Doctoral Dissertation

Site-Directed Mutagenesis Studies Using Photo-Chemical Nucleic Acid Editing

by

Siddhant Sethi

Submitted to

Japan Advanced Institute of Science and Technology

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Under the guidance of

Professor Kenzo Fujimoto

School of Materials Science

Japan Advanced Institute of Science and technology

June 2018

Site-Directed Mutagenesis Studies Using Photo-Chemical Nucleic Acid Editing

Abstract

Siddhant Sethi
1540152
Fujimoto Laboratory

Introduction

Nucleic acid chemistry and its biological applications have a great scope in the development of futuristic drugs and cure for many diseases. Enzymatic methods for the nucleic acid manipulation have proven to be very useful *in vitro* but they have their disadvantages in the *in vivo* applications.[1] Therefore, chemical methods to edit nucleic acids were developed but the drawbacks of those chemical methods are numerous.[2]

Thus, to overcome these problems, photochemical methods to edit the nucleic acids have been devised which utilize single base modified nucleo-base to specifically target a desired sequence of DNA/RNA and edit that sequence at a single point. Fujmoto's group has discovered a novel compound, 3-Cyanovinylcarbazole, which can be easily incorporated as a nucleo-base and upon irradiation of 366nm radiation, forms a crosslink with the pyrimidine and lead to deamination to afford the transformation of the cytosine to uracil. The crosslink is photo-reversible and can be easily converted back by 312nm irradiation.[3-6]

The experiments have already shown success in the small nucleic acid sequences.[7-8] Moreover, the feasibility of the photo-crosslinking reaction using 3-Cyanovinylcarbazole *in vivo* has also been reported.[9]

Major drawback of this method is that the deamination step takes place at 90ºC, which is not a feasible condition for the *in vivo* applications. Therefore, in this research, the focus on development new method, such that, the deamination, which takes 200 years in physiological conditions without any external factor, can be carried out at 37ºC, i.e. physiological conditions in shorter time.

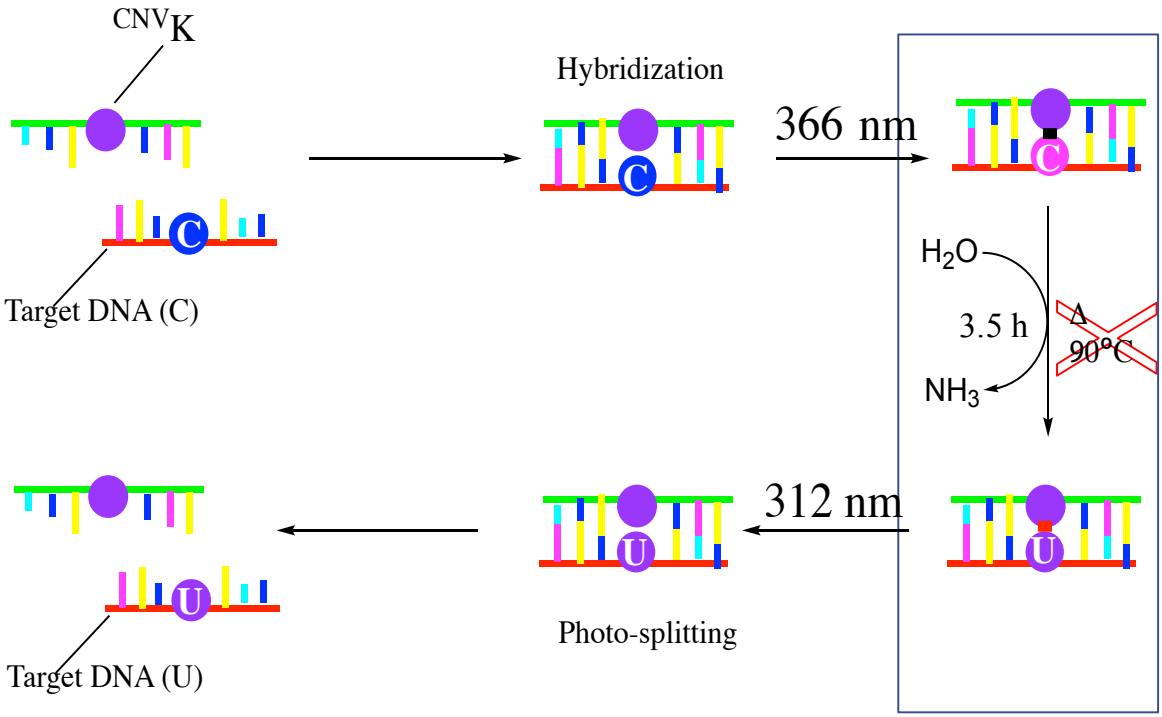


Figure 1: Schematic overview of photo-chemical DNA editing.

Results and Discussion

Chapter 2: The focus was to find the best counter-base of cytosine, based on hydrogen bonding, for photo-chemical site-directed mutagenesis using 3-cyanovinylcarbazole as the photo-active nucleoside which can crosslink with the target cytosine to afford cytosine to uracil transformation at physiological conditions. Different counter bases like guanine (G), inosine (I), 2-aminopurine (P), nebularine (R), and 5-nitroindole (N) were used to find the best counter base. Among all the bases, it was found that P, R, and N are not suitable counter base for photo-chemical cytosine to uracil transformation as when using these bases, no deamination reaction takes place. While in case of G, the deamination reaction is very slow and only 5% conversion is observed in 72h reaction time. Thus, the best counter base among the bases was found to be inosine which gives 35-40% in 72h reaction time at 37ºC having the optimal hydrogen bonding pattern before and after photo-cross-linking.

Chapter 3: In this chapter, the role of hydrophilicity and polarity of photo-cross-linker was discussed. Various derivatives of vinyl carbazole, like 3-cyanovinyl carbazole (^{CNV}K), 3-amidovinylcarbazole (^{NH2V}K), 3-methoxyvinylcarbazole (^{OMeV}K), and 3-carboxylvinylcarbazole (^{OHV}K), were used for studying the micro-environment around the

target cytosine crosslinked to photo-cross-linker during the deamination of cytosine. It was discovered that the hydrophilicity and polarity of the photo-cross-linker plays a crucial role in the deamination of cytosine to uracil via photo-cross-linking. ^{OMeV}K having the least hydrophilicity gave the least rate of reaction for the deamination reaction at varying temperature (90, 70, 50, and 37 ºC) while the highest reaction rate was observed with ^{OHV}K, which is most polar among the cross-linkers based on the polarity index (Log P). Thus, hydrophilicity and polarity around target cytosine are deciding factors in case of deamination reaction of cytosine via photo-cross-linking.

Chapter 4: Based on the findings of chapter 1 and 2, the overall micro-environment around the target cytosine for the mutation of cytosine to uracil via vinylcarbazole based photo-cross-linking was studied. A combination of counter bases (guanine (G), inosine (I), and cytosine (C)) and photo-cross-linkers (^{CNV}K, ^{NH2V}K, and ^{OHV}K) were used in the ODN to study the best match for acceleration of deamination of cytosine to uracil at physiological conditions. It turned out that the best combination of counter base and photo-cross-linker is inosine and ^{OHV}K which could give ~70% conversion of cytosine to uracil in 7 days at physiological conditions, which could be extended to ~90% in 20 days. The micro-environment around cytosine, including hydrogen bonding, hydrophilicity, and polarity of counter base and photo-cross-linker are key players for the photo-cross-link assisted deamination of cytosine to uracil.

Chapter 5: Based on the previous chapters we realized that the micro-environment around the target cytosine is deciding factor for rate of cytosine to uracil conversion via photo-cross-linking. Although, the reaction rate is very rather slow at physiological conditions even when inosine is counter base and ^{OHV}K is photo-cross-linker. Thus, a different approach to accelerate the rate of deamination reaction was used in which the ODN containing photo-cross-linker was divided into two parts between the counter base and photo-cross-linker. The adjoining part was modified with phosphate group at the terminal of counter base to increase the hydrophilicity near the cytosine. It was observed that upon the phosphate group modification near cytosine, ~100% conversion of cytosine to uracil was observed in just 24 h. Furthermore, we removed the ODN with counter base and modified the photo-cross-linker end with phosphate group to study the rate of reaction without hydrogen bonding and high hydrophilicity. It was found that the rate of reaction increased multifold with the modification giving ~100% conversion from cytosine to uracil in 3h at physiological conditions.

These results indicate that the deamination of cytosine to uracil is feasible at physiological conditions and heating to very high temperature is no more necessary to achieve the site-directed mutagenesis via photo-cross-linked cytosine. This has opened vast opportunities to use this enzyme free system in the biological samples at reduced cost and complexity to afford specific and site-directed cytosine to uracil conversions for the treatment of various genetic disorders like Leigh's syndrome.

Future Prospects

In this study, I have developed a refined way to carry out site-directed mutagenesis using enzyme free photo-chemical methods at physiological conditions. This technique has wide applications in the field of anti-sense technology, RNAi, RNA/DNA editing in cell for genome engineering to eradicate certain genetic disorders arising due to single T→C point mutations.

Keywords

- Genome editing
- 3cyanovinylcarbazole
- Site-directed mutagenesis
- Photo-cross-linking
- Cytosine deamination

References

1. M. Cascalho; *The Journal of Immunology* **2004**, 172(11), 6513-6518
2. C. Auerbach et al.; *Science* **1947**, 105 (2723): 243–7
3. K. Fujimoto et al.; *J. Am. Chem. Soc.* **2013**, 135, 16161-16167.
4. K. Fujimoto et al.; *Molecular BioSystems.* **2012**, 8 (2), 491-494.
5. K. Fujimoto et al.; *Chem. Commun.* **2010**, 46, 7545-47
6. K. Fujimoto et al.; *Chem. Commun.* **2010,** 11, 1661-64
7. K. Fujimoto et al.; *Chem. Lett.* **2008**, 37, 94-95
8. K. Fujimoto et al.; *Chem. Commun.* **2006**, 30, 3223-3225
9. K. Fujimoto et al.; *Biomater. Sci.*, **2014**, 2, 1154-1157

Table of Contents

Chapter 1: General Introduction

History of DNA

Deoxyribonucleic Acid (DNA) traces back its roots form the times when there were only living chemicals. DNA had the necessary properties, which were required for it to work as the genetic material to carry out the information from one generation to other without any significant loss of the information or the material carrying it, that were self-replication ability and the expressivity of the information at the vital time and place. Although, the evidences for these facts were found at a very later stage by Griffith, 1928[1] and Hershey-Chase, 1958[2], who discovered that DNA, and not protein, is the actual genetic material.

A few decades ago, before it was established that DNA is the genetic material, DNA was discovered by Friedrich Miescher (1869), who named it as nuclein due to its presence in the nucleus of cell, from pus cells.[3] A decade later, Albrecht Kossei in 1878 was able to identify the major constituents of DNA, that is, Adenosine, Thymine, Guanine and Cytosine.[4] Further in 1929, Phoebus Levene[5] found out that the DNA and RNA molecule contains sugar and phosphates groups besides nucleobases (Fig 1.1 (A)). But a while ago, in 1927, Nikolai Kolstov[6] suggested the presence of two strands, which are mirror images, and that they replicate in a semi-conservative fashion, which was confirmed the experiments of Matthew Meselson and Franklin Stahl in 1958[7]. A decade after Kolstov, the X-Ray Diffraction photographs of DNA molecule taken by William Astbury in 1937[8] suggested the duplex structure of DNA. Which was later confirmed by the single X-ray Diffraction images taken by Rosalind Franklin and Raymond Gosling in 1952[9]. Using these images, Watson and Crick proposed the Double Helix model of DNA in 1953.[10] Crick also came up with the concept of central dogma, which gave a relation between DNA, RNA, and protein, in 1958[11] (Fig 1.1 (B)). Later, in the year 1977, Sanger[12] reported the method for sequencing DNA, more commonly known as “Sanger’s Method” and Maxam and Gilbert[13] gave chemical degradation method of DNA sequencing.

