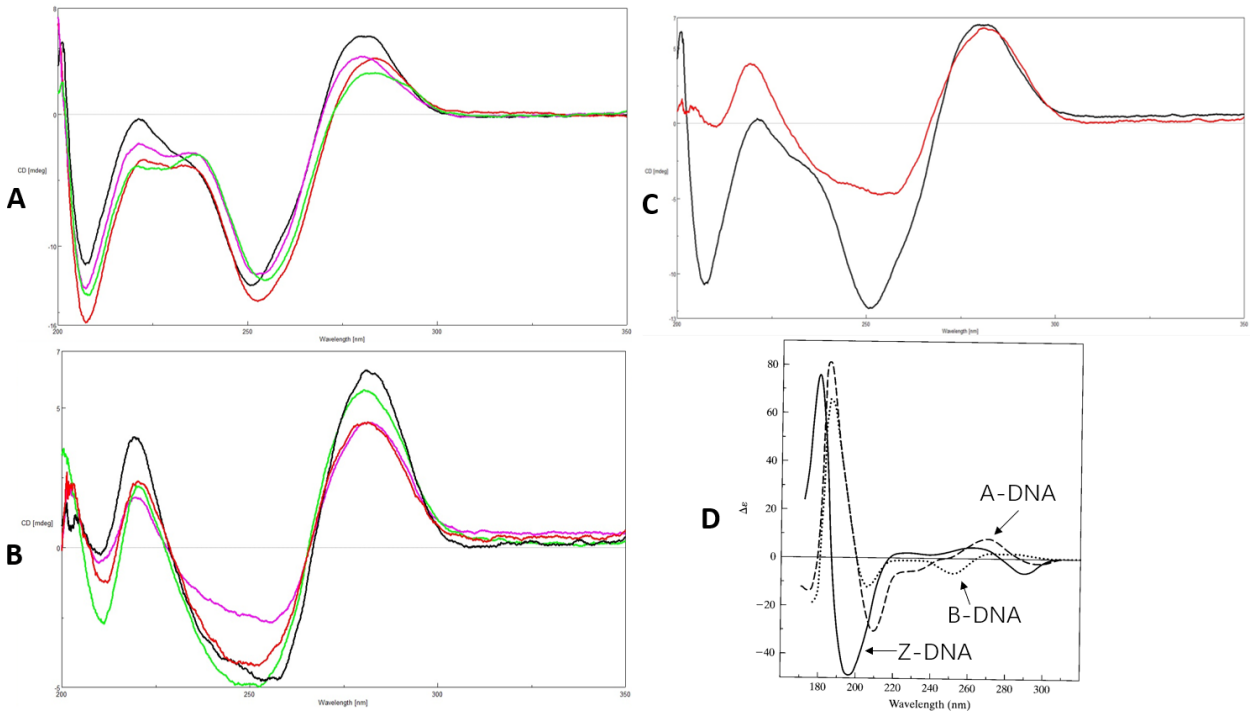


Figure 4.11 CD spectra for native and Se-modified DNAs

(A) CD spectra for native and Se-modified Dickerson dodecamer DNAs. DNA 4-1 is black, DNA 4-2 is red, DNA 4-3 is purple, and DNA 4-4 is green;

(B) CD spectra for native and Se-modified (5'-GTGTACAC-3')$_2$ DNAs. DNA 4-5 is black, DNA 4-6 is red, DNA 4-7 is purple, and DNA 4-8 is green;

(C) Overlapped CD spectra for native Dickerson dodecamer and (5'-GTGTACAC-3')$_2$ DNAs. DNA 4-1 is black, DNA 4-5 is red.

(D) CD spectra for different conformations of DNA duplex[100].

The CD result is very interesting. First of all, the spectra of native DNA and Se-modified DNA are very similar to each other despite different sequences including Dickerson dodecamer and (5'-GTGTACAC-3')$_2$. This result indicates that Se-modifications does not only keep the structure of DNA duplex in Bst/DNA polymerase complexes in crystals but also keep the conformation of DNA duplexes by themselves in aqueous. On the other hand, the CD spectra of

these two different sequences are also much alike each other despite some small differences. When comparing them with the CD spectra of different DNA duplex conformations in **Figure 4.11D**[100-101], both of the DNA duplexes look like B-form DNA in water solution with peak at about 275 nm and valley at about 250 nm, despite the fact that the DNA duplexes show different conformations in crystallization. This result could indicate these two different sequences of DNA duplex both prefer B-form conformation in aqueous instead of A-form conformation. In this way, Bst may only trap the B-form conformation DNA duplexes in aqueous and lock their conformation while packing to crystals instead of changing the conformation of $(5'\text{-GTGTACAC-}3')_2$ duplexes when binding to them. $(5'\text{-GTGTACAC-}3')_2$ duplexes show A-form conformation when crystallized by themselves because only A-form DNA pack to form crystals. When DNA duplexes were packing in crystals, the conformation of DNA is locked. In this way, some of the DNA duplexes in aqueous will change the conformation from B-form to A-form so that the balance of conformations in water solution could be kept. This is our explanation that why the conformation of $(5'\text{-GTGTACAC-}3')_2$ DNA duplexes is found A-form when crystallized by itself, while the conformation of the same sequence is found B-form in the crystal of DNA/Bst DNA polymerase complexes.

4.4 Conclusion

Se-modifications do not perturb the structure of the DNA duplex in Bst DNA polymerase complex. We co-crystallized Bst DNA polymerase with both B-form preferred Dickerson dodecamer and A-form preferred $(5'\text{-GTGTACAC-}3')_2$ DNAs with different Se-modifications and determined the structure of complexes. No significant structure perturbation was found between native DNA and Se-modified DNA in either Dickerson dodecamer or $(5'\text{-GTGTACAC-}3')_2$ DNA when binding to Bst DNA polymerase. On the other hand, when Bst DNA polymerase

binding to DNA duplex, it trapped the B-form DNA duplexes with sugar pucker 2'-endo conformation in aqueous and lock the conformation for crystal growth, even if the DNA duplex itself may pack to crystals in A-form duplexes and lock the 3'-endo conformations.

REFERENCES

1. Watson, J. D.; Crick, F. H., Molecular structure of nucleic acids. *Nature* **1953,** *171* (4356), 737-738.
2. Benoff, B.; Yang, H.; Lawson, C. L.; Parkinson, G.; Liu, J.; Blatter, E.; Ebright, Y. W.; Berman, H. M.; Ebright, R. H., Structural basis of transcription activation: the CAP-αCTD-DNA complex. *Science* **2002,** *297* (5586), 1562-1566.
3. Storz, G., An expanding universe of noncoding RNAs. *Science* **2002,** *296* (5571), 1260-1263.
4. Eddy, S. R., Non–coding RNA genes and the modern RNA world. *Nature Reviews Genetics* **2001,** *2* (12), 919.
5. Opalinska, J. B.; Gewirtz, A. M., Nucleic-acid therapeutics: basic principles and recent applications. *Nature Reviews Drug Discovery* **2002,** *1* (7), 503.
6. DiPaolo, J. A.; Alvarez-Salas, L. M., Advances in the development of therapeutic nucleic acids against cervical cancer. *Expert opinion on biological therapy* **2004,** *4* (8), 1251-1264.
7. Ng, E. W.; Shima, D. T.; Calias, P.; Cunningham Jr, E. T.; Guyer, D. R.; Adamis, A. P., Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease. *Nature reviews drug discovery* **2006,** *5* (2), 123.
8. Calin, G. A.; Sevignani, C.; Dumitru, C. D.; Hyslop, T.; Noch, E.; Yendamuri, S.; Shimizu, M.; Rattan, S.; Bullrich, F.; Negrini, M., Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proceedings of the National Academy of Sciences* **2004,** *101* (9), 2999-3004.
9. Egli, M., Nucleic acid crystallography: current progress. *Current opinion in chemical biology* **2004,** *8* (6), 580-591.
10. Berman, H. M.; Gelbin, A.; Westbrook, J., Nucleic acid crystallography: A view from the nucleic acid database. *Progress in biophysics and molecular biology* **1996,** *66* (3), 255-288.
11. Egli, M.; Pallan, P. S., Insights from crystallographic studies into the structural and pairing properties of nucleic acid analogs and chemically modified DNA and RNA oligonucleotides. *Annu. Rev. Biophys. Biomol. Struct.* **2007,** *36*, 281-305.
12. Ferré-D'Amaré, A. R.; Zhou, K.; Doudna, J. A., Crystal structure of a hepatitis delta virus ribozyme. *Nature* **1998,** *395* (6702), 567.
13. Stein, C. A.; Castanotto, D., FDA-approved oligonucleotide therapies in 2017. *Molecular Therapy* **2017,** *25* (5), 1069-1075.
14. Lin, L.; Sheng, J.; Huang, Z., Nucleic acid X-ray crystallography via direct selenium derivatization. *Chemical Society Reviews* **2011,** *40* (9), 4591-4602.
15. Campbell, N. H.; Parkinson, G. N., Crystallographic studies of quadruplex nucleic acids. *Methods* **2007,** *43* (4), 252-263.
16. Hassan, A. E.; Sheng, J.; Zhang, W.; Huang, Z., High fidelity of base pairing by 2-selenothymidine in DNA. *Journal of the American Chemical Society* **2010,** *132* (7), 2120-2121.
17. Sheng, J.; Salon, J.; Gan, J.; Huang, Z., Synthesis and crystal structure study of 2′ -Se-adenosine-derivatized DNA. *Science China Chemistry* **2010,** *53* (1), 78-85.
18. Hassan, A. E.; Sheng, J.; Jiang, J.; Zhang, W.; Huang, Z., Synthesis and crystallographic analysis of 5-Se-thymidine DNAs. *Organic letters* **2009,** *11* (12), 2503-2506.
19. Jiang, J.; Sheng, J.; Carrasco, N.; Huang, Z., Selenium derivatization of nucleic acids for crystallography. *Nucleic acids research* **2006,** *35* (2), 477-485.