The discovery of double helical structure of DNA and the sequencing methods of DNAs opened a huge portal in the field of DNA related research. It led to the complete profiling of human genome, DNA nanotechnology, genetic engineering, artificial gene, molecular medicines, and so on.

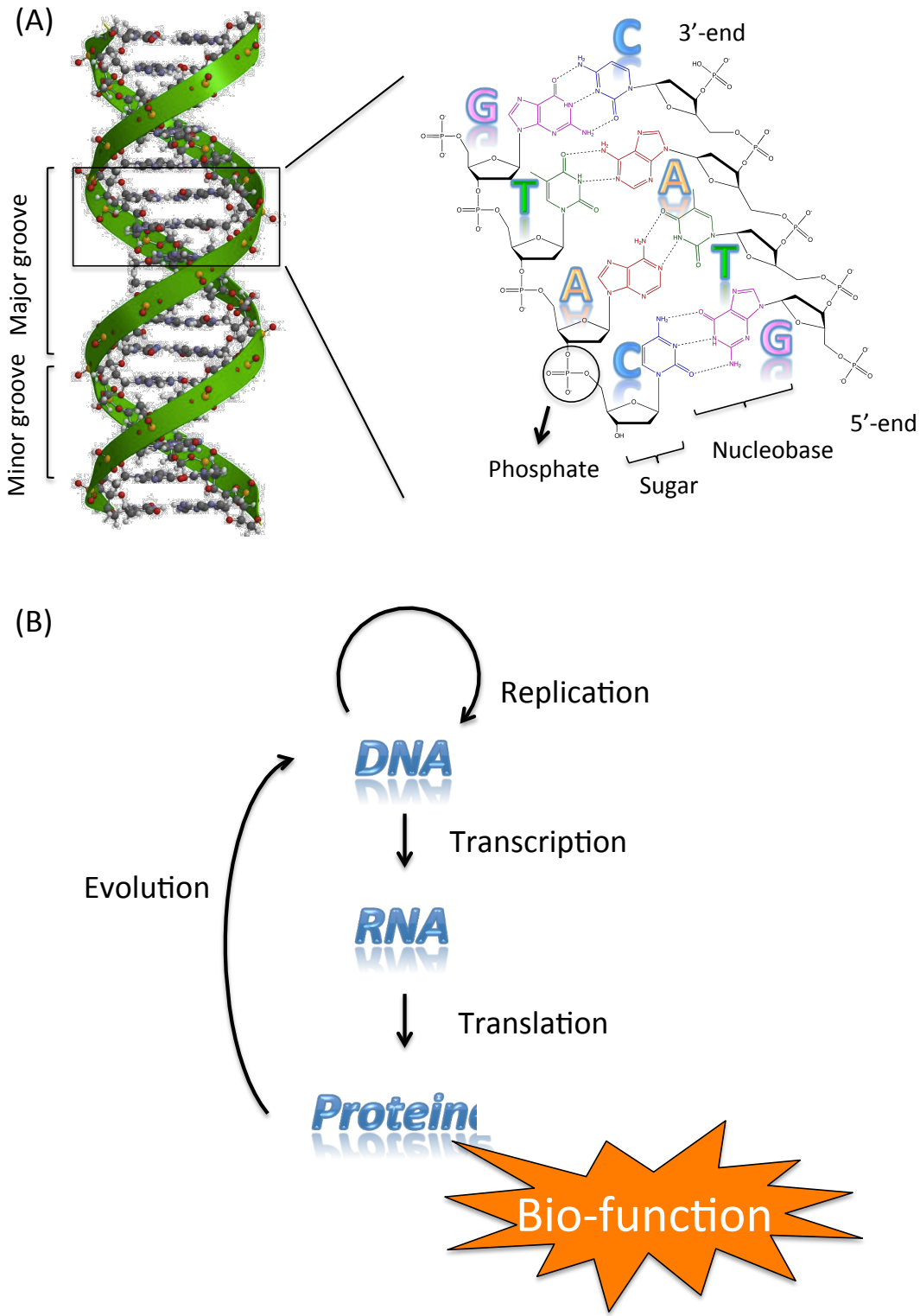


Figure 1.1: (A) Molecular model of DNA helical duplex; The conformation of DNA is double helical structure formed Watson Crick interaction. (B)The scheme of central dogma.

Mutations in DNA

Mutation is any change in the DNA, which results in a variant form that may be transmitted to subsequent generations. It can be caused by alteration of large part of gene or a single nucleotide in DNA[14]. The mutations occurring in chromosomes during meiosis can be numerical or structural.

A mutation in a single base (point mutation) generally affects the function of single gene and may produce three kinds of mutations: *Missense mutation*, where, due to the mutation, one of the amino acid in a protein is replaced by other; *Nonsense mutation* where, due to mutation, a normal codon is replaced by stop codon in mid-sequence leading to early translation termination; *Frame shift mutation,* leads to change in reading frame.[14]

Structural Mutations[15] occur due to deletion or rearrangement of part of the chromosome. They include: *Deletion*, one segment of chromosome gets deleted. This aberration includes Wolf-Hirschhorn syndrome (partial 4p deletion) and Jacobsen syndrome (11q terminal deletion); *Duplication*, single or more segment of the chromosome is present in multiple copies. This includes Charcot-Marie-Tooth disease type 1A due to multiple copies of peripheral myelin protein 22 gene (PMP22) at chromosome 17; *Translocations*, a segment of chromosome is translocated to another chromosome. There could be reciprocal translocation (between two chromosomes segments) or Robertsonian translocation (attaching on one chromosome on the centromere of other); *Inversions*, a part of the chromosome is inverted and reattached; *Rings*, circularization of a segment of chromosome; *Isochromosome*, formation of mirror copies of chromosome (or a part thereof).

Numerical Mutations[15] are caused due to unequal chromosome division resulting in an extra chromosome or a deficiency in chromosomes. Examples of these mutations are Down syndrome (trisomy 21) and Turner syndrome.

Site-directed mutagenesis

Mutagenesis is the process of inducing change in a cell by the altering its DNA. It results in the altered gene or a specific nucleotide. Mutations provide the way for evolution if they produce an enhanced protein or beneficial functions.[16]

Site directed mutagenesis also called oligonucleotide directed mutagenesis is an in-vitro method of inducing specific mutation in double stranded plasmid DNA by using oligonucleotide primers. It is important in studying protein activity and screening mutations.[17]

In this method, an oligonucleotide with required base switching is annealed onto a copy of single stranded DNA incorporated into a suitable vector. The purpose of the primer is served by this annealed strand which lead to synthesis of complimentary strand of the plasmid vector. This mismatched double stranded recombinant plasmid is transformed into bacteria where the cellular enzymes repair the mismatch. About half of the transformant will restore the normal sequence while the rest retains the desired mutation. The transformants are checked by sequencing of the plasmid DNA.[18]

One of the ways to mutagenesis is single primer method where a primer is annealed to the single-stranded template, which is shortly extended with Klenow fragment, which then transfects the bacteria. The mismatched base may be removed by the 5'-3' exonuclease activity. So, the ends of new strands are ligated after extending the strand. In the double primer method, two primers are used one is mutagenic and the other non-mutagenic primer is used which is 5' to the mutagenic primer (Fig 1.2). It protects the mismatched base after extension and ligation.[19]

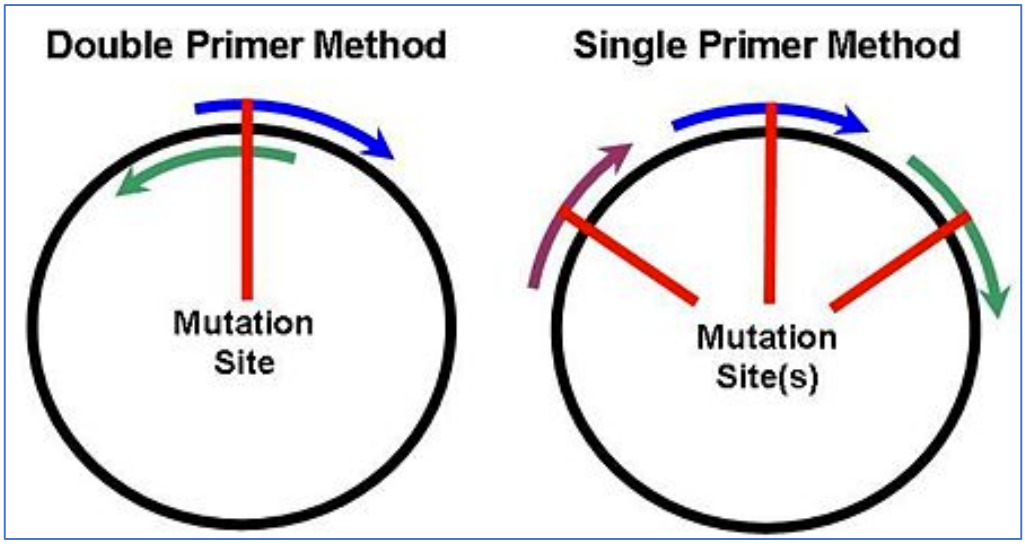


Figure 1.2: Single primer and double primer method for site-directed mutagenesis

The PCR-plasmid based method for site-directed mutagenesis, though advantageous, has certain limitations like availability of restriction sites, incompatibility and large size of gene as compared to plasmid, low fidelity and high cost of polymerase, etc. when applied to mammalian systems.[20-22] Therefore, new methods for site directed mutagenesis were discovered which include chemical methods and enzymatic methods.

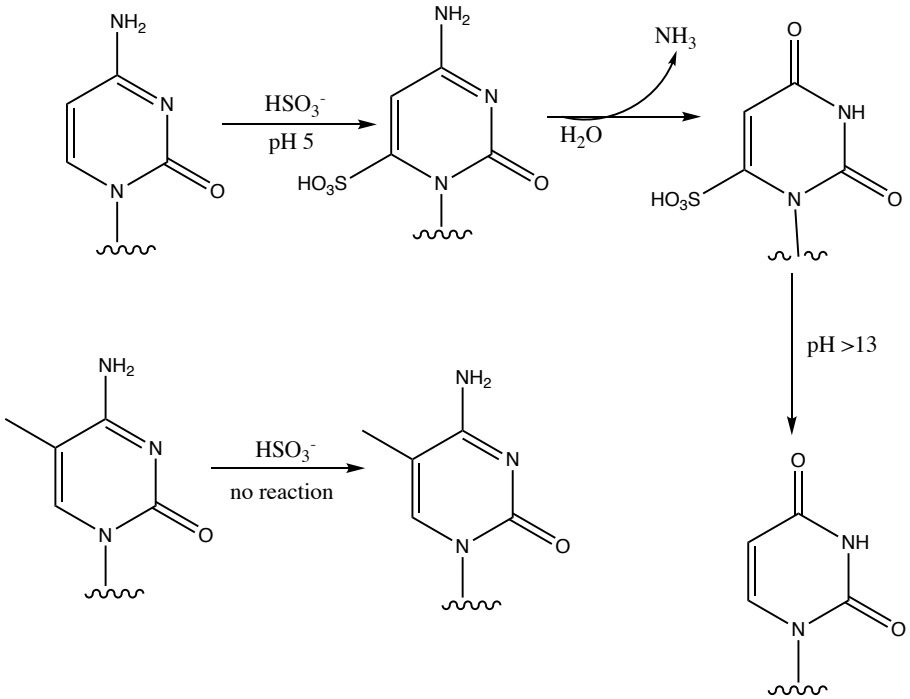
Chemical methods for mutagenesis

The chemical methods used for inducing mutation in the DNA are based on the cytosine to uracil or 5-methyl cytosine to thymine conversion in the DNA strand. The mutation methods often use chemicals like sodium bisulphite, nitrous acid, formic acid, hydrazine, ethyl methyl sulfonate etc. Myers et al. (1985) and Walton et al. (1991) reviewed the reports that bisulfite, nitrous acid, and hydrazine can be used for generating random point mutations in DNA strand with a frequency of a few base pair mutation per kbp of DNA.[23-24]

Sodium bisulphite

In 1970, Shapiro et al. and Hayatsu et al reported that sodium bisulphite can react with ssDNA leading to the addition of sulphonate group on the 5,6-double bond of cytosine and be used for the purpose of inducing mutagenesis in the DNA catalysing the deamination of the cytosine to uracil in the ssDNA through the removal of 4-amino group. Varying the pH allows the restriction of the backward reaction. The reaction of bisulfite with cytosine containing DNA

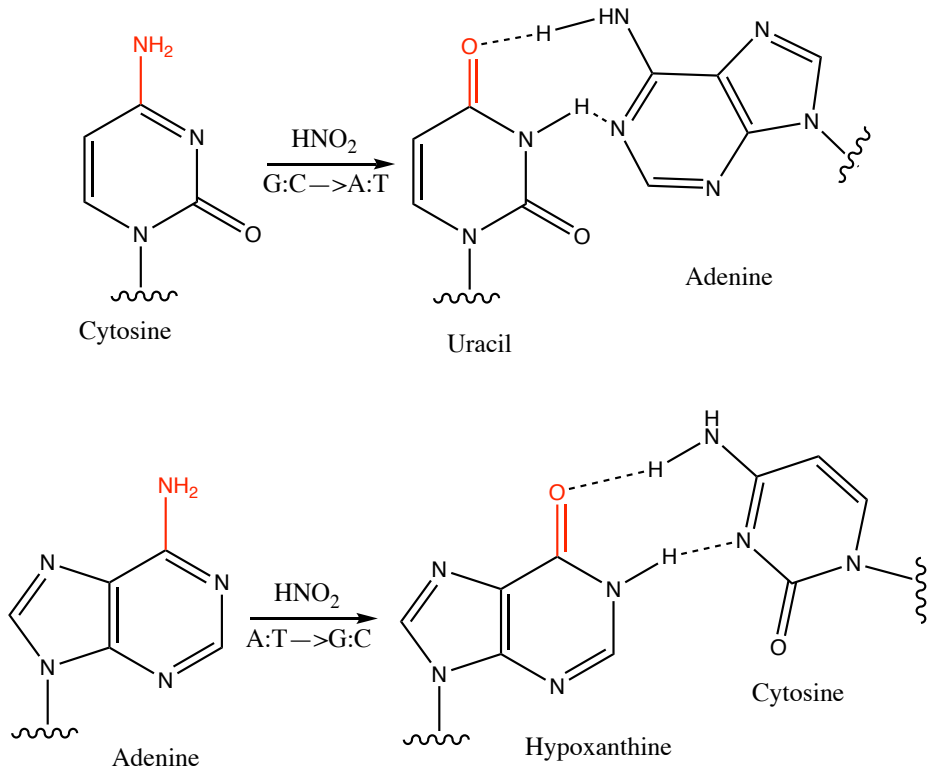
strand is shown in scheme 1.1. Sodium bisulfite lead to deamination of cytosine in ssDNA by forming an adduct with cytosine at pH 5 which promotes the deamination of cytosine. The bisulfite adduct can be degraded at pH 13.[25-26] Sodium bisulfite gained momentum due to its application in the epigenotyping of methyl cytosine selective hydrolysis.[27-28] Epigenetics, or specifically, identification of methylation of DNA in the genome is carried out using sodium bisulphite which selectively deaminate cytosine to uracil which is read as thymine after PCR unless the cytosine is methylated, in which case the deamination does not take place and the sequenced DNA can be shown to be methylated or not.[29]



Scheme 1.1: Bisulfite treatment of cytosine and methyl-cytosine containing DNA strand.

Nitrous Acid

Nitrous acid, generated from nitrosamines, nitrites, and nitrated, are powerful deaminating agents which can deaminate cytoine, guanine, or adenine leading to formation of uracil, xanthine, or hypoxanthine, respectively (Scheme 1.2).[30-31] This further leads to mismatch in the base pairing of the original bases causing point mutations in the dsDNA which in turn leads to mutation in the protein.



Scheme 1.2: Mutation through nitrous acid.

Formaldehyde

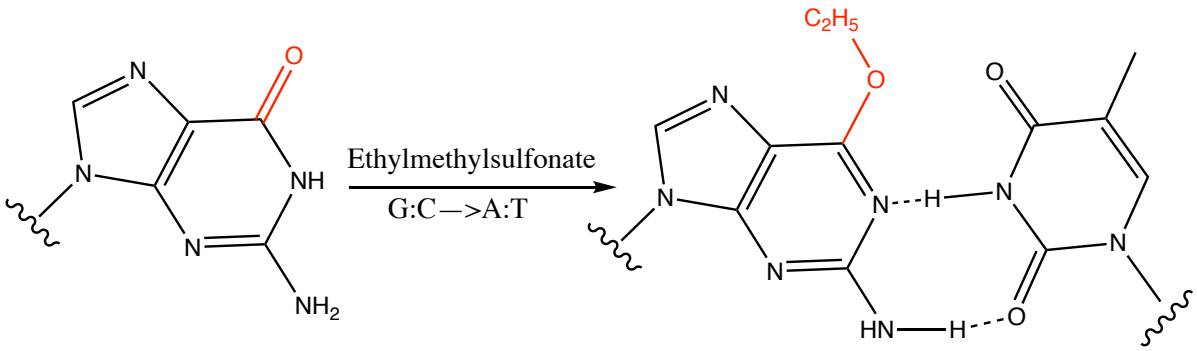
Formaldehyde, an omni-present compound in food, air, and water, is a highly reactive compound which can easily and rapidly react with amines, thiols, amides, hydroxyls, etc. and act as an electrophile to form adducts between DNA, RNA, and/or proteins.[32] Formaldehyde can also react with the amino group of guanine, cytosine, and thymine in the DNA strand to form hydroxtmethyl derivatives of the nucleobases which leads to substitution mutation from G:C→A:T or A:T→C:G.[33-35] Apart from forming hydroxylmethyl derivatives, formaldehyde also induce cross-linking in the DNA strands, both inter- and intra-, leading to chromosomal aberrations and joining of telomeres of chromosomes.[36]

Besides formaldehyde, other common aldehydes like acetaldehyde, crotonaldehyde, glyoxal, methylglyoxal etc have also been reported to form exocyclic alkyl derivatives of guanine leading to mutation at G:C base pair.[37-40]

Ethylmethyl sulfonate

Besides the mentioned chemicals, ethyl methyl sulfonate (EMS) has been extensively used in the past for directed mutation of G:C→A:T transformation in the whole organisms with a

frequency of 5-6 bp/kbp,[41] while Lai et al. in 2004 showed for the first time that EMS can be used for inducing mutations in the defined DNA strand *in vitro*.[42] Ethyl methyl suphonate acts as an alkylating agent and attacks the carbonyl group of nucleoside as an electrophile. Due to inability of alkylated carbonyl group, the natural bonding gets disrupted leading to base pair mutation from G:C→A:T (Scheme 1.3).



Scheme 1.3: Ethylmethylsulfonate induced mutation in DNA.

The chemical methods used for mutagenesis have certain disadvantages which include, but are not limited to, deleterious chemicals, drastic conditions of reactions like extreme pH, random mutations, no control over the site and frequency of mutation, carcinogenicity, lethality due to mutation etc.

Enzymatic tools for targeted genome editing

Due to incompetency of previous chemical method and PCR based method to induce targeted genome engineering in the eukaryotic cells, specifically mammalian cells, alternate methods were developed to tackle the targeted genome editing through mutagenesis, repair of genes, and correction of deleterious mutations *in vivo*. Homologous gene targeting was first shown by Gerald Fink in 1978 in the yeast cells.[43] And subjected to mammalian cells by Oliver Smithies et al.[44] Thereafter, various homing endonucleases and recombinases have been discovered which are effective in targeted homologous gene engineering. Some of them are discussed herein:

Meganucleases

Meganucleases were discovered in the yeast cells by the discovery of two proteins which acted as endonucleases and triggered the recombination in the yeast cells.[45-48]

The peculiar quality of the mentioned proteins was high specificity as the recognition site of these nucleases were rather long (12-40 bp).[49-50] In the past decades, many members of homing endonuclease or mega nucleases have been identified in almost all the cells.[51] This high specificity of meganucleases gives them high precision and low toxicity as compared to other restriction enzymes. Due to these factors, the meganucleases have shown promises in the field of genetic engineering and gene editing.[52-53] Since, the types of meganuclease and their recognition sites were limited, the scope of use was restricted. Thus, modified meganucleases were designed to target specific genes at desired locations by introducing small variation in the amino acid sequence of the native meganucleases based on the recognition site.[54-56] another approach used for engineering meganucleases was to produce chimeric nucleases by fusing domains from different enzymes.[57-59] Despite the advantages of the homing endonuclease, the applications of the enzymes are limited due to high cost involved in the production of engineered meganucleases and requirement of specific meganuclease for each gene.

Zinc-finger Nucleases (ZFN)

The ultimate aim for gene engineering is to repair a defective gene rather than replace it by using a viral vector as mostly only a small portion of the gene is defected.[60] ZFN are artificial restriction enzymes having two domains: zinc-finger DNA-binding domain and DNA cleavage domain which are fused together.[61] The zinc-finger domains are engineered to target complex genomes as they are designed to bind with specific DNA sequence.[62] Typically, the DNA-binding domain has 3-6 zinc finger repeats that can identify 9-18 bp DNA segment. The sequence this domain recognize must be in the form of $(GNN)_N$ otherwise the assembly fails to function as the zinc finger repeats might overlap and the context dependence is lost.[63] The DNA-cleavage domain is usually a type IIs recognition enzyme (FokI) which is non-specific in nature and function only upon forming a dimer.[64-65] ZFNs are used through 2 approaches, homologous recombination (HR) or non-homologous end joining (NHEJ). Wherein, NHEJ is usually error-prone and can create localized insertion or deletions in the DNA strand. Whereas, when homologous recombination approach is used, the replacement of a particular section of the gene is possible (Fig 1.3).[66]

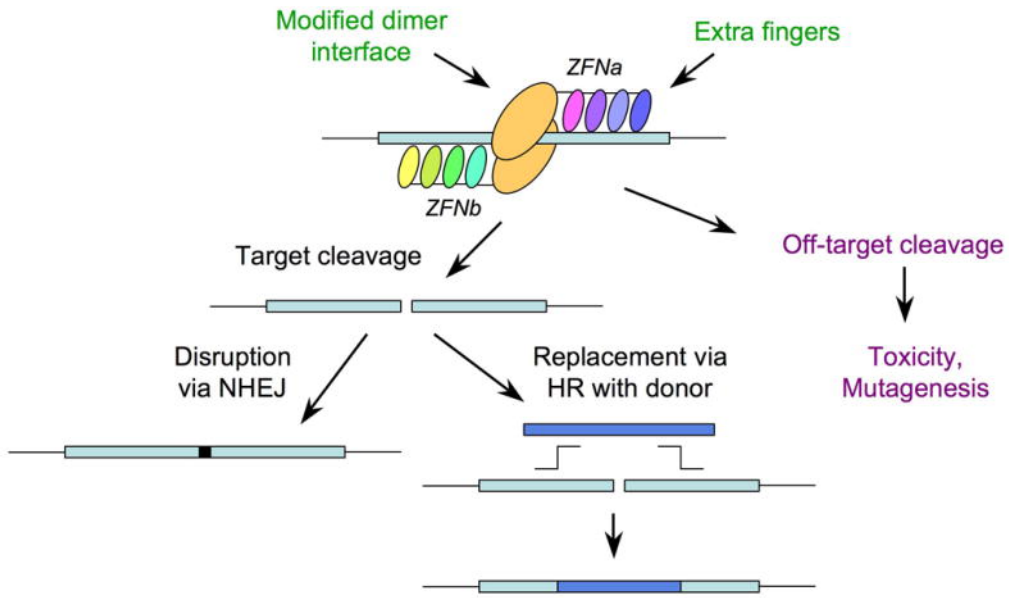


Figure 1.3: Mechanism of working of the Zinc-finger nuclease.