20. Moroder, H.; Kreutz, C.; Lang, K.; Serganov, A.; Micura, R., Synthesis, oxidation behavior, crystallization and structure of 2 ‘-methylseleno guanosine containing RNAs. *Journal of the American Chemical Society* **2006,** *128* (30), 9909-9918.
21. Salon, J.; Sheng, J.; Gan, J.; Huang, Z., Synthesis and crystal structure of 2′ -Se-modified guanosine containing DNA. *The Journal of organic chemistry* **2010,** *75* (3), 637-641.
22. Salon, J.; Sheng, J.; Jiang, J.; Chen, G.; Caton-Williams, J.; Huang, Z., Oxygen replacement with selenium at the thymidine 4-position for the Se base pairing and crystal structure studies. *Journal of the American Chemical Society* **2007,** *129* (16), 4862-4863.
23. Sheng, J.; Jiang, J.; Salon, J.; Huang, Z., Synthesis of a 2 ‘-Se-thymidine Phosphoramidite and Its Incorporation into Oligonucleotides for Crystal Structure Study. *Organic letters* **2007,** *9* (5), 749-752.
24. Wilds, C. J.; Pattanayek, R.; Pan, C.; Wawrzak, Z.; Egli, M., Selenium-assisted nucleic acid crystallography: use of phosphoroselenoates for MAD phasing of a DNA structure. *Journal of the American Chemical Society* **2002,** *124* (50), 14910-14916.
25. Teplova, M.; Wilds, C. J.; Wawrzak, Z.; Tereshko, V.; Du, Q.; Carrasco, N.; Huang, Z.; Egli, M., Covalent incorporation of selenium into oligonucleotides for X-ray crystal structure determination via MAD: proof of principle. *Biochimie* **2002,** *84* (9), 849-858.
26. Salon, J.; Gan, J.; Abdur, R.; Liu, H.; Huang, Z., Synthesis of 6-Se-guanosine RNAs for structural study. *Organic letters* **2013,** *15* (15), 3934-3937.
27. Abdur, R.; Gerlits, O. O.; Gan, J.; Jiang, J.; Salon, J.; Kovalevsky, A. Y.; Chumanevich, A. A.; Weber, I. T.; Huang, Z., Novel complex MAD phasing and RNase H structural insights using selenium oligonucleotides. *Acta Crystallographica Section D: Biological Crystallography* **2014,** *70* (2), 354-361.
28. Thompson, R. A.; Spring, A. M.; Sheng, J.; Huang, Z.; Germann, M. W., The importance of fitting in: conformational preference of selenium 2′ modifications in nucleosides and helical structures. *Journal of Biomolecular Structure and Dynamics* **2015,** *33* (2), 289-297.
29. Höbartner, C.; Micura, R., Chemical synthesis of selenium-modified oligoribonucleotides and their enzymatic ligation leading to an U6 SnRNA stem− loop segment. *Journal of the American Chemical Society* **2004,** *126* (4), 1141-1149.
30. Höbartner, C.; Rieder, R.; Kreutz, C.; Puffer, B.; Lang, K.; Polonskaia, A.; Serganov, A.; Micura, R., Syntheses of RNAs with up to 100 Nucleotides Containing Site-Specific 2 ‘-Methylseleno Labels for Use in X-ray Crystallography. *Journal of the American Chemical Society* **2005,** *127* (34), 12035-12045.
31. Serganov, A.; Keiper, S.; Malinina, L.; Tereshko, V.; Skripkin, E.; Höbartner, C.; Polonskaia, A.; Phan, A. T.; Wombacher, R.; Micura, R., Structural basis for Diels-Alder ribozyme-catalyzed carbon-carbon bond formation. *Nature Structural and Molecular Biology* **2005,** *12* (3), 218.
32. Freisz, S.; Lang, K.; Micura, R.; Dumas, P.; Ennifar, E., Binding of aminoglycoside antibiotics to the duplex form of the HIV‐1 genomic RNA dimerization initiation site. *Angewandte Chemie International Edition* **2008,** *47* (22), 4110-4113.
33. Haeberli, P.; Berger, I.; Pallan, P. S.; Egli, M., Syntheses of 4′ -thioribonucleosides and thermodynamic stability and crystal structure of RNA oligomers with incorporated 4′ -thiocytosine. *Nucleic acids research* **2005,** *33* (13), 3965-3975.
34. Zhang, W.; Huang, Z., Synthesis of the 5′ -se-thymidine phosphoramidite and convenient labeling of DNA oligonucleotide. *Organic letters* **2011,** *13* (8), 2000-2003.

35. Carrasco, N.; Huang, Z., Enzymatic Synthesis of Phosphoroselenoate DNA Using Thymidine 5 ‘-(α-P-seleno) triphosphate and DNA Polymerase for X-ray Crystallography via MAD. *Journal of the American Chemical Society* **2004,** *126* (2), 448-449.
36. Carrasco, N.; Caton‐Williams, J.; Brandt, G.; Wang, S.; Huang, Z., Efficient enzymatic synthesis of phosphoroselenoate RNA by using adenosine 5′‐(α‐P‐seleno) triphosphate. *Angewandte Chemie International Edition* **2006,** *45* (1), 94-97.
37. Brandt, G.; Carrasco, N.; Huang, Z., Efficient substrate cleavage catalyzed by hammerhead ribozymes derivatized with selenium for X-ray crystallography. *Biochemistry* **2006,** *45* (29), 8972-8977.
38. Egli, M.; Pallan, P. S.; Pattanayek, R.; Wilds, C. J.; Lubini, P.; Minasov, G.; Dobler, M.; Leumann, C. J.; Eschenmoser, A., Crystal structure of homo-DNA and nature's choice of pentose over hexose in the genetic system. *Journal of the American Chemical Society* **2006,** *128* (33), 10847-10856.
39. Salon, J.; Jiang, J.; Sheng, J.; Gerlits, O. O.; Huang, Z., Derivatization of DNAs with selenium at 6-position of guanine for function and crystal structure studies. *Nucleic acids research* **2008,** *36* (22), 7009-7018.
40. Caton‐Williams, J.; Huang, Z., Synthesis and DNA‐polymerase incorporation of colored 4‐selenothymidine triphosphate for polymerase recognition and DNA visualization. *Angewandte Chemie International Edition* **2008,** *47* (9), 1723-1725.
41. Sharma, V. K.; Sharma, R. K.; Singh, S. K., Antisense oligonucleotides: modifications and clinical trials. *MedChemComm* **2014,** *5* (10), 1454-1471.
42. Lin, L.; Sheng, J.; Momin, R. K.; Du, Q.; Huang, Z., Facile synthesis and anti-tumor cell activity of Se-containing nucleosides. *Nucleosides, Nucleotides and Nucleic Acids* **2009,** *28* (1), 56-66.
43. Zamecnik, P. C.; Stephenson, M. L., Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxynucleotide. *Proceedings of the National Academy of Sciences* **1978,** *75* (1), 280-284.
44. Robinson, R., RNAi therapeutics: how likely, how soon? *PLoS biology* **2004,** *2* (1), e28.
45. Marwick, C., First antisense drug will treat CMV retinitis. *Jama* **1998,** *280* (10), 871-871.
46. Morrison, C., Alnylam prepares to land first RNAi drug approval. *Nature Reviews Drug Discovery* **2018,** *17*, 156-157.
47. Xie, X.; Liang, J.; Pu, T.; Xu, F.; Yao, F.; Yang, Y.; Zhao, Y.-L.; You, D.; Zhou, X.; Deng, Z., Phosphorothioate DNA as an antioxidant in bacteria. *Nucleic acids research* **2012,** *40* (18), 9115-9124.
48. Rahman, S. A.; Baba, T.; Kodama, T.; Islam, M. A.; Obika, S., Hybridizing ability and nuclease resistance profile of backbone modified cationic phosphorothioate oligonucleotides. *Bioorganic & medicinal chemistry* **2012,** *20* (13), 4098-4102.
49. De Clercq, E.; Eckstein, F.; Merigan, T., Interferon induction increased through chemical modification of a synthetic polyribonucleotide. *Science* **1969,** *165* (3898), 1137-1139.
50. Crooke, S. T.; Bennett, C. F., Progress in antisense oligonucleotide therapeutics. *Annual Review of Pharmacology and Toxicology* **1996,** *36* (1), 107-129.
51. Prakash, T. P., An overview of sugar‐modified oligonucleotides for antisense therapeutics. *Chemistry & biodiversity* **2011,** *8* (9), 1616-1641.
52. Prakash, T. P.; Bhat, B., 2'-Modified oligonucleotides for antisense therapeutics. *Current topics in medicinal chemistry* **2007,** *7* (7), 641-649.

53. Obika, S.; Morio, K.-i.; Nanbu, D.; Hari, Y.; Itoh, H.; Imanishi, T., Synthesis and conformation of 3′, 4′-BNA monomers, 3′-O, 4′-C-methyleneribonucleosides. *Tetrahedron* **2002,** *58* (15), 3039-3049.
54. Obika, S.; Nanbu, D.; Hari, Y.; Morio, K.-i.; In, Y.; Ishida, T.; Imanishi, T., Synthesis of 2′-O, 4′-C-methyleneuridine and-cytidine. Novel bicyclic nucleosides having a fixed C 3,-endo sugar puckering. *Tetrahedron Letters* **1997,** *38* (50), 8735-8738.
55. Koshkin, A. A.; Fensholdt, J.; Pfundheller, H. M.; Lomholt, C., A simplified and efficient route to 2 '-O, 4 '-C-methylene-linked bicyclic ribonucleosides (locked nucleic acid). *The Journal of organic chemistry* **2001,** *66* (25), 8504-8512.
56. Watts, J. K., Locked nucleic acid: tighter is different. *Chemical Communications* **2013,** *49* (50), 5618-5620.
57. Obika, S.; Andoh, J.-i.; Sugimoto, T.; Miyashita, K.; Imanishi, T., Synthesis of a conformationally locked AZT analogue, 3′-azido-3′-deoxy-2′-O, 4′-C-methylene-5-methyluridine. *Tetrahedron letters* **1999,** *40* (35), 6465-6468.
58. Steffens, R.; Leumann, C. J., Synthesis and thermodynamic and biophysical properties of tricyclo-DNA. *Journal of the American Chemical Society* **1999,** *121* (14), 3249-3255.
59. Guschlbauer, W.; Jankowski, K., Nucleoside conformation is determined by the electronegativity of the sugar substituent. *Nucleic Acids Research* **1980,** *8* (6), 1421-1433.
60. Aboul-Fadl, T., Antisense oligonucleotides: the state of the art. *Current medicinal chemistry* **2005,** *12* (19), 2193-2214.
61. Herdewijn, P., Heterocyclic modifications of oligonucleotides and antisense technology. *Antisense and Nucleic Acid Drug Development* **2000,** *10* (4), 297-310.
62. Peacock, H.; Kannan, A.; Beal, P. A.; Burrows, C. J., Chemical modification of siRNA bases to probe and enhance RNA interference. *The Journal of organic chemistry* **2011,** *76* (18), 7295-7300.
63. Verma, S.; Eckstein, F., Modified oligonucleotides: synthesis and strategy for users. Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA: 1998.
64. Sun, H.; Jiang, S.; Caton-Williams, J.; Liu, H.; Huang, Z., 2-Selenouridine triphosphate synthesis and Se-RNA transcription. *Rna* **2013,** *19* (9), 1309-1314.
65. Christofferson, A.; Zhao, L.; Sun, H.; Huang, Z.; Huang, N., Theoretical studies of the base pair fidelity of selenium-modified DNA. *The Journal of Physical Chemistry B* **2011,** *115* (33), 10041-10048.
66. Zhang, W., Structure and Function Studies of Selenium Substituted Nucleic Acids. **2012**.
67. Nowotny, M.; Yang, W., Stepwise analyses of metal ions in RNase H catalysis from substrate destabilization to product release. *The EMBO journal* **2006,** *25* (9), 1924-1933.
68. Nemoto, N.; Miyamoto-Sato, E.; Yanagawa, H., Fluorescence labeling of the C‐terminus of proteins with a puromycin analogue in cell‐free translation systems. *FEBS letters* **1999,** *462* (1-2), 43-46.
69. Tyagi, S., Imaging intracellular RNA distribution and dynamics in living cells. *natuRe methods* **2009,** *6* (5), 331.
70. Chen, T.; Wu, C. S.; Jimenez, E.; Zhu, Z.; Dajac, J. G.; You, M.; Han, D.; Zhang, X.; Tan, W., DNA micelle flares for intracellular mRNA imaging and gene therapy. *Angewandte Chemie* **2013,** *125* (7), 2066-2070.
71. Paige, J. S.; Wu, K. Y.; Jaffrey, S. R., RNA mimics of green fluorescent protein. *Science* **2011,** *333* (6042), 642-646.