ZFN has been used for gene manipulation in various plants and animals like tobacco, soybean, *Drosophila, C. elegans*, zebrafish, mice, rat, rabbit, pig, cattle, etc.[67=76] Besides animal and plants, the technique has shown to be effective in the clinical trials for a method to treat HIV/AIDS by gene disruption of CCR5 gene in the CD4+ human T cells.[77] Beside all the advantages, there are many problems associated with ZFN like off-target cleavage (fig 1), cytotoxicity, and immunogenicity.[78-79]

Transcription Activator-like Effector-based Nuclease (TALEN)

Transcription activation-like effector nucleases are special kind of engineered restriction enzymes made by fusing TAL effector protein, a highly conserved 33-34 amino acid protein with 12th and 13th position (repeat variable diresidue, RVD) variable for recognition of DNA and binding, with a DNA cleavage domain from FokI restriction enzyme.[80-85] The TALEN system bind to target in a very specific fashion due to the RVD domain, which even with slight change in sequence can lead to change in specificity of DNA binding.[74] This direct relationship between amino acid and DNA recognition can lead to efficient protein engineering through gene editing.[86] A simple working mechanism of TALEN has been shown in Fig 1.4,[87] wherein a pair of TALENs are required for the working with complimentary FokI units on each side to cut the

recognized DNA double strand at palindromic sequences. Based on these excisions, NHEJ or HR based repair of the genome can be accomplished.[88]

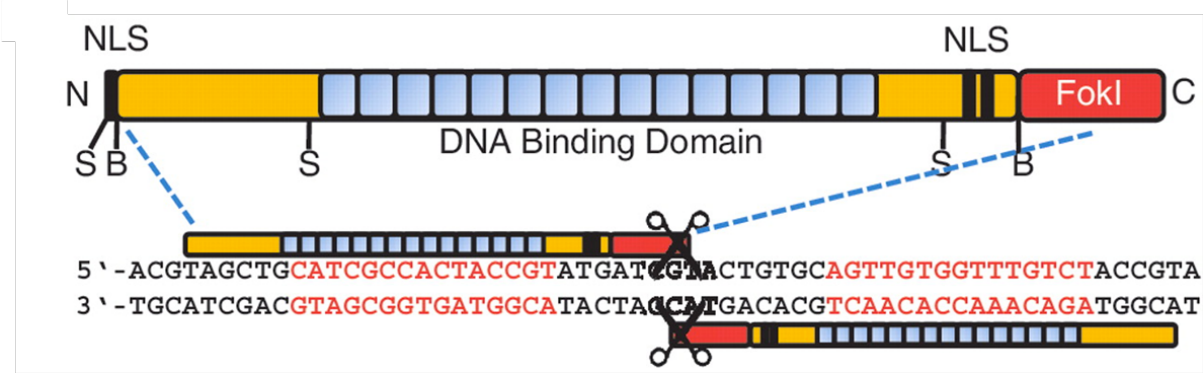


Figure 1.4: DNA binding domain of TALEN fused with DNA-cleavage domain (FokI) bound to dsDNA to introduce double stranded break required for genetic engineering.

TALEN, although very young, has been already used in many organisms including, but not limited to, fruit fly, cricket, rat, frog, pig, rice, human somatic and pluripotent stem cells.[89-96] TALEN, though very useful, has some disadvantages like high cost of producing the system, off-target cleavage, time and labour extensive process.[97]

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)

CRISPR is short palindromic sequence of DNA which act as immune system of the bacteria and archaea bacteria by identifying the attacking viral genome, where the spacers are produced after each viral genome that has attacked, along with Cas (CRISPR associated protein) that is translated based on the CRISPR sequences to digest the alien DNA.[98-101] Upon viral invasion, the Cas protein fuse with crRNA (CRISPR RNA) to digest and neutralise the viral DNA and in the process, produce short segments of viral genome to identify the same virus in case of subsequent attacks. The small segment of DNA gets incorporated in the bacterial genome near the CRISPR array and later transcribe Cas and crRNA to tackle the viral attacks later (Fig 1.5).[102]

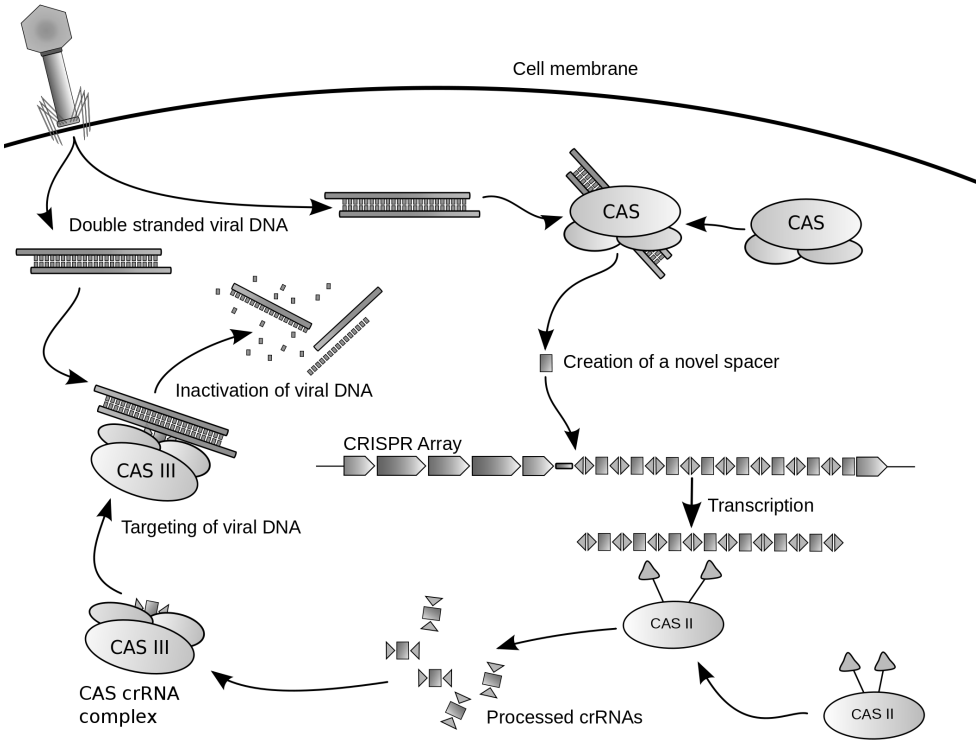


Figure 1.5: Diagrammatic sketch of mechanism of CRISPR-Cas system as bacterial immune system.

This system has been manipulated to be used as genetic engineering tool.[103] CRISPR-Cas9 system has been engineered to modify genome by excising the genome at specific sites recognized by the Cas9 protein fused with guide RNA (gRNA) or single gRNA (sgRNA), synthesized artificially to recognize specific sequence which is a fusion of crRNA and transcrRNA.[104-105] In order for gRNA to recognize and cut the DNA strand, it further requires a 3-5bp DNA segment adjacent to CRISPR spacer called protospacer adjacent motifs (PAM).[106] For the purpose of genome editing, the CRISPR-Cas9 system is chiefly used with a gRNA which identifies a specific sequence of DNA and cut it near the PAM sequence to induce a double stranded break in the genome which is then repaired either by NHEJ of homology directed repair (HDR) mechanism requiring a donor DNA template to be inserted at the site of DSB (Fig 1.6).[107]

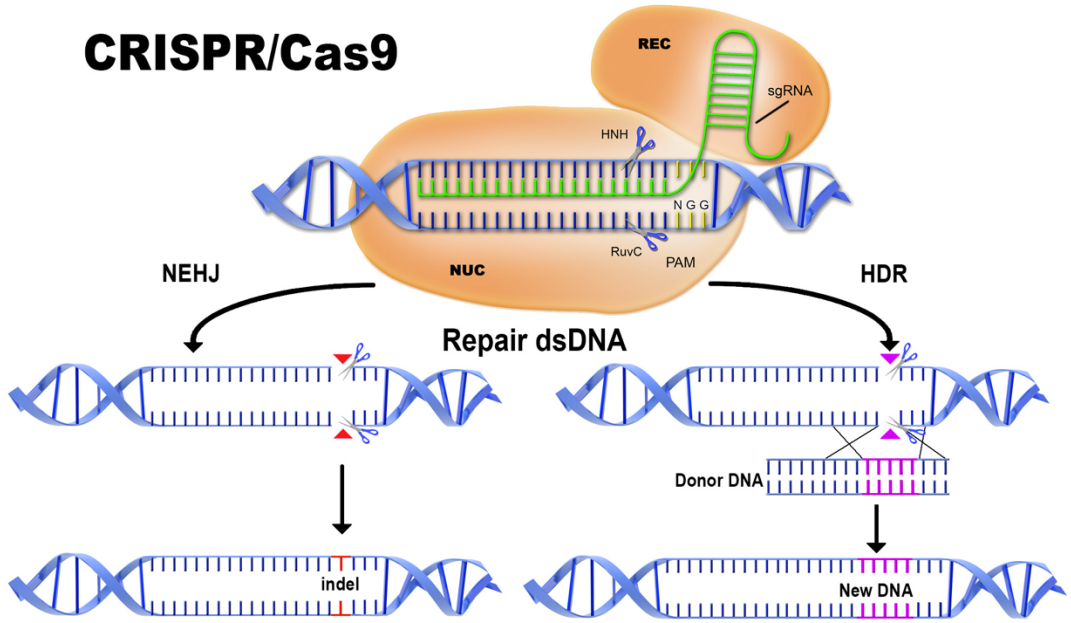


Figure 1.6: Mechanism of CRISPR-Cas9 system for genome editing by NEHJ and HDR.