72. Franzini, R. M.; Kool, E. T., Efficient nucleic acid detection by templated reductive quencher release. *Journal of the American Chemical Society* **2009,** *131* (44), 16021-16023.
73. Gonçalves, M. S. T., Fluorescent labeling of biomolecules with organic probes. *Chemical reviews* **2008,** *109* (1), 190-212.
74. Perlette, J.; Tan, W., Real-time monitoring of intracellular mRNA hybridization inside single living cells. *Analytical chemistry* **2001,** *73* (22), 5544-5550.
75. Wu, Y.; Yang, C. J.; Moroz, L. L.; Tan, W., Nucleic acid beacons for long-term real-time intracellular monitoring. *Analytical chemistry* **2008,** *80* (8), 3025-3028.
76. Papadopoulou, A.; Green, R. J.; Frazier, R. A., Interaction of flavonoids with bovine serum albumin: a fluorescence quenching study. *Journal of agricultural and food chemistry* **2005,** *53* (1), 158-163.
77. Sitepu, I.; Ignatia, L.; Franz, A.; Wong, D.; Faulina, S.; Tsui, M.; Kanti, A.; Boundy-Mills, K., An improved high-throughput Nile red fluorescence assay for estimating intracellular lipids in a variety of yeast species. *Journal of microbiological methods* **2012,** *91* (2), 321-328.
78. Paunesku, T.; Vogt, S.; Maser, J.; Lai, B.; Woloschak, G., X‐ray fluorescence microprobe imaging in biology and medicine. *Journal of cellular biochemistry* **2006,** *99* (6), 1489-1502.
79. Sakurai, K.; Eba, H., Micro X-ray fluorescence imaging without scans: Toward an element-selective movie. *Analytical chemistry* **2003,** *75* (2), 355-359.
80. Abnet, C. C.; Lai, B.; Qiao, Y.-L.; Vogt, S.; Luo, X.-M.; Taylor, P. R.; Dong, Z.-W.; Mark, S. D.; Dawsey, S. M., Zinc concentration in esophageal biopsy specimens measured by x-ray fluorescence and esophageal cancer risk. *Journal of the National Cancer Institute* **2005,** *97* (4), 301-306.
81. Daly, M. J.; Gaidamakova, E. K.; Matrosova, V. Y.; Vasilenko, A.; Zhai, M.; Leapman, R. D.; Lai, B.; Ravel, B.; Li, S.-M. W.; Kemner, K. M., Protein oxidation implicated as the primary determinant of bacterial radioresistance. *PLoS biology* **2007,** *5* (4), e92.
82. Chen, J.; Marks, E.; Lai, B.; Zhang, Z.; Duce, J. A.; Lam, L. Q.; Volitakis, I.; Bush, A. I.; Hersch, S.; Fox, J. H., Iron accumulates in Huntington's disease neurons: protection by deferoxamine. *PloS one* **2013,** *8* (10), e77023.
83. Yang, L.; McRae, R.; Henary, M. M.; Patel, R.; Lai, B.; Vogt, S.; Fahrni, C. J., Imaging of the intracellular topography of copper with a fluorescent sensor and by synchrotron x-ray fluorescence microscopy. *Proceedings of the National Academy of Sciences* **2005,** *102* (32), 11179-11184.
84. Reith, F.; Etschmann, B.; Grosse, C.; Moors, H.; Benotmane, M. A.; Monsieurs, P.; Grass, G.; Doonan, C.; Vogt, S.; Lai, B., Mechanisms of gold biomineralization in the bacterium Cupriavidus metallidurans. *Proceedings of the National Academy of Sciences* **2009**, pnas. 0904583106.
85. Kashiv, Y.; Austin II, J. R.; Lai, B.; Rose, V.; Vogt, S.; El-Muayed, M., Imaging trace element distributions in single organelles and subcellular features. *Scientific reports* **2016,** *6*, 21437.
86. Rayman, M. P., The importance of selenium to human health. *The lancet* **2000,** *356* (9225), 233-241.
87. MacDonald, R. S., The role of zinc in growth and cell proliferation. *The Journal of nutrition* **2000,** *130* (5), 1500S-1508S.
88. Kornberg, A., Biologic synthesis of deoxyribonucleic acid. *Science* **1960,** *131* (3412), 1503-1508.
89. Kornberg, A.; Baker, T. A., *DNA replication*. Wh Freeman New York: 1992; Vol. 3.

90. Klenow, H.; Henningsen, I., Selective elimination of the exonuclease activity of the deoxyribonucleic acid polymerase from Escherichia coli B by limited proteolysis. *Proceedings of the National Academy of Sciences* **1970,** *65* (1), 168-175.
91. Ollis, D.; Brick, P.; Hamlin, R.; Xuong, N.; Steitz, T. A., Structure of large fragment of Escherichia coli DNA polymerase I complexed with dTMP. *nature* **1985,** *313* (6005), 762.
92. 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. *The EMBO journal* **1991,** *10* (1), 25-33.
93. Delarue, M.; Poch, O.; Tordo, N.; Moras, D.; Argos, P., An attempt to unify the structure of polymerases. *Protein Engineering, Design and Selection* **1990,** *3* (6), 461-467.
94. Lawyer, F. C.; Stoffel, S.; Saiki, R. K.; Myambo, K.; Drummond, R.; Gelfand, D. H., Isolation, characterization, and expression in Escherichia coli of the DNA polymerase gene from Thermus aquaticus. *Journal of Biological Chemistry* **1989,** *264* (11), 6427-6437.
95. Stenesh, J.; Roe, B., DNA polymerase from mesophilic and thermophilic bacteria: I. Purification and properties of DNA polymerase from Bacillus licheniformis and Bacillus stearothermophilus. *Biochimica et Biophysica Acta (BBA)-Nucleic Acids and Protein Synthesis* **1972,** *272* (2), 156-166.
96. Kim, Y.; Eom, S. H.; Wang, J.; Lee, D.-S.; Suh, S. W.; Steitz, T. A., Crystal structure of Thermus aquaticus DNA polymerase. *Nature* **1995,** *376* (6541), 612.
97. Korolev, S.; Nayal, M.; Barnes, W. M.; Di Cera, E.; Waksman, G., Crystal structure of the large fragment of Thermus aquaticus DNA polymerase I at 2.5-A resolution: structural basis for thermostability. *Proceedings of the National Academy of Sciences* **1995,** *92* (20), 9264-9268.
98. Kiefer, J. R.; Mao, C.; Hansen, C. J.; Basehore, S. L.; Hogrefe, H. H.; Braman, J. C.; Beese, L. S., Crystal structure of a thermostable Bacillus DNA polymerase l large fragment at 2.1 Å resolution. *Structure* **1997,** *5* (1), 95-108.
99. Kiefer, J. R.; Mao, C.; Braman, J. C.; Beese, L. S., Visualizing DNA replication in a catalytically active Bacillus DNA polymerase crystal. *Nature* **1998,** *391* (6664), 304.
100. Riazance, J. H.; Baase, W. A.; Johnson Jr, W. C.; Hall, K.; Cruz, P.; Tinoco Jr, I., Evidence for Z-form RNA by vacuum UV circular dichroism. *Nucleic acids research* **1985,** *13* (13), 4983-4989.
101. Holm, A. I. S.; Nielsen, L. M.; Hoffmann, S. V.; Nielsen, S. B., Vacuum-ultraviolet circular dichroism spectroscopy of DNA: a valuable tool to elucidate topology and electronic coupling in DNA. *Physical Chemistry Chemical Physics* **2010,** *12* (33), 9581-9596.

APPENDIX: TABLES AND FIGURES

The Classics Suite (Qiagen) Screen Kit Conditions

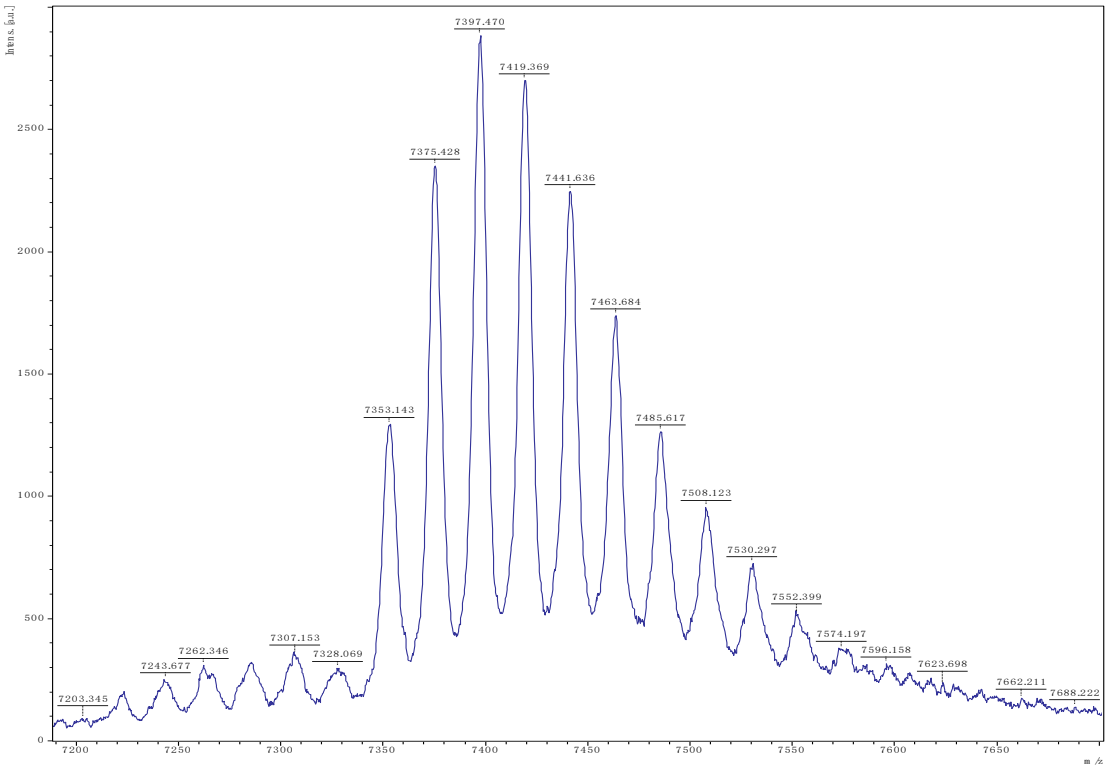
#	Salt	Buffer	Precipitant
1	0.01 M Cobalt chloride	0.1 M Sodium acetate pH 4.6	1.0 M 1,6-Hexanediol
2		0.1 M tri-Sodium acetate pH 5.6	2.5 M 1,6-Hexanediol
3	0.2 M Magnesium chloride	0.1 M Tris pH 8.5	3.4 M 1,6-Hexanediol
4			5% (v/v) Isopropanol; 2.0 M Ammonium sulfate
5		0.1 M HEPES sodium salt pH 7.5	10% (v/v) Isopropanol; 20% (w/v) PEG 4000
6	0.2 M Calcium chloride	0.1 M sodium acetate pH 4.6	20% (v/v) Isopropanol
7		0.1M tri-Sodium citrate pH 5.6	20% (v/v) Isopropanol; 20% (w/v) PEG 4000
8	0.2 M tri-Sodium citrate	0.1 M HEPES sodium salt pH 7.5	20% (v/v) Isopropanol
9	0.2 M tri-Sodium citrate	0.1 M sodium cacodylate pH 6.5	30% (v/v) Isopropanol
10	0.2 M Magnesium chloride	0.1 M HEPES sodium salt pH 7.5	30% (v/v) Isopropanol
11	0.2 M Ammonium acetate	0.1 M Tris-HCl pH 8.5	30% (v/v) Isopropanol
12			10% (v/v) Ethanol; 1.5 M Sodium chloride
13		0.1 M Tris pH 8.5	20% (v/v) Ethanol
14			25% (v/v) Ethylene glycol
15	0.02 M Calcium chloride	0.1 M Sodium acetate pH 4.6	30% (v/v) MPD
16	0.2 M Sodium chloride	0.1 M Sodium acetate pH 4.6	30% (v/v) MPD
17	0.2 M Ammonium acetate	0.1 M tri-Sodium acetate pH 5.6	30% (v/v) MPD
18	0.2 M Magnesium acetate	0.1 M Sodium cacodylate pH 6.5	30% (v/v) MPD
19	0.2 M tri-Sodium citrate	0.1 M HEPES sodium salt pH 7.5	30% (v/v) MPD
20	0.5 M Ammonium sulfate	0.1 M HEPES pH 7.5	30% (v/v) MPD
21	0.2 M Ammonium phosphate	0.1 M Tris pH 8.5	50% (v/v) MPD
22		0.1 M HEPES pH 7.5	70% (v/v) MPD
23		0.1 M Tris pH 8.5	25% (v/v) tert-Butanol
24		0.1 M tri-Sodium citrate pH 5.6	35% (v/v) tert-Butanol
25			0.4 M Ammonium phosphate