To transfect the target cells, a plasmid is usually used having CRISPR sequence, for transcription of crRNA, PAM sequence, and code for Cas9 protein gene. Multiple crRNA and transcrRNA can be packed into a single plasmid, with one RNA for one application, to make sgRNA along with Cas9 gene (Fig 1.7).[108-109]

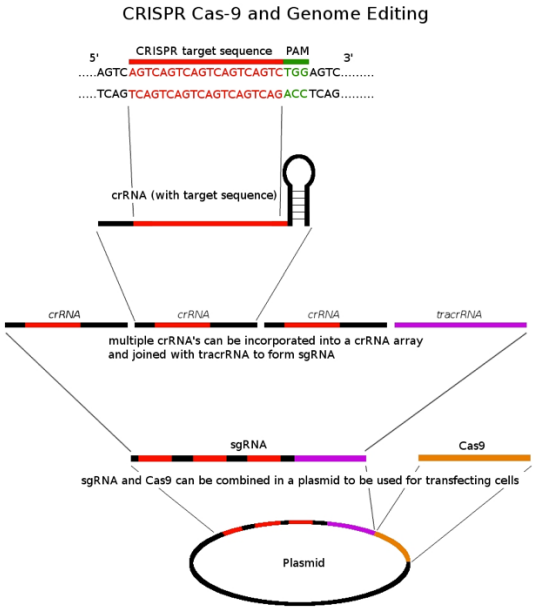
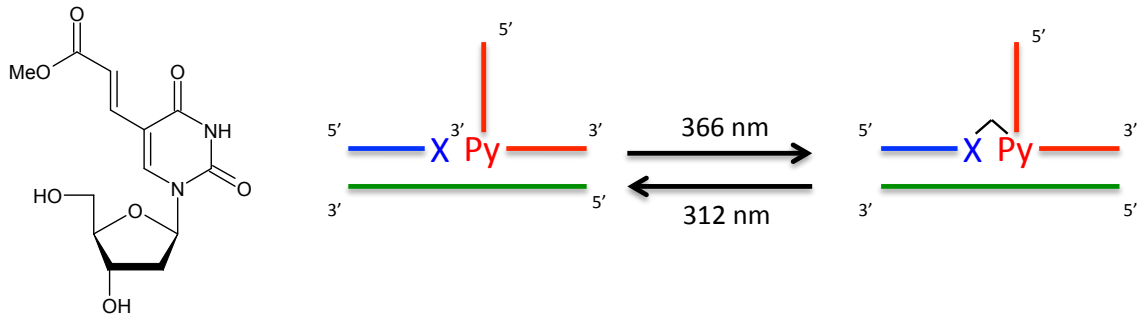


Figure 1.7: Structural representation of plasmid used for CRISPR-Cas9 showing multiple crRNA and transcrRNA combined to form one sgRNA for cell transfection.

CRISPR-Cas9 being a versatile tool for genome engineering has certain limitations. The major limitations are off-target cleavage activity and cytotoxicity has been observed. [110-113]

Photo-chemical methods

There are many drawbacks of the chemical and enzymatic methods available for genetic engineering, thus the need to develop a non-toxic, cheap, precise, accurate, and easy method for site-directed mutagenesis was of utmost importance. Thus, to overcome these problems, photochemical methods to edit the nucleic acids have been devised which utilize single base modified nucleo-base to specifically target a desired sequence of DNA/RNA and edit that sequence at a single point. Fujimoto et al. (2000) reported 5-vinyldeoxyuridine (^{V}U), a modification in uridine, which could photo-crosslink with pyrimidine in another DNA/RNA strand at the terminal position lading to ligation of two DNA strands without the use of enzymes and showing no side reaction by photoradiation of 366nm UV radiation.[114] This reaction ligated T and ^{V}U via [2+2] photocyclization.[115] Moreover, it was further reported that another artificial nucleotide 5-Carboxyvinyldeoxyuridine (^{CV}U),[116] was more responsive than ^{V}U. The photochemical reaction of ^{CV}U is shown in Fig. 6.1A. The template-directed photochemical ligation with ^{CV}U and other short ODNs, dependent on sequence of template, was achieved using ^{CV}U based photo-ligation to create self-assembled long dsDNA.[117]



Scheme 1.4: Reversible photo-ligation using ^{CV}U.

The photo-ligation technique had various advantages in terms of DNA nano-technology, DNA sensors, and DNA computing.[118-120] In terms of DNA manipulation as well, ^{CV}U was used for

5-methylcytosine to thymine transformation through photo-ligation and heat treatment but the drawback of ^{CV}U was requirement of prolonged irradiation with UV for photo-ligation.[121]

Therefore, Fujimoto's group discovered a novel compound, 3-Cyanovinylcarbazole, which can be easily incorporated as a nucleo-base and upon irradiation of 366nm for a few seconds, forms a crosslink with the pyrimidine and lead to deamination to afford the transformation of the cytosine to uracil. The crosslink is photo-reversible and can be easily converted back by 312nm irradiation (Fig 1.8).[122-125]

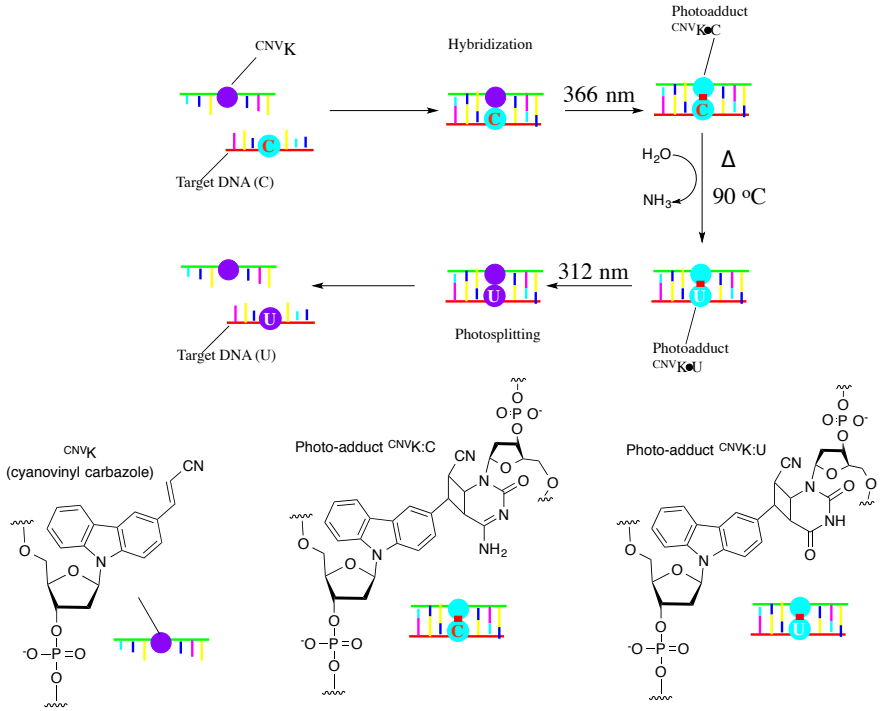
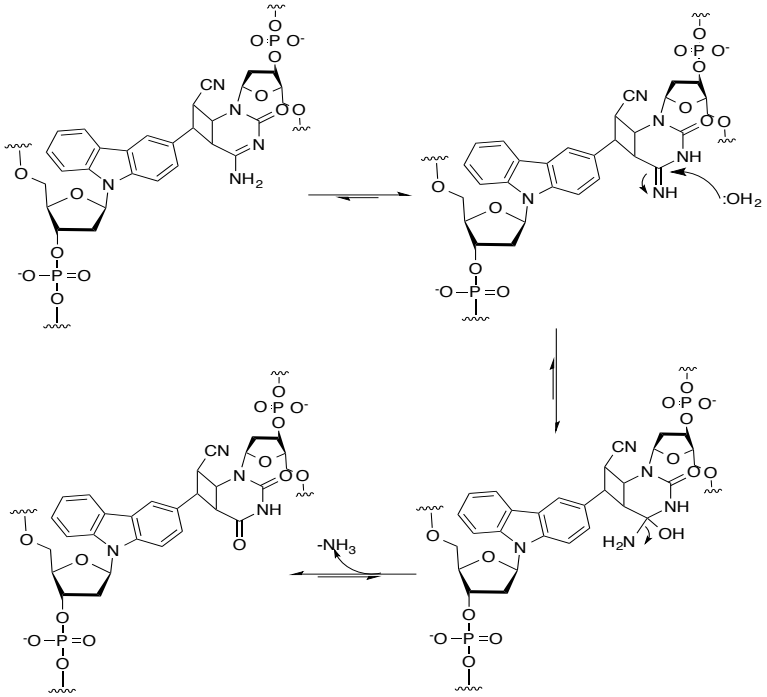


Figure 1.8: Schematic diagram of photochemical DNA/RNA editing

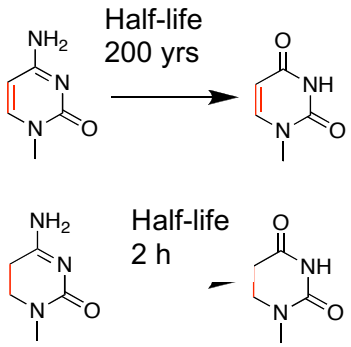
The approach for this reaction (Scheme 1.2) is through the formation of a cyclobutane ring which is formed between the trans double bonds of cyanovinyl group of cyanocarbazole and the double bonds present in the purine base of DNA. The mechanism follows the wherein, after the formation of cyclobutane ring, the amino group on cytosine undergoes amino-imine tautomerism. The imine moiety gets attacked by water molecule, forming a hydroxyl bond which is converted to keto group via removal of ammonia, ultimately leading to formation of uracil in place of cytosine.



Scheme 1.2: Mechanism of deamination in photo-chemical DNA editing.

The experiments have already shown success in the small nucleic acid sequences.[126-127] Moreover, the feasibility of the photo-crosslinking reaction using 3-Cyanovinylcarbazole *in vivo* has also been reported.[128]

Major drawback of this method is that the deamination step takes place at 90ºC, which is not a feasible condition for the *in vivo* applications. Therefore, in this research, the focus on development new method, such that, the deamination, which takes 200 years in physiological conditions without any external factor due to presence of conjugation in the cytosine molecule which makes it difficult for nucleophile to attack and disrupt the conjugation, can be carried out at 37ºC, i.e. physiological conditions in short time when the double bond at position 4-5 is degraded to single bond to make dihydrocytosine which doesn't possess conjugation and attack of nucleophile is facilitated (Scheme 1.3).[129]



Scheme 1.3: Half-life of cytosine and dihydrocytosine in nature.

Objective of this study

Genetic engineering for treatment of genetic disorders and creating desired mutations has been an upcoming field in today's world. Many techniques, for manipulating genome, have been developed since the discovery of DNA as genetic material. These methods, like enzymatic genome editing, chemical genome editing etc., have proven to be useful in their niche but these methods had certain disadvantages in terms of applicability, cost, ease of use, and cytotoxicity due to which new methods were needed to be developed which should be fast, convenient, cheap, safe, and easy to use. In this study, I have developed one such technique using 3-cyanovinylcarbazole and its derivatives for photo-chemical site-directed mutagenesis. The photo-chemical DNA manipulation doesn't require enzymes and relatively cheaper than most available techniques for the purpose. The deamination of cytosine to uracil in a site-specific and precise manner is the prime objective of this technique.

In chapter 2, I have tried to improve the reactivity of CNVK mediated photo-cross-link induced deamination of cytosine at physiological conditions, the evaluation of base pairing in cytosine has been carried out with respect to its deamination. Guanine has been replaced with 4 different counter bases (inosine, 2-aminopurine, 5-nitroindole, and nebularine), showing distinct hydrogen bonding pattern with target cytosine, which were incorporated at -1 position with respect to CNVK in the CNVK-modified photo-responsive oligodeoxyribonucleotides to ascertain the role of hydrogen bonding in the deamination at physiological conditions. Among the counter bases, inosine has shown highest acceleration towards the photo-induced deamination reaction.