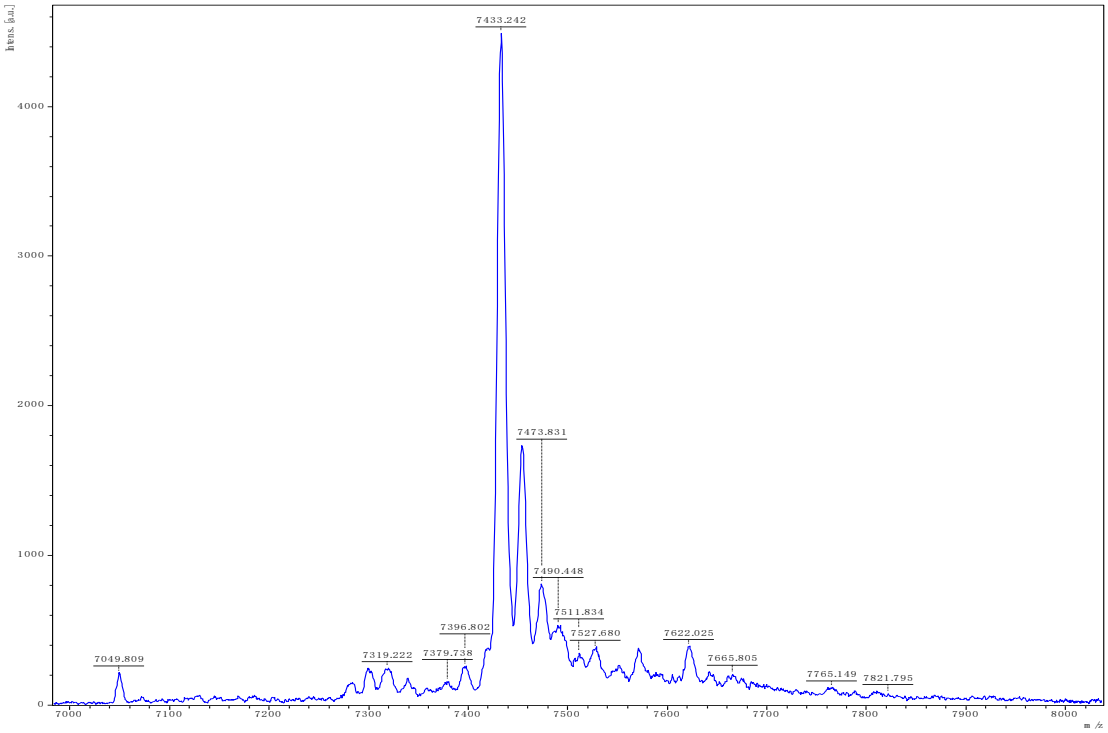
26		0.1 M tri-Sodium citrate pH 5.6	1.0 M Ammonium phosphate
27		0.1M Tris·HCl pH 8.5	2.0 M Ammonium phosphate
28		0.1 M HEPES pH 7.5	2.0 M Ammonium formate
29		0.1 M Sodium acetate pH 4.6	2.0 M Ammonium sulfate
30		0.1M Tris·HCl pH 8.5	2.0 M Ammonium sulfate
31			2.0 M Ammonium sulfate
32	0.1 M Sodium chloride	0.1 M HEPES pH 7.5	1.6 M Ammonium sulfate
33	0.01 M Cobalt chloride	0.1 M MES pH 6.5	1.8 M Ammonium sulfate
34	0.2 M K/Na tartate	0.1 M tri-Sodium citrate pH 5.6	2.0 M Ammonium sulfate
35			1.0 M Imidazole pH 7.0
36			0.4 M K/Na tartate
37		0.1 M HEPES sodium salt pH 7.5	0.8 M K/Na tartate
38		0.1 M Imidazole pH 6.5	1.0 M Sodium acetate
39	0.05 M Cadmium sulfate	0.1 M HEPES pH 7.5	1.0 M Sodium acetate
40		0.1 M Sodium cacodylate pH 5.6	1.4 M Sodium acetate
41		0.1 M Sodium acetate pH 4.6	2.0 M Sodium chloride
42	0.1 M Sodium phosphate; 0.1 M Potassium phosphate	0.1 M MES pH 6.5	2.0 M Sodium chloride
43		0.1 M HEPES pH 7.5	4.3 M Sodium chloride
44		0.1 M HEPES sodium salt pH 7.5	1.4 M tri-Sodium citrate
45			1.6 M tri-Sodium citrate pH 6.5
46		0.1 M HEPES sodium salt pH 7.5	0.8 M Sodium phosphate 0.8 M Potassium phosphate
47		0.1 M Sodium acetate pH 4.6	2.0 M Sodium formate
48			4.0 M Sodium formate
49		0.1 M Bicine pH 9.0	2% (v/v) Dioxane; 10% (w/v) PEG 20000
50		0.1 M MES pH 6.5	10% (v/v) Dioxane; 1.6 M Ammonium sulfate
51			35% (v/v) Dioxane
52	0.5 M Sodium chloride	0.1 M tri-Sodium citrate pH 5.6	2% (v/v) Ethylene imine polymer
53		0.1 M Tris pH 8.5	12% (v/v) Glycerol; 1.5 M Ammonium sulfate
54	0.5 M Sodium chloride; 0.01 M Magnesium chloride	0.01 M CTAB	

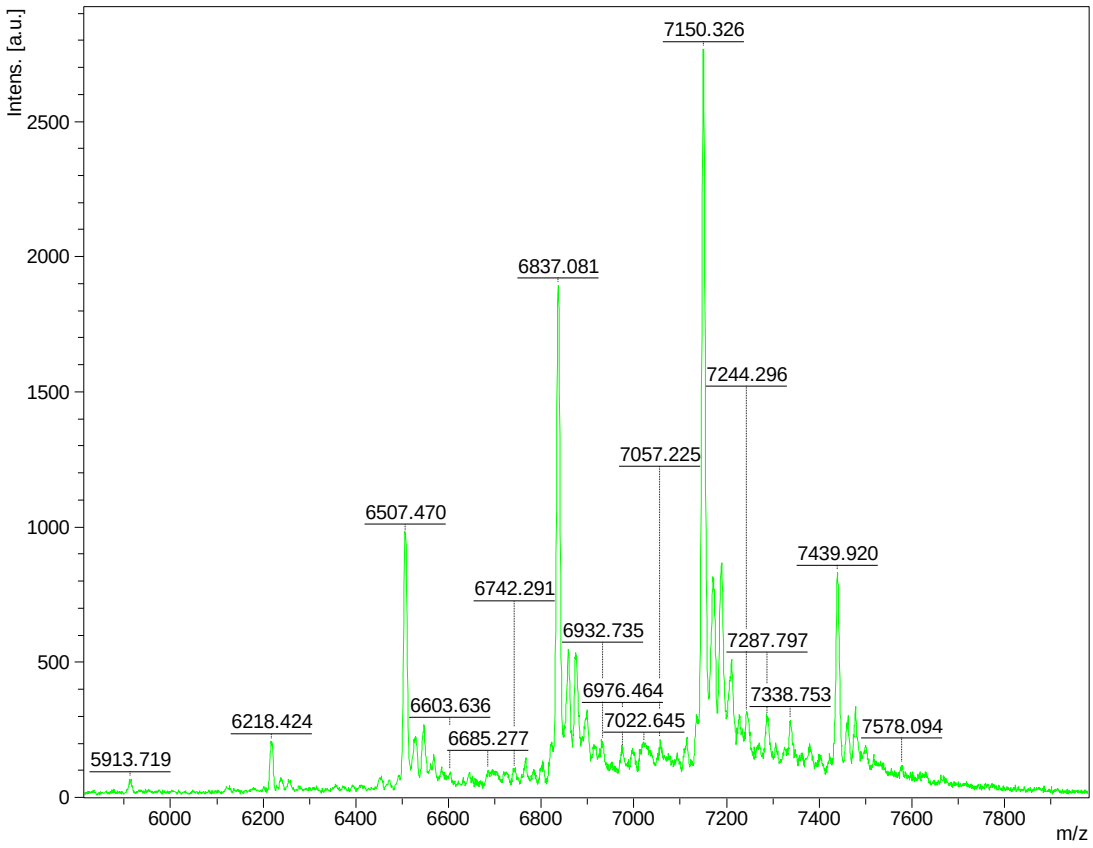
55	0.01 M Ferric chloride	0.1 M tri-Sodium citrate pH 5.6	10% (v/v) Jeffamine M-600
56		0.1 M HEPES pH 7.5	20% (v/v) Jeffamine M-600
57	0.5 M Ammonium sulfate	0.1 M tri-Sodium citrate pH 5.6	1.0 M Lithium sulfate
58	0.01 M Nickel chloride	0.1 M Tris pH 8.5	1.0 M Lithium sulfate
59		0.1 M HEPES sodium salt pH 7.5	1.5 M Lithium sulfate
60		0.1 M Bicine pH 9.0	2.0 M Magnesium chloride
61			0.2 M Magnesium formate
62		0.1 M MES pH 6.5	1.6 M Magnesium sulfate
63		0.1 M Tris·HCl pH 8.5	8% (w/v) PEG 8000
64		0.1 M HEPES pH 7.5	10% (w/v) PEG 8000
65	0.5 M Lithium sulfate		15% (w/v) PEG 8000
66	0.2 M Zinc acetate	0.1 M Sodium cacodylate pH 6.5	18% (w/v) PEG 8000
67	0.2 M Calcium acetate	0.1 M Sodium cacodylate pH 6.5	18% (w/v) PEG 8000
68	0.2 M Magnesium acetate	0.1 M Sodium cacodylate pH 6.5	20% (w/v) PEG 8000
69	0.05 M Potassium phosphate		20% (w/v) PEG 8000
70	0.2 M Ammonium sulfate	0.1 M Sodium cacodylate pH 6.5	30% (w/v) PEG 8000
71	0.2 M Sodium acetate	0.1 M Sodium cacodylate pH 6.5	30% (w/v) PEG 8000
72	0.2 M Ammonium sulfate		30% (w/v) PEG 8000
73		0.1 M HEPES sodium salt pH 7.5	2% (v/v) PEG 400; 2.0 M Ammonium sulfate
74	0.2 M Calcium chloride	0.1 M HEPES sodium salt pH 7.5	28% (v/v) PEG 400
75	0.1 M Cadmium chloride	0.1 M Sodium acetate pH 4.6	30% (v/v) PEG 400
76	0.2 m Magnesium chloride	0.1 M HEPES sodium salt pH 7.5	30% (v/v) PEG 400
77	0.2 M tri-Sodium citrate	0.1 M Tris·HCl pH 8.5	30% (v/v) PEG 400
78	0.1 M Sodium chloride	0.1 M Bicine pH 9.0	20% (w/v) PEG 550 MME
79	0.01 M Zinc sulfate	0.1 M MES pH 6.5	25% (w/v) PEG 550 MME
80			10% (w/v) PEG 1000; 10% (w/v) PEG 8000
81			30% (w/v) PEG 1500
82	0.01 M Nickel chloride	0.1 M Tris pH 8.5	20% (w/v) PEG 2000 MME
83	0.2 M Ammonium sulfate	0.1 M Sodium acetate pH 4.6	30% (w/v) PEG 2000 MME
84		0.1 M Sodium acetate pH 4.6	8% (w/v) PEG 4000

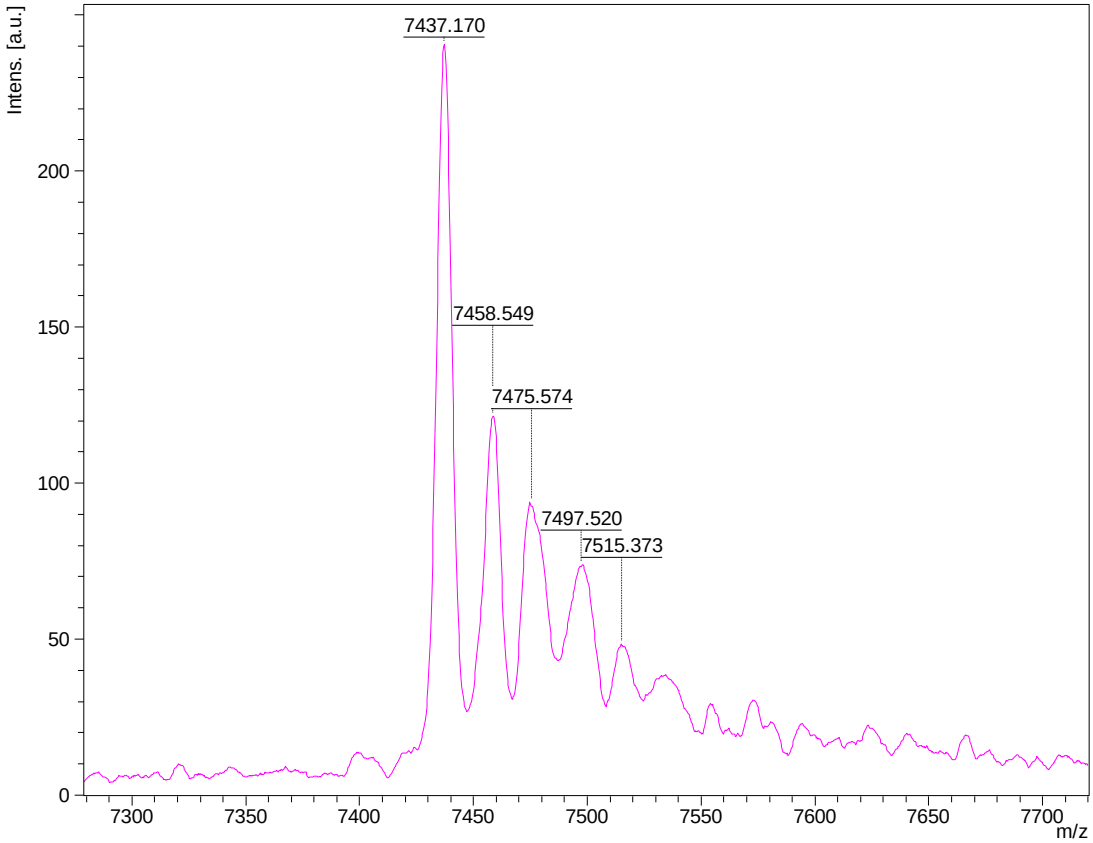
85	0.2 M Ammonium sulfate	0.1 M Sodium acetate pH 4.6	25% (w/v) PEG 4000
86	0.2 M Ammonium acetate	0.1 M Sodium acetate pH 4.6	30% (w/v) PEG 4000
87	0.2 M Ammonium acetate	0.1 M tri-Sodium citrate pH 5.6	30% (w/v) PEG 4000
88	0.2 M Magnesium chloride	0.1 M Tris·HCl pH 8.5	30% (w/v) PEG 4000
89	0.2 M Lithium sulfate	0.1 M Tris·HCl pH 8.5	30% (w/v) PEG 4000
90	0.2 M Sodium acetate	0.1 M Tris·HCl pH 8.5	30% (w/v) PEG 4000
91	0.2 M Ammonium sulfate		30% (w/v) PEG 4000
92	0.2 M Ammonium sulfate	0.1 M MES pH 6.5	30% (w/v) PEG 5000 MME
93		0.1 M HEPES pH 7.5	10% (w/v) PEG 6000; 5% (v/v) MPD
94			10% (w/v) PEG 6000; 2.0 M Sodium chloride
95		0.1 M HEPES pH 7.5	20% (w/v) PEG 10000; 8% (v/v) Ethylene glycol
96		0.1 M MES pH 6.5	12% (w/v) PEG 20000