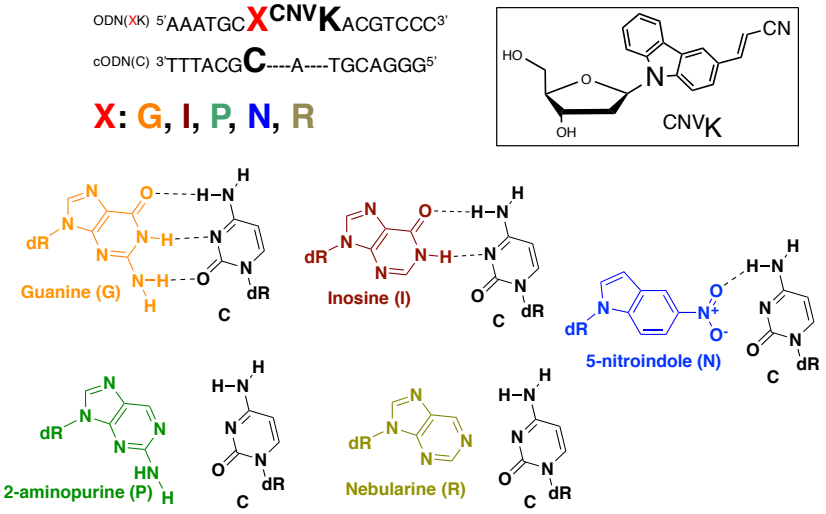


Figure 1.9: Overview of nucleobases used as counter base of cytosine with ODN sequence and ^{CNV}K structure.

In chapter 3, I tried to developed new photo-cross-linkers based on CNVK, 3-methoxycarbonlycarbazole, 3-carboxyvinylcarbazole, and 3-carbonylamidevinylcarbazole. The use of 3-carboxyvinylcarbazole resulted in greater acceleration of the deamination reaction than that achieved with CNVK. The most likely factors affecting the ability of ultrafast photo-responsive nucleosides to accelerate the deamination reaction are polarity and hydrophilicity of the oligodeoxyribonucleotides that contain photo-cross-linker.

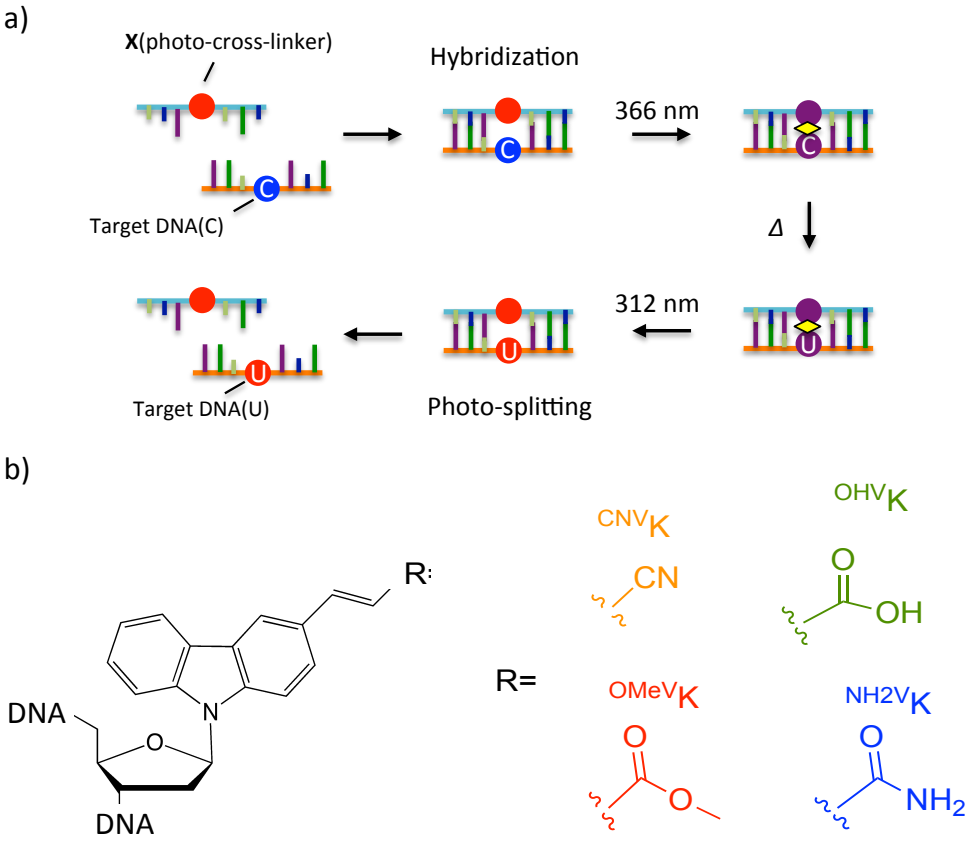


Figure 1.10: a) Photochemical Cytosine to Uracil transition in duplex. X shows photo-cross-linker. b) Structure of 3-vinylcarbazole derivatives.

In chapter 4, I have shown the combined effect of hydrogen bonding and polarity by taking varying combinations of counter-bases (inosine, guanine, and cytosine) and vinylcarbazole derivatives (3-cyanovinyl carbazole, 3-amidovinylcarbazole, and 3-carboxyl vinyl carbazole) and incorporating them in the ODN to accomplish deamination at physiological conditions.

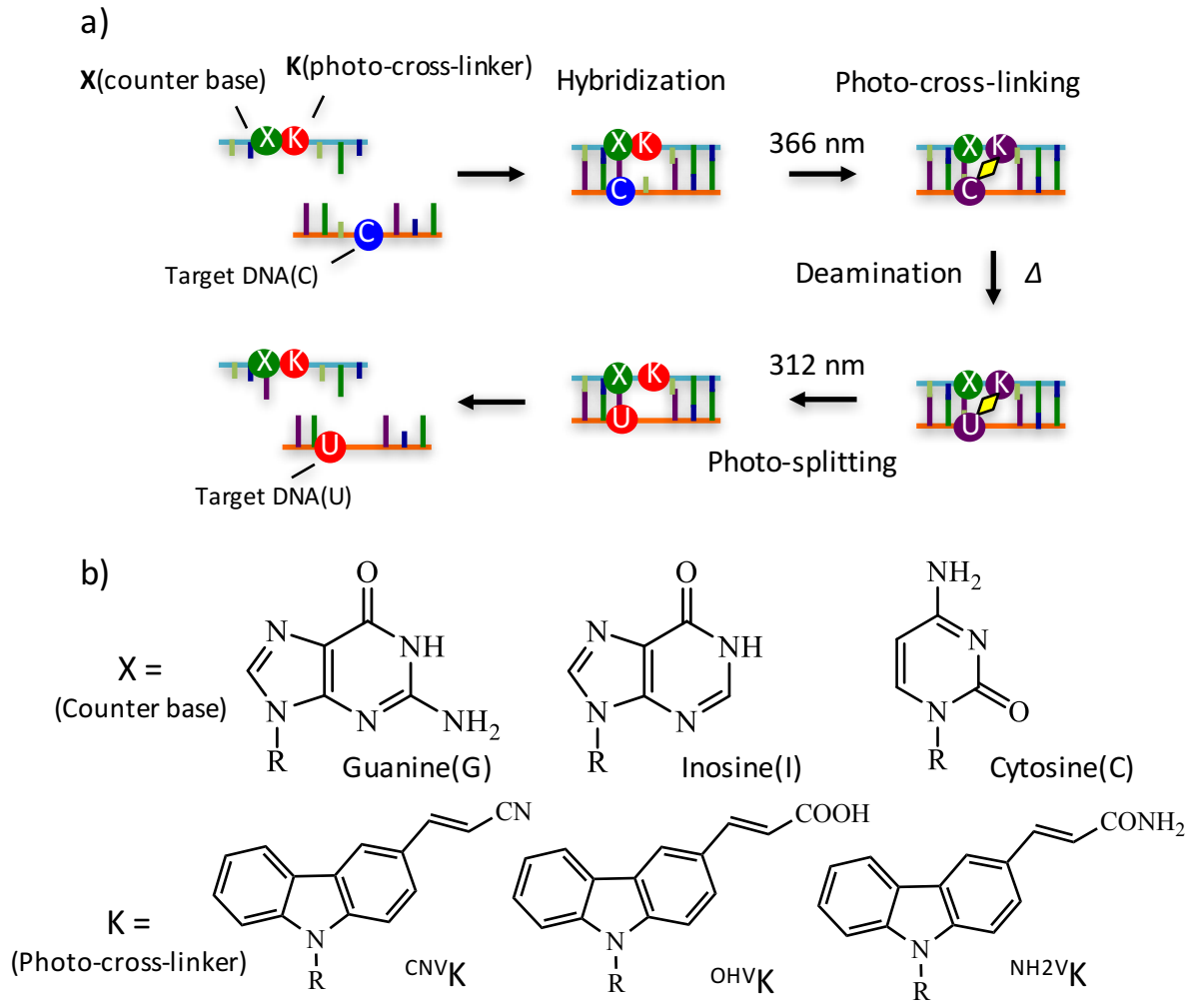


Figure 1.11: a) Overview of photo-cross-linking based DNA editing. b) Structures of counter bases and modified photo-cross-linker.

In chapter 5, I tried a different approach towards changing the micro-environment around the target cytosine by breaking the photo-active ODN into two pieces and introducing a free phosphate group, firstly only on the counter base to check the effect of increased freedom of movement and hydrophilicity and then by further adding the free phosphate group on 3' end of photo-cross-linker. It was observed that upon using phosphate group, the efficiency of deamination reaction at 37ºC has increased drastically as the deamination reaction is now reaching upto 90% completion in merely 6h at physiological conditions.

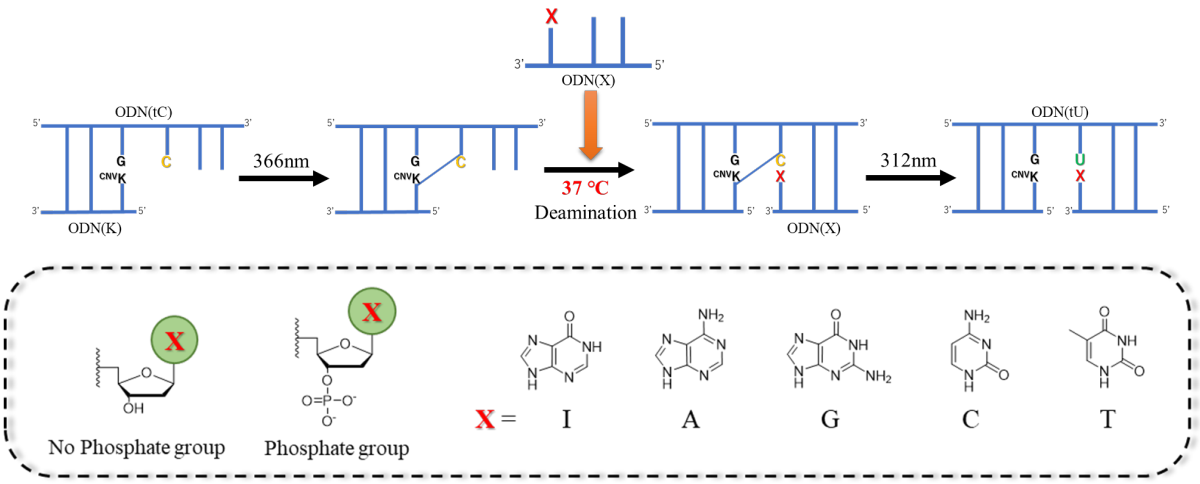


Figure 1.12: Overall view of photo-crosslinking based deamination reaction and structure of modified counter bases.