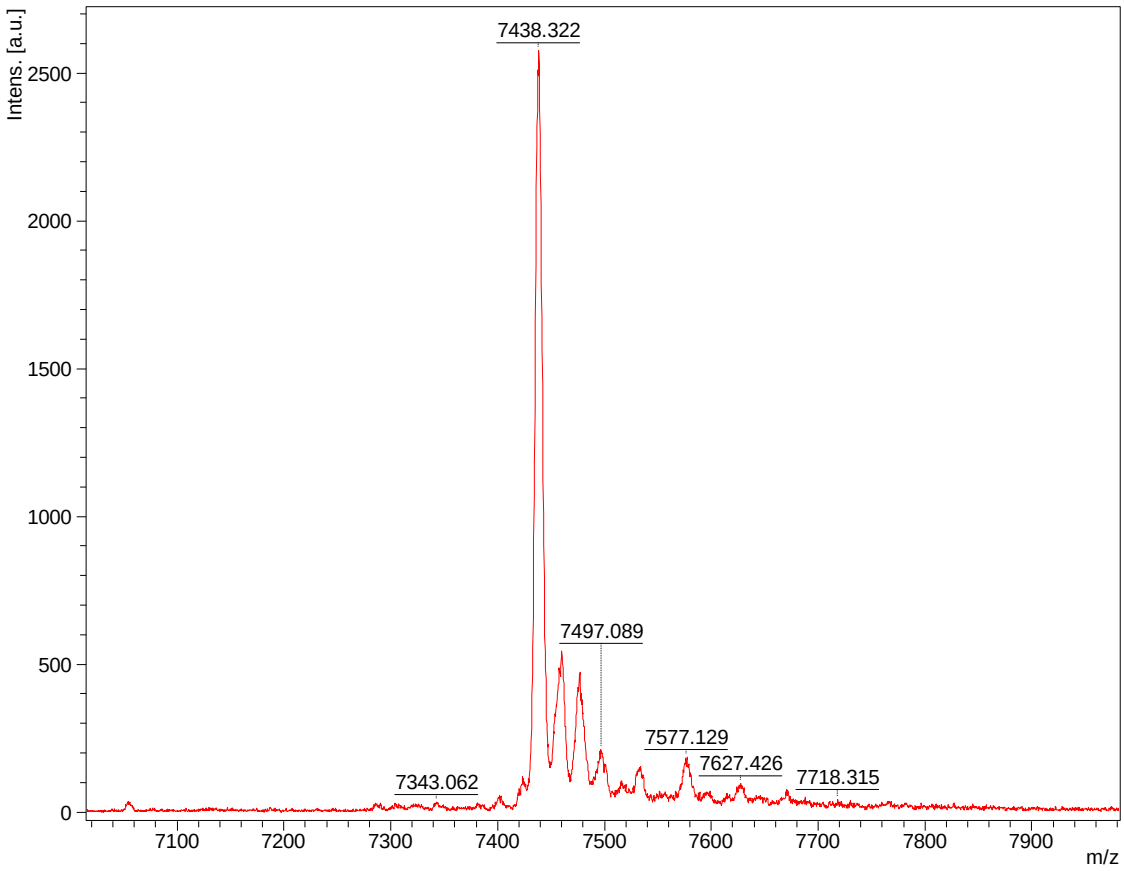
Mass Spectrum for Important Sequences

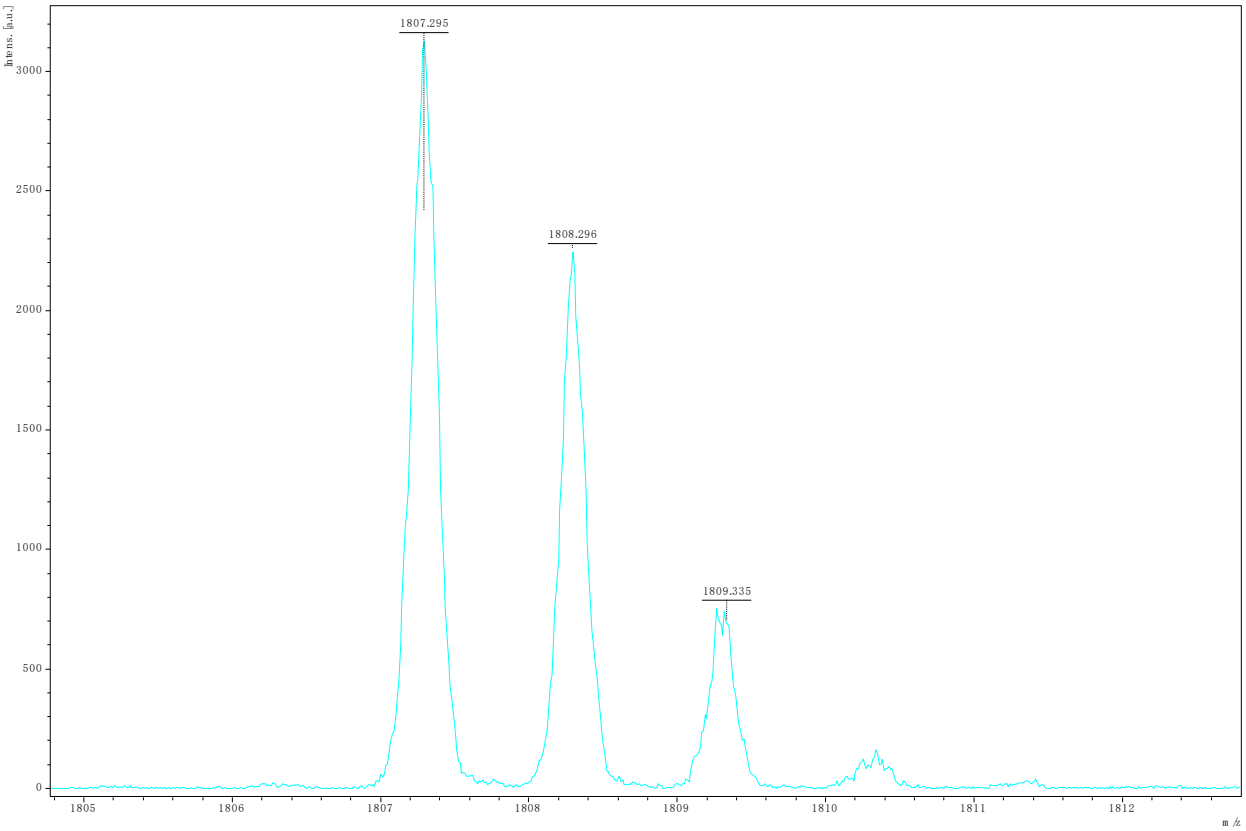


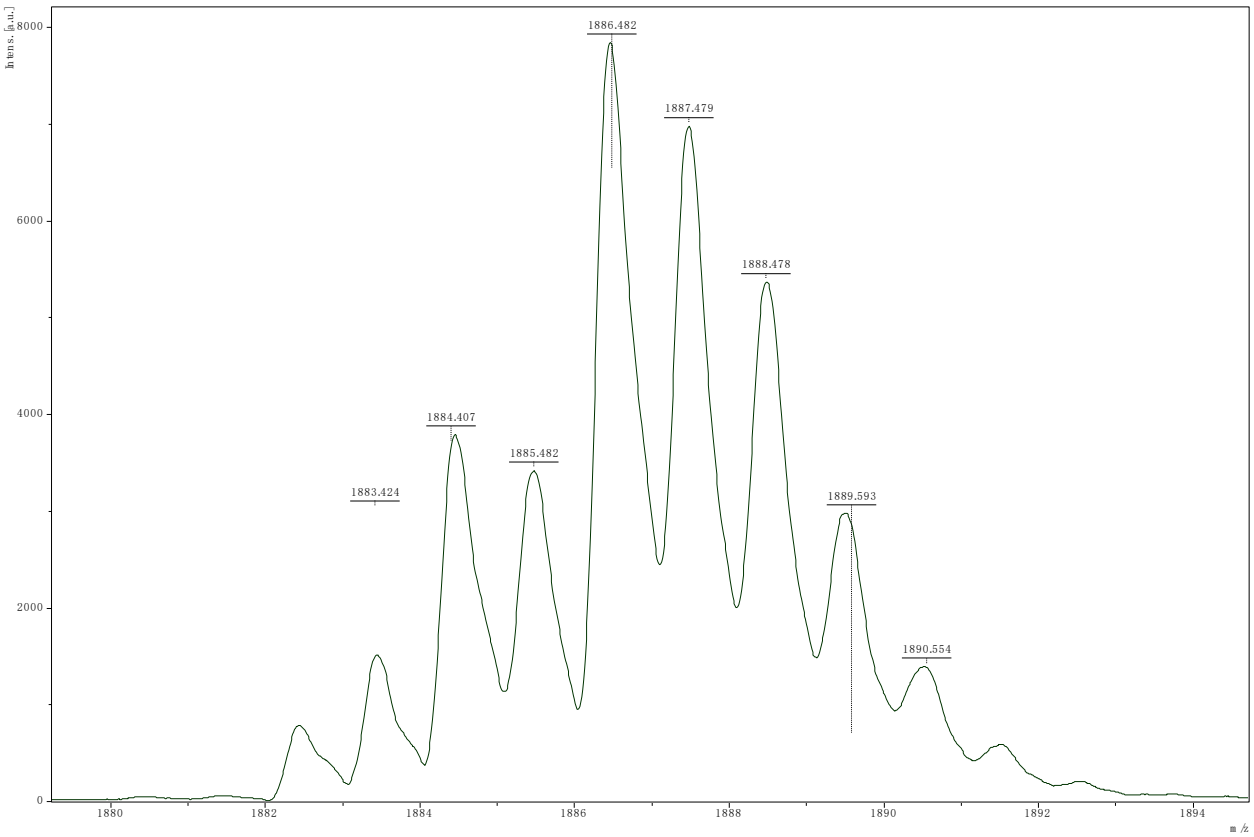


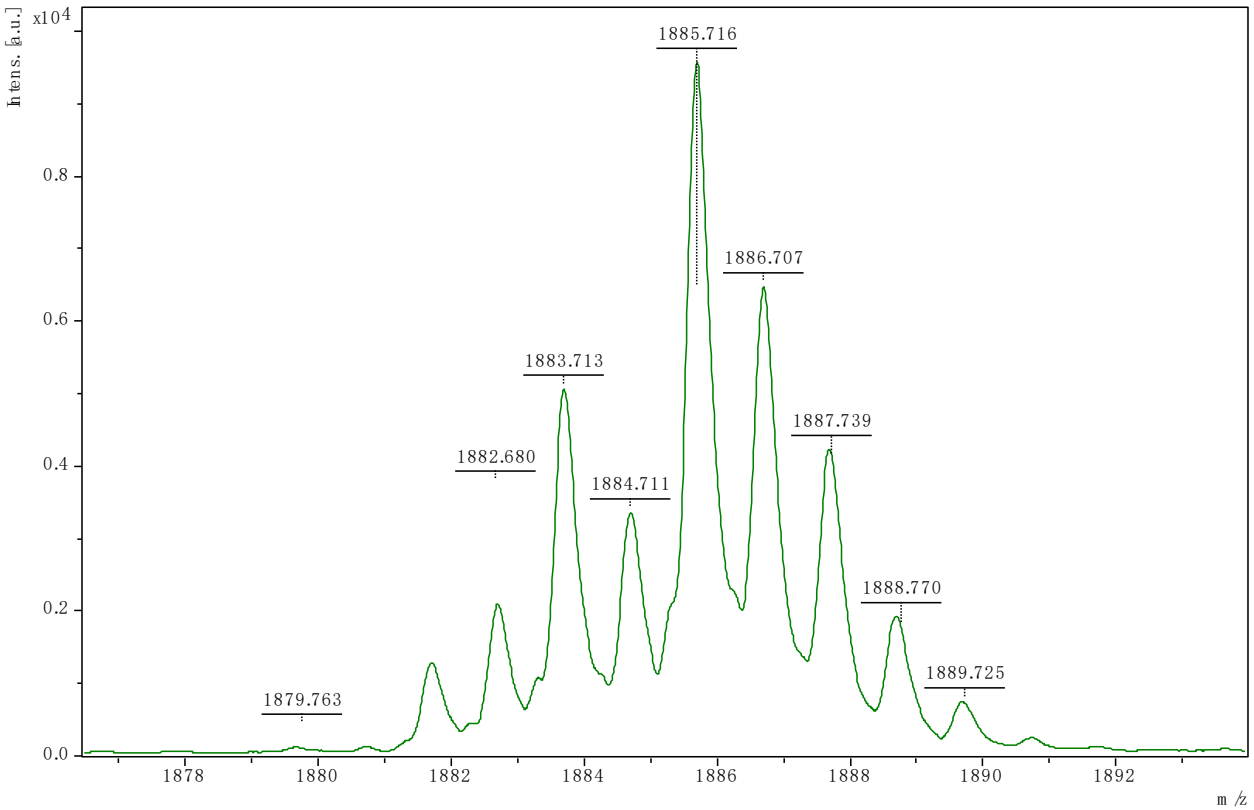

Intens. [a.u.]
2500
2000
1500
1000
500
0
7150.326
6837.081
7244.296
7057.225
6507.470
7439.920
6742.291
6932.735
7287.797
6976.464
7338.753
6603.636
6218.424
7022.645
7578.094
6685.277
5913.719
6000
6200
6400
6600
6800
7000
7200
7400
7600
7800
m/z

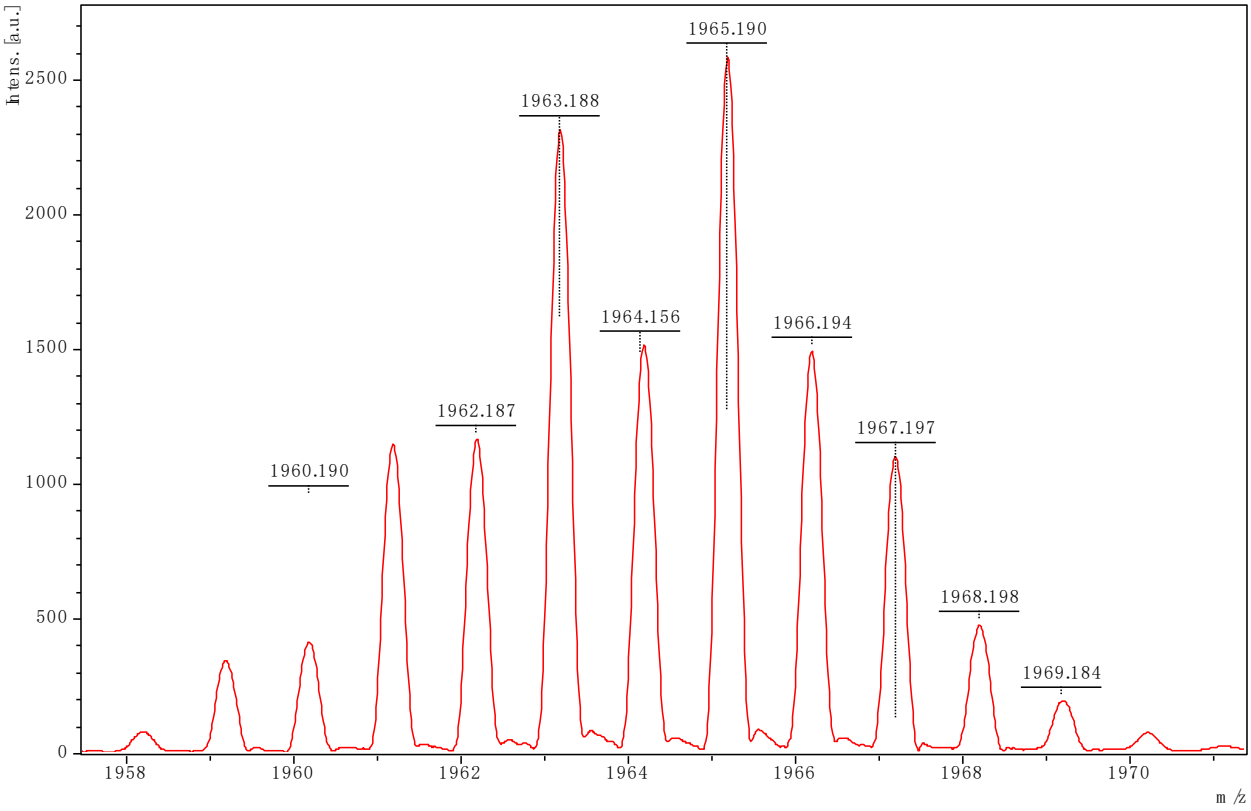

Intens. [a.u.]
200
150
100
50
0
7437.170
7458.549
7475.574
7497.520
7515.373
7300
7350
7400
7450
7500
7550
7600
7650
7700
m/z

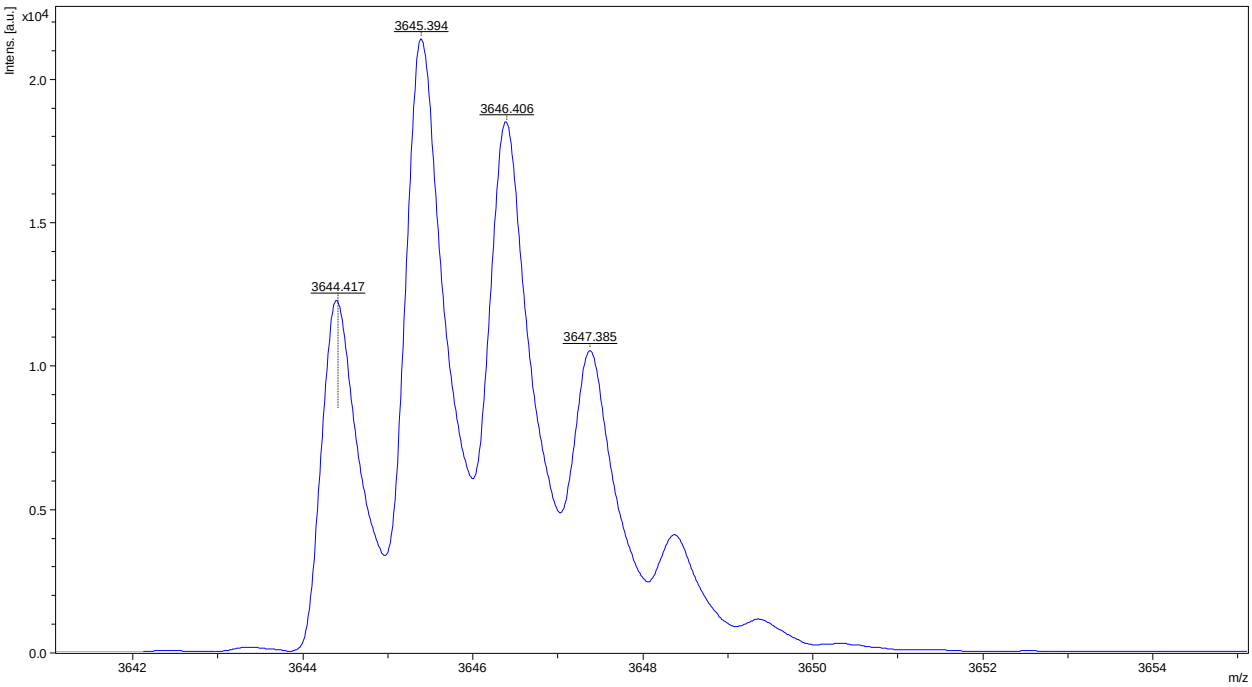

Intens. [a.u.]
2500
2000
1500
1000
500
0
7438.322
7497.089
7577.129
7627.426
7718.315
7343.062
7100
7200
7300
7400
7500
7600
7700
7800
7900
m/z

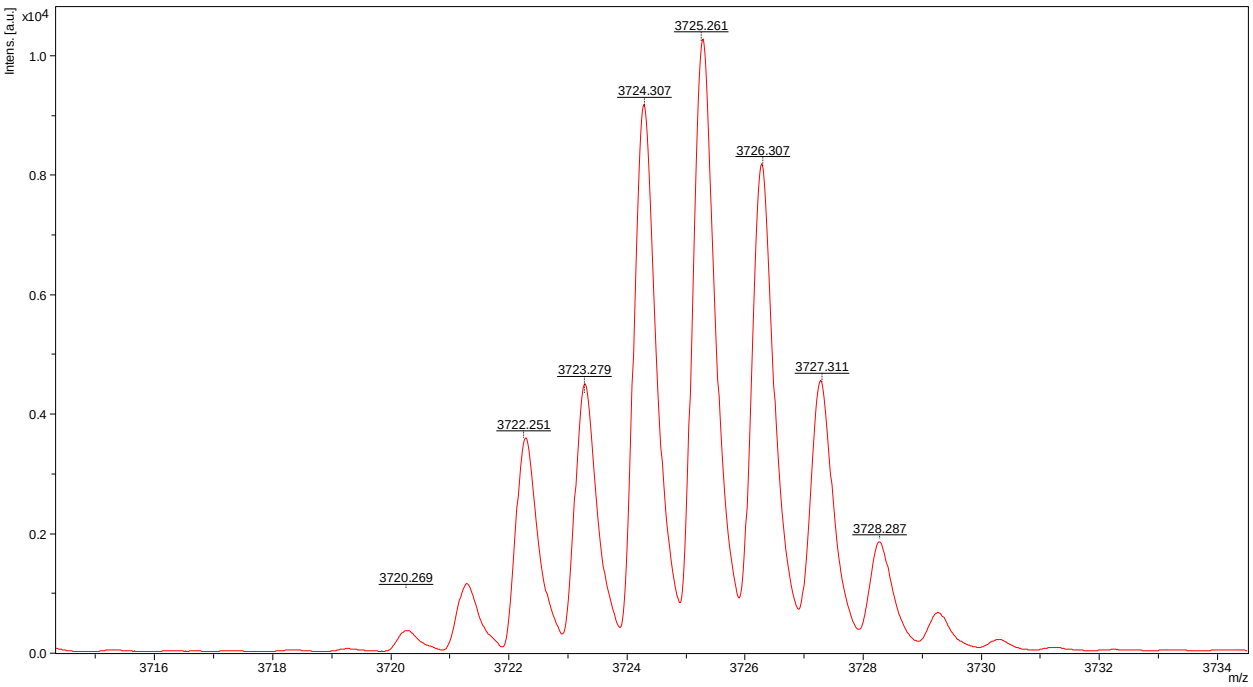

Intens. [a.u.]
3000
2500
2000
1500
1000
500
0
1807.295
1808.296
1809.335
1805
1806
1807
1808
1809
1810
1811
1812
m/z

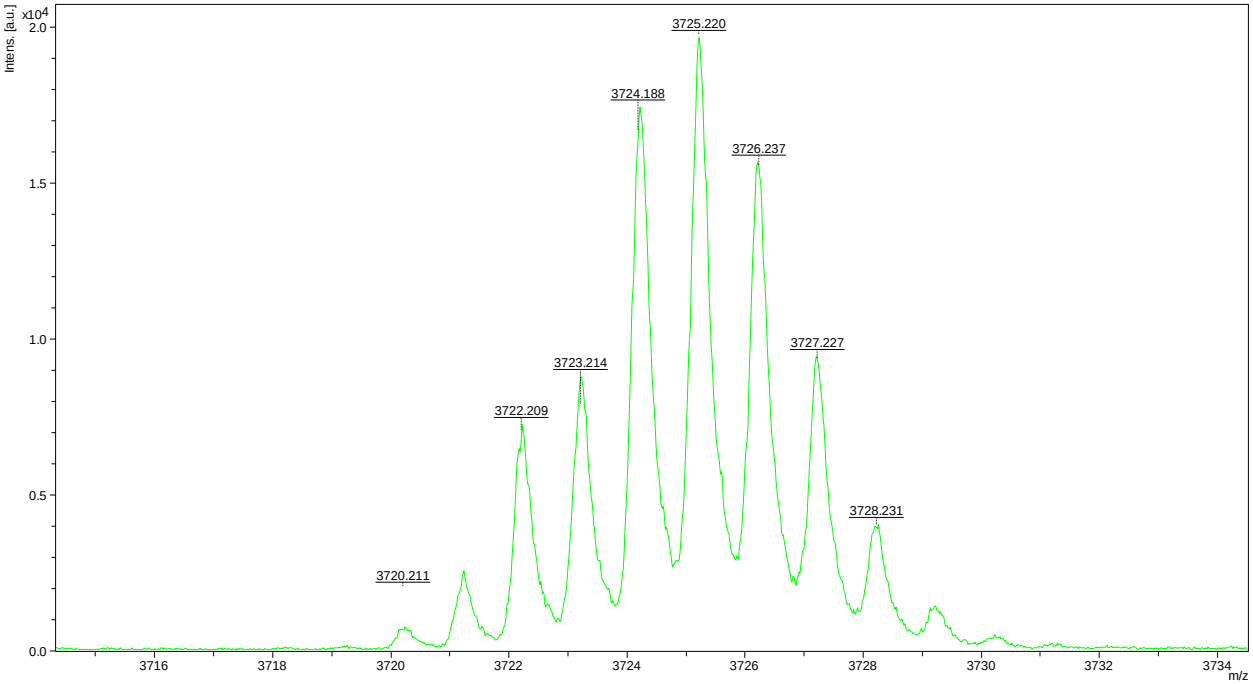

Intens. [a.u.]
8000
6000
4000
2000
0
1880
1882
1884
1886
1888
1890
1892
1894
m/z
1886.482
1887.479
1888.478
1884.407
1885.482
1883.424
1889.593
1890.554

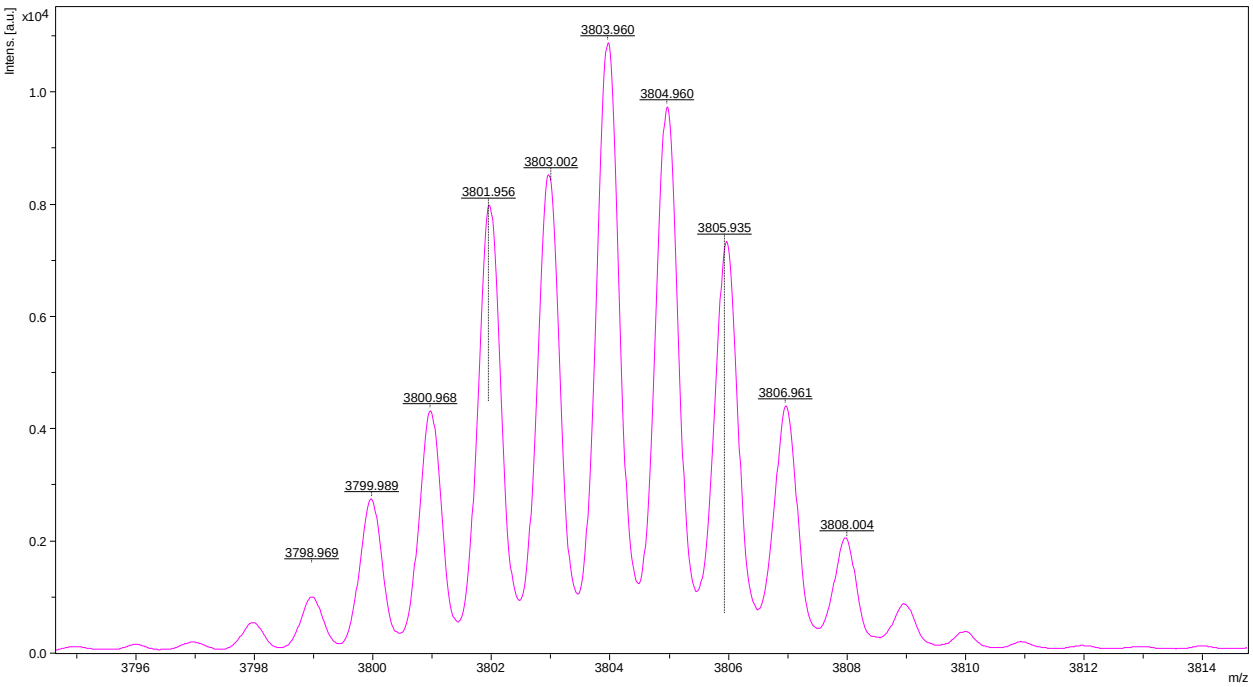

Intens. [a.u.]
x10^4
0.8
0.6
0.4
0.2
0.0
1878
1880
1882
1884
1886
1888
1890
1892
m/z
1885.716
1886.707
1883.713
1887.739
1882.680
1884.711
1888.770
1889.725
1879.763