References

1. Griffith F **(1928)**. *The Journal of Hygiene (London)* 27 (2): 113-59.
2. Hershey A D, Chase M **(1952)**. *J Gen Physiol* 36 (1): 39-56.
3. Ralf D (**2008**). *Hum Genet*, 122, 565-581.
4. Jones M E (**1953**) *Yale. J. Biol. Med*, 26(1), 80-97
5. Levene: Simoni R D, Hill R L, Vaugha M (**2002**). *J. Biol. Chem*, 277(22), 11.
6. Soyfer V N (**2001**) *Nat Rev Genet*, 2, 723-729.
7. Pray L **(2008**) *Nature Education,* 1(1):98
8. Hall K (**2011**). *Stud Hist Phlios Biol Biomed Sci*, 42(2), 119-128.
9. Olby R C (**1994**). *Dover publications inc.*, New York.
10. Watson J D, Crick F H (**1953)** *Nature*, 171, 737-738
11. Crick F H (**1958**). *Symp. Soc. Exp. Biol XII*, 139-163.
12. Sanger F, Nicklen S, Coulson A R (**1977)** *Proc Natl Acad Sci,* 74(12), 5463-5467.
13. Gaastra W (**1984**) *Methods in Molecular Biology*, 2, 333-341
14. Lodish H, Berk A, Zipursky SL, et al. (**2000**). *W. H. Freeman*, New York.
15. Obe G, Pfeiffer P, Savage J R K, Johannes C, Goedecke W, Jeppesen P, Natarajan A T, Martinez-Lopez W, Folle G A, Drets M E, (**2002)**. *Mutation Research* 504 17–36
16. Auerbach C, Robson J M, Carr J G, (**1947**). *Science* 105(2723), 243–7.
17. Kunkel T A, (**1985**) *Proc Natl Acad Sci U.S.A.* 82(2):488-492
18. Nelson D L, Lehninger A L, Cox M M, (**2005**). *W.H. Freeman.* New York
19. Carter P, (**1986**) *Biochem. J.*, 237, 1-7
20. Jones D H, Howard B H, **(1990)** *BioTechniques,* 8, 178-183.
21. Inouye S, Inouye M, **(1987)** *Synthesis and Applications of DNA and RNA* (Narang, S. A., ed.), Academic, New York, 181-206.
22. Lai D, Pestka S, **(1996)** In: Trower M.K. (eds) *In Vitro Mutagenesis Protocols. Methods In Molecular Medicine ™*, vol. 57. Humana Press
23. Walton C R, Booth K, Stockley P G, **(1991)** In: *McPherson MJ (ed.). Directed mutagenesis: A practical approach.* Oxford: IRL Press. 135-162
24. Myers R M, Lerman L S, Maniatis T, **(1985)** *Science*, 229, 242-247
25. Shapiro R, Servis R E, Welcher M, **(1970)** *J. Am. Chem. Soc.,* 90, 422-424
26. Hayatsu H, Wataya Y, Kai K, **(1970)** *J. Am. Chem. Soc.,* 90, 724-726
27. Snow E T, Foote R S, Mitra S, **(1984)** *J. Biol. Chem.*, 259, 8095-8100

28. Tanaka K, Okamoto A, **(2007)** *Bioorg. Med. Chem. Lett.,* 17, 1912-1915
29. Clark S J, Harrison J, Paul C L, Frommer M, **(1994)** *Nucleic Acids Res.*, 22, 2990-2997.
30. Lehninger A L, **(1988),** *Principles of biochemistry,* Worth Publishers Inc.
31. Sidorkina O, Saparbaev M, Laval J, **(1997)** *Mutagenesis,* 12(1), 23-28.
32. Kawanishi M, Matsuda T, Yagi T, **(2014)** *Frontiers in Environmental Science,* 2, 36.
33. Beland F A, Fullerton N F, Heflich R H, **(1984)** *J. Chromatogr*. 308, 121–131.
34. Cheng G, Shi Y, Sturla S J, Jalas J R, McIntee E J, Villalta P W, et al. **(2003)** *Chem. Res. Toxicol*. 16, 145–152.
35. Zhong W, Que Hee S S, **(2004)** *Mutat. Res*. 563, 13–24.
36. Matsuda T, Kawanishi M, Yagi T, Matsui S, Takebe H, **(1998)** *Nucleic Acids Res*. 26, 1769–1774.
37. Matsuda T, Yagi T, Kawanishi M, Matsui S, Takebe H, **(1995)** *Carcinogenesis* 16, 2389–2394.
38. Murata-Kamiya N, Kamiya H, Kaji H, Kasai H, **(1997)** *Nucleic Acids Res*. 25, 1897–1902.
39. Kawanishi M, Matsuda T, Sasaki G, Yagi T, Matsui S, Takebe H, **(1998)** *Carcinogenesis* 19, 69–72.
40. Kawanishi M, Matsuda T, Nakayama A, Takebe H, Matsui S, Yagi T, **(1998)** *Mutat. Res*. 417, 65–73.
41. Lai Y P, Huang J, Wang L F, Li J, Wu Z R, **(2004)** *Biotechnol. Bioeng.,* 86, 622-627
42. Frommer M, McDonald L E, Millar D S, Collis C M, Watt F, Grigg G W, Molloy P L, Paul C L, **(1992)** *Proc. Natl. Acad. Sci. U.S.A.* 80, 1579-1583
43. *Hinnen A, Hicks J B, Fink G R,* **(1978)** *Proc Natl Acad Sci U S A.,* 75(4), 1929-33.
44. Smithies O, Gregg R G, Boggs S S, Koralewski M A, Kucherlapati R S, **(1985)** *Nature.* 317(6034), 230-234.
45. Kostriken R, Strathern J N, Klar A J, Hicks J B, Heffron F, **(1983)** *Cell*, 35, 167-174.
46. Jacquier A, Dujon B, **(1985)** "An intron-encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene." *Cell,* 41, 383-394.
47. Haber J E, **(1998)** *Annu. Rev. Genet.,* 32, 561-599.
48. Klar A J, Strathern J N, Abraham J A, **(1984)** *Cold Spring Harb. Symp. Quant. Biol.,* 49, 77-88.
49. Nickoloff J A, Chen E Y, Heffron F, **(1986)** *Proc. Natl. Acad. Sci. U.S.A.,* 83, 7831-7835.

50. Colleaux L, D'Auriol L, Galibert F, Dujon B, **(1988)** *Proc. Natl. Acad. Sci. U.S.A.,* 85, 6022-6026.
51. Thierry A, Dujon B, **(1992)** *Nucleic Acids Res.,* 20, 5625-5631.
52. Epinat J C, Arnould S, Chames P, Rochaix P, Desfontaines D, Puzin C, Patin A, Zanghellini A, Paques F, **(2003)** *Nucleic Acids Research*. 31(11), 2952-2962.
53. Arnould S, Perez C, Cabaniols J P, Smith J, Gouble A, Grizot S, Epinat J C, Duclert A, Duchateau P, **(2007)** *Journal of Molecular Biology*. **371** (1): 49-65.
54. Seligman L M, Chisholm K M, Chevalier B S, Chadsey M S, Edwards S T, Savage J H, Veillet A L, **(2002)**. *Nucleic Acids Research*. 30(17), 3870-3879.
55. Sussman D, Chadsey M, Fauce S, Engel A, Bruett A, Monnat R, Stoddard B L, Seligman L M, **(2004)**. *Journal of Molecular Biology*. 342(1), 31-41
56. Rosen L E, Morrison H A, Masri S, Brown M J, Springstubb B, Sussman D, Stoddard B L, Seligman L M, **(2006)**. *Nucleic Acids Research*. 34(17), 4791-4800.
57. Arnould S, Chames P, Perez C, Lacroix E, Duclert A, Epinat J C, Stricher F, Petit A S, Patin A, **(2006).** *Journal of Molecular Biology*. 355(3), 443-58.
58. Smith J, Grizot S, Arnould S, Duclert A, Epinat J-C, Chames P, Prieto J, Redondo P, Blanco F J **(2006)**. *Nucleic Acids Research*. **34**(22), 149.
59. Chevalier B S, Kortemme T, Chadsey M S, Baker D, Monnat R J, Stoddard B L, **(2002)**. *Molecular Cell*. 10(4): 895-905
60. Carroll D, **(2008)** *Gene therapy* 15(22) 1463-1468.
61. Tan W S, Carlson D F, Walton M W, Fahrenkrug S C, Hackett P B, **(2012)** *Adv Genet*., 80, 37-97
62. M. Baker, **(2012)** *Nat Meth,* 9(1), 23-26
63. Ramirez C L, Foley J E, Wright D A, et al. **(2008)**. *Nat. Methods*. 5(5), 374-375.
64. Kim Y G, Cha J, Chandrasegaran S, **(1996).** *Proc Natl Acad Sci USA*. 93(3), 1156–60.
65. Bitinaite J, Wah D A, Aggarwal A K, Schildkraut I, **(1998).** *Proc Natl Acad Sci USA*. 95(18), 10570-10575.
66. Cathomen T, Joung J K, **(2008)** *Mol Ther.* 16(7), 1200-1207.
67. Townsend J A, Wright D A, Winfrey R J, Fu F, Maeder M L, Joung J K, Voytas D F, **(2009).** *Nature*. 459(7245), 442-445.
68. Curtin S J, Zhang F, Sander J D, Haun W J, Starker C, Baltes N J, Reyon D, et al. **(2011).** *Plant Physiology*. 156(2), 466-473.

69. Bibikova M, Beumer K, Trautman J, Carroll D, **(2003).** *Science*. 300 (5620): 764.
70. Wood A J, Lo T W, Zeitler B, Pickle C S, Ralston E J, Lee A H, Amora R, Miller J C, et al. **(2011).** *Science*. 333 (6040): 307.
71. Ekker S C, **(2008)**. *Zebrafish*. **5**(2): 1121-1123.
72. Goldberg A D, Banaszynski L A, Noh K M, Lewis P W, Elsaesser S J, Stadler S, Dewell S, Law M, Guo X, Li X, Wen D, Chapgier A, Dekelver R C, Miller J C, Lee Y L, et al. **(2010).** *Cell*. 140 (5): 678–691.
73. Geurts A M, Cost G J, Freyvert Y, Zeitler B, Miller J C, Choi V M, Jenkins S S, Wood A, Cui X, Meng X, Vincent A, Lam S, Michalkiewicz M, Schilling R, Foeckler J, et al. **(2009)**. *Science*. 325(5939): 433–433.
74. Flisikowska T, Thorey I S, Offner S, Ros F, Lifke V, Zeitler B, Rottmann O, Vincent A, Zhang L, Jenkins S, Niersbach, H, Kind A J, Gregory P D, Schnieke A E, Platzer J, **(2011)**. Milstone, David S., ed. *PLoS ONE*. 6(6) 21045.
75. Hauschild J, Petersen B, Santiago Y, Queisser A –L, Carnwath J W, Lucas-Hahn A, Zhang L, Meng X, Gregory P D, Schwinzer R, Cost G J, Niemann H, **(2011)**. *Proc. Nat. Acad. Sci*. 108(29): 12013-12017.
76. Yu S, Luo J, Song Z, Ding F, Dai Y, Li N, **(2011)**. *Cell Research*. 21, 1638-1640.
77. Tebas P, Stein D, Tang W W, Frank I, Wang S Q, Lee Gary, Spratt S K, Surosky R T, Giedlin M A, Nichol G, Holmes M C, Gregory P D, Ando D G, et al. **(2014).** *New England Journal of Medicine*. 370 (10): 901–910.
78. Durai S, Mani M, Kandavelou K, Wu J, Porteus MH, Chandrasegaran S **(2005)**. *Nucleic Acids Res*. 33(18): 5978–90.
79. Kim H, Kin J, **(2014)** *Nature Reviews Genetics,* 15(5), 321-334
80. Boch J **(2011).** *Nature Biotechnology*. 29(2): 135–6.
81. Boch J, Bonas U **(2010)**. *Annual Review of Phytopathology*. 48: 419–36.
82. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A, Bonas U **(2009).** *Science*. 326 (5959): 1509–12.
83. Moscou M J, Bogdanove A J, **(2009)**. *Science*. 326 (5959): 1501.
84. Christian M, Cermak T, Doyle EL, Schmidt C, Zhang F, Hummel A, Bogdanove AJ, Voytas DF **(2010)**. *Genetics*. 186 (2): 757–61.
85. Li T, Huang S, Jiang W Z, Wright D, Spalding M H, Weeks D P, Yang B **(2011)**. *Nucleic Acids Research*. 39 (1): 359–72.