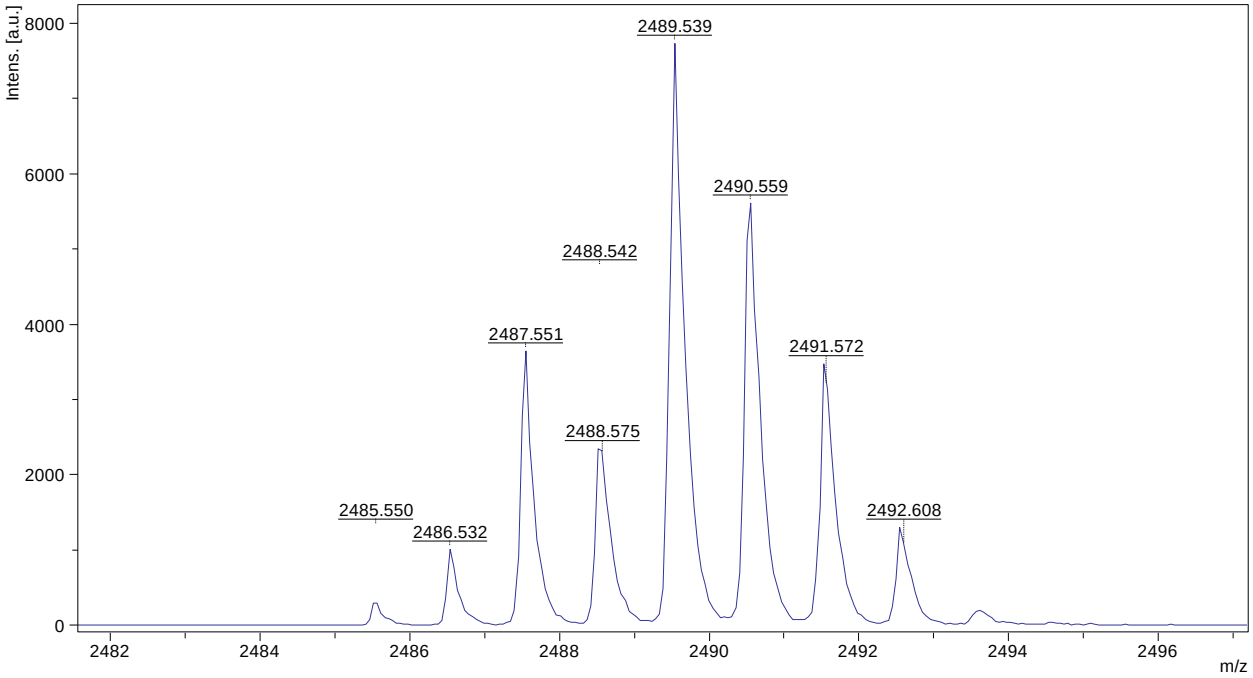

Intens. [a.u.]
2500
2000
1500
1000
500
0
1960.190
1962.187
1963.188
1964.156
1965.190
1966.194
1967.197
1968.198
1969.184
1958
1960
1962
1964
1966
1968
1970
m/z

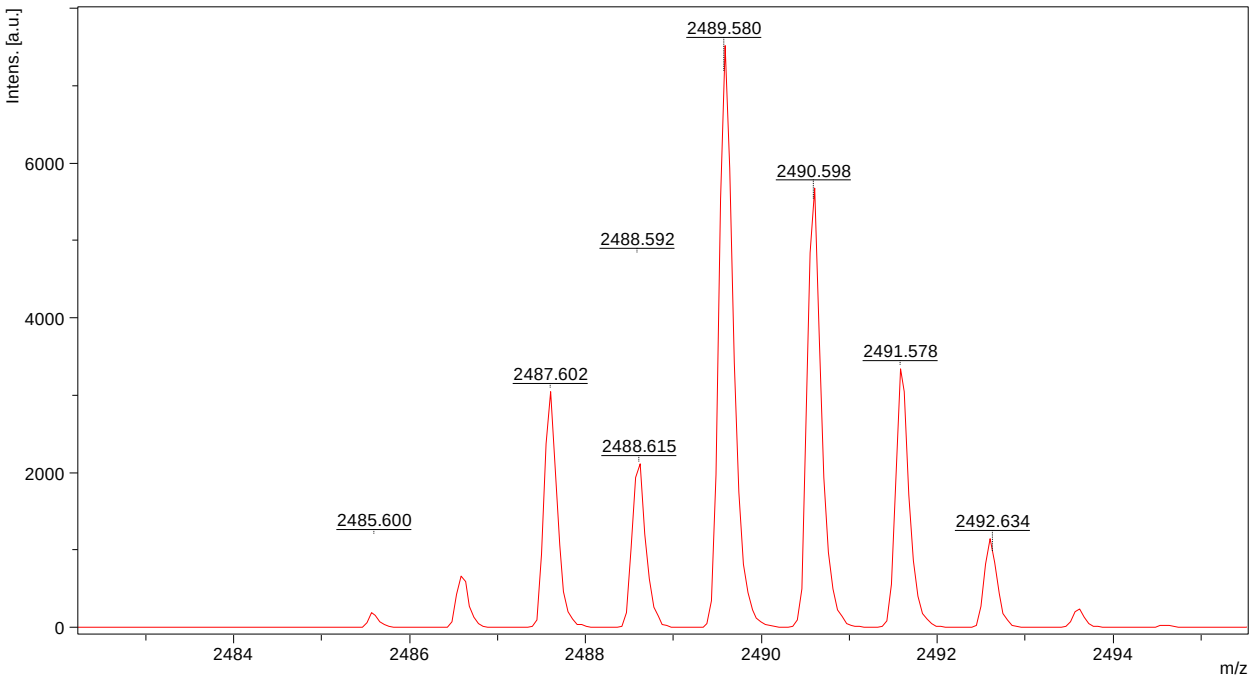

Intens. [a.u.]
x10^4
2.0
1.5
1.0
0.5
0.0
3644.417
3645.394
3646.406
3647.385
3642
3644
3646
3648
3650
3652
3654
m/z

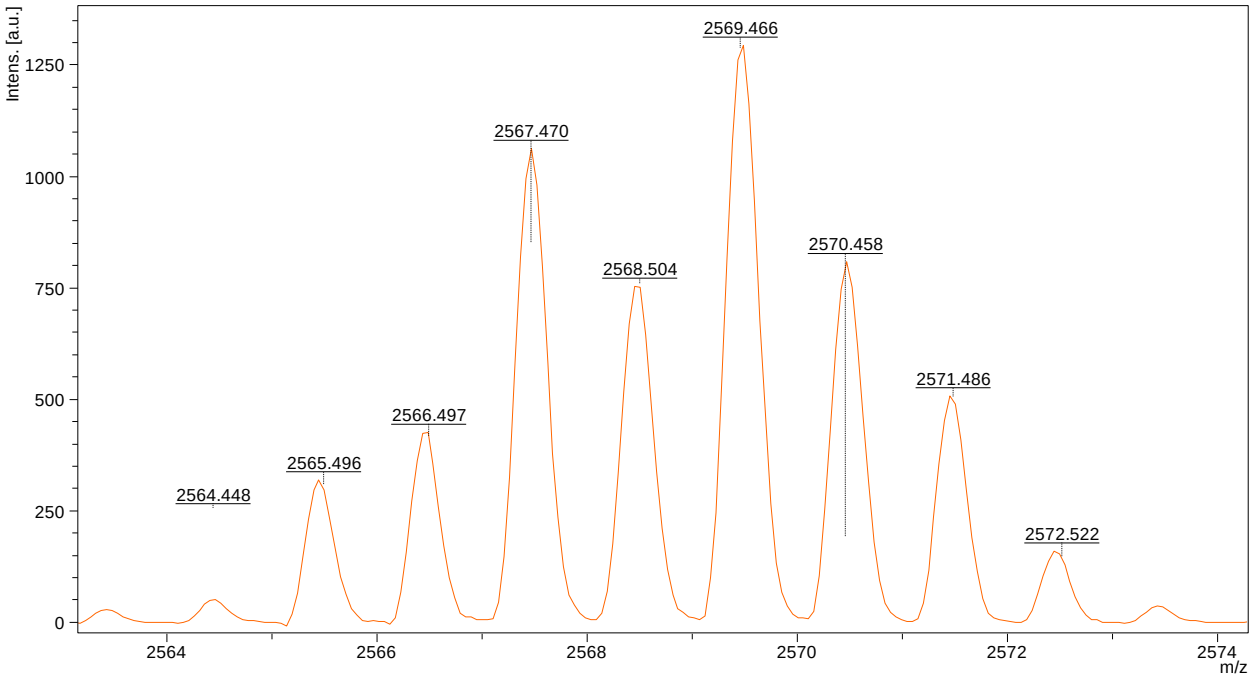

Intens. [a.u.]
x10^4
1.0
0.8
0.6
0.4
0.2
0.0
3725.261
3724.307
3726.307
3723.279
3727.311
3722.251
3728.287
3720.269
3716
3718
3720
3722
3724
3726
3728
3730
3732
3734
m/z


Intens. [a.u.]
x10^4
2.0
1.5
1.0
0.5
0.0
3725.220
3724.188
3726.237
3727.227
3723.214
3722.209
3728.231
3720.211
3716
3718
3720
3722
3724
3726
3728
3730
3732
3734
m/z


Intens. [a.u.]
x10^4
1.0
0.8
0.6
0.4
0.2
0.0
3803.960
3804.960
3803.002
3801.956
3805.935
3800.968
3806.961
3799.989
3808.004
3798.969
3796
3798
3800
3802
3804
3806
3808
3810
3812
3814
m/z


Intens. [a.u.]
8000
6000
4000
2000
0
2489.539
2490.559
2488.542
2487.551
2491.572
2488.575
2485.550
2486.532
2492.608
2482
2484
2486
2488
2490
2492
2494
2496
m/z


Intens. [a.u.]
6000
4000
2000
0
2489.580
2490.598
2488.592
2491.578
2487.602
2488.615
2485.600
2492.634
2484
2486
2488
2490
2492
2494
m/z


Intens. [a.u.]
1250
1000
750
500
250
0
2569.466
2567.470
2570.458
2568.504
2571.486
2566.497
2565.496
2564.448
2572.522
2564
2566
2568
2570
2572
2574
m/z