86. Juillerat A, Pessereau C, Dubois G, Guyot V, Marechal A, Valton J, Daboussi F, Poirot L, Duclert A, Duchateau P **(2015)**. *Scientific Reports*. 5: 8150.

87. Zhang F, Cong L, Lodato S, Kosuri S, Church GM, Arlotta P **(2011).** *Nature Biotechnology*. 29 (2): 149–53

88. Joung J K, Sander J D. **(2013)** *Nature reviews Molecular cell biology*. 14(1):49-55.

89. Liu J, Li C, Yu Z, Huang P, Wu H, Wei C, Zhu N, Shen Y, Chen Y, Zhang B, Deng WM, Jiao R, **(2012)** *J Genet Genomics.* 39(5):209-15.

90. Lei Y, Guo X, Liu Y, Cao Y, Deng Y, Chen X, Cheng CH, Dawid IB, Chen Y, Zhao H, **(2012)** *Proc Natl Acad Sci U S A.* 109(43):17484-9.

91. Tesson L, Usal C, Ménoret S, Leung E, Niles BJ, Remy S, Santiago Y, Vincent AI, Meng X, Zhang L, Gregory PD, Anegon I, Cost GJ, **(2011)** *Nat Biotechnol.* 29(8):695-6.

92. Watanabe T, Ochiai H, Sakuma T, Horch HW, Hamaguchi N, Nakamura T, Bando T, Ohuchi H, Yamamoto T, Noji S, Mito T, **(2012)** *Nat Commun.* 3():1017.

93. Carlson DF, Tan W, Lillico SG, Stverakova D, Proudfoot C, Christian M, Voytas DF, Long CR, Whitelaw CB, Fahrenkrug SC, **(2012)** *Proc Natl Acad Sci U S A.* 109(43):17382-7.

94. Li T, Liu B, Spalding MH, Weeks DP, Yang B, **(2012)** *Nat Biotechnol.* 30(5):390-2.

95. Reyon D, Tsai SQ, Khayter C, Foden JA, Sander JD, Joung JK, **(2012)** *Nat Biotechnol.* 30(5):460-5.

96. Hockemeyer D, Wang H, Kiani S, Lai CS, Gao Q, Cassady JP, Cost GJ, Zhang L, Santiago Y, Miller JC, Zeitler B, Cherone JM, Meng X, Hinkley SJ, Rebar EJ, Gregory PD, Urnov FD, Jaenisch R, **(2011)** *Nat Biotechnol.* 29(8):731-4.

97. Nemudryi AA, Valetdinova KR, Medvedev SP, Zakian SM. **(2014)** *Acta Naturae*. 6(3):19-40.

98. Barrangou R **(2015).** *Current Opinion in Immunology*. 32: 36–41

99. Redman M, King A, Watson C, King D **(2016).** *Arch. Dis. Chil. Educ. Prac. Ed*. 101 (4): 213–5.

100. Mohanraju P, Makarova KS, Zetsche B, Zhang F, Koonin EV, van der Oost J **(2016)**. *Science*. 353 (6299): 5147

101. Marraffini LA, Sontheimer EJ **(2010)**. *Nature Reviews Genetics*. 11(3): 181–90.

102. Horvath P, Barrangou R **(2010)**. *Science*. 327 (5962): 167–70.

103. Zhang F, Wen Y, Guo X **(2014).** *Human Molecular Genetics*. 23 (R1): R40–6.

104. Barrangou R **(2015).** *Genome Biology*. 16: 247

105. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E **(2012)**. *Science*. 337 (6096): 816-21.

106. Bolotin A, Quinquis B, Sorokin A, Ehrlich SD **(2005)**. *Microbiology*. 151 (Pt 8): 2551-61.

107. Ledford H **(2016)** *Nature*. 531(7593): 156-9.

108. "CRISPR/Cas9 Plasmids". *www.systembio.com. Retrieved 2018-01-20.*

109. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F **(2013)**. *Nature Protocols*. 8(11): 2281–308.

110. Li J, Shou J, Guo Y, et al. **(2015)** *Journal of Molecular Cell Biology*. 7(4):284-298.

111. Yen S, Zhang M, Deng J, Usman S, Smith C, Parker-Thornburg J, Swinton P, Martin J, Behringer R, **(2014)**. *Developmental Biology,* 393(1), 3-9.

112. Oliver D, Yuan S, McSwiggin H, Yan W **(2015)** *PLoS ONE* 10(6): e0129457.

113. Yang H, Wang H, Shivalila CS, Cheng AW, Shi L, Jaenisch R, **(2013)** *Cell,* 154(6), 1370-1379.

114. Fujimoto K, Matuda S, Takahashi N, Saito I, **(2000)** *J. Am. Chem. Soc*. *122*, 5646-5647.

115. Ogasawara S, Fujimoto K, **(2005)** *Chem. Bio. Chem*., *10*, 1756.

116. Fujimoto K, Matuda S, Ogawa N, Hayashi M, Saito I, **(2000)** *Tetrahedron Lett*., *41*, 6451-6454.

117. Nakamura S, Ogasawara S, Matsuda S, Saito I, Fujimoto K, **(2012)** *Molecules*. 17, 163-178.

118. Nakamura F, Ito E, Sakao Y, Ueno N, Gatuna I N, Ohuchi F S, Hara M, **(2003)** *Nano. Lett., 3*, 1083-1086

119. Becerril H A, Stoltenberg R M, Wheeler D R, Davis R C, Harb J N, Woolley A T, **(2005)** *J. Am. Chem. Soc*., *127*, 2828-2829.

120. Ogasawara S, Ami T, Fujimoto K, **(2008)** *J. Am. Chem. Soc.*, 130, 10050-10051.

121. Matsumura T, Ogino M, Nagayoshi K, Fujimoto K, **(2008)** *Chem. Lett.,* 37(1), 94-95

122. Fujimoto K, Yamada A, Yoshimura Y, Tsukaguchi T, Sakamoto T, **(2013)** *J. Am. Chem. Soc.* 135 (43), 16161-16167

123. Fujimoto K, Hiratsuka-Konishi K, Sakamoto T, Ohtake T, Shinohara K-I, Yoshimura Y, **(2012)** *Mol. BioSyst.* 8, 491.

124. Fujimoto K, Hiratsuka-konishi K, Sakamoto T, Yoshimura Y, **(2010)** *Chem. Commun.* 46, 7545-47

125. Fujimoto K, Hiratsuka-konishi K, Sakamoto T, Yoshimura Y, **(2010)** *Chem.Biochem.,* 11, 1661-64

126. Matsumura T, Ogino M, Nagayoshi K, Fujimoto K, **(2008)** *Chem. Lett.* 37:1, 94-95

127. Fujimoto K, Matsuda S, Yoshimura Y, Matsumura T, Hayashi M, Saito I, **(2006)** *Chem. Commun.*, 0, 3223.

128. Sakamoto T, Shigeno A, Ohtaki Y, Fujimoto K, **(2014)** *Biomater. Sci.,* 2, 1154

129. Frederico L A, Kunkel T A, Shaw B R, **(1990)** *Biochemistry* 29, 2532-37.

Chapter 2: Effect of nucleobase change on cytosine deamination through DNA photo-cross-linking reaction via 3-cyanovinylcarbazole nucleoside

Introduction

Genome editing is an emerging technique in the field of genetic engineering. Genome editing using engineered nucleases (GEEN) has been awarded "method of the year, 2011" by Nature Methods.[1] Major methods available for genome editing are meganuclease,[2] zincfinger nuclease (ZFN),[3] transcription activator-like effector-based nuclease (TALEN),[4] and CRISPR-Cas system.[5] These methods have been considered better than the classic site-directed mutagenesis, which include PCR-plasmid and recombination based methods as the classical methods cannot be used for mammalian cells.[6] GEEN based genome editing exploits the double strand break, using engineered nucleases, and repair mechanism of cellular machinery to introduce mutation.[2] Compared to normally used endonucleases, the engineered nucleases like meganucleases, ZFN, and TALEN are capable of creating site-specific double stranded breaks in the genomic DNA.[7,8] While engineered meganucleases are engineered by modifying DNA interacting amino acids, it is very expensive to produce sequence specific enzymes for all possible sequences.[9] ZFN and TALEN are based on cutting DNA recognition and cleaving portions of enzymes and joining the cleaving part with another sequence recognition peptide providing specificity.[10] In CRISPR/Cas system include plasmids with CRISPR sequence and Cas9 (CRISPR associated proteins) which can be incorporated into eukaryotic cells to induce desired editings in genome.[11] These methods have the disadvantages such as ZFN have been shown to be toxic to the cellular system[12] and require enzymes which are teadious and expensive to produce.

In 2010, we reported an enzyme-free reaction for cytosine deamination, which use 3-cyanovinylcarbazole nucleoside (^{CNV}K) mediated photo-cross-link based deamination reaction.[13] This reaction involves the use of an artificial nucleobase, 3-cyanovinylcarbazole, which upon incorporation in oligodeoxyribonucleotide (ODN), can make the ODN photo-responsive towards UV radiation. The major drawback of this protocol is that the deamination step takes place at elevated temperature (90 ºC) to afford the cytosine deamination, which is not suitable for cellular applications. In our previous study, it has been shown that the rate of ^{CNV}K based photo-cross-linking is affected by number of hydrogen bonds between the counter base and target pyrimidine.[14] Thus, in this study, role of hydrogen bonding has been evaluated to improve the reactivity of the photo-cross-link mediated cytosine deamination at physiological conditions by using inosine (I), 2-aminopurine (P), nebularine (R), and 5-nitroindole (N), in place of guanine (G), in the photo responsive ODN as counter-base of target cytosine.

Materials and Methods

Preparation of oligonucleotide. The phosphoramidite of ^{CNV}K was synthesized based on the method prescribed in the literatute.[15] Phosphoramidite monomers of adenine, cytosine, thymidine, guanine, deoxyinosine, 2-aminopurine, nebularine, and 5-nitroindole were purchased form GlenResarch (USA). Oligodeoxyribonucleotides were synthesized using automatic DNA synthesizer (3400 DNA Synthesizer, Applied BioSystems, USA) and purified by reverse phase HPLC (PU-980, HG-980-31, DG-980-50, UV-970system (Jasco, Japan) equipped with Cosmosil® 5C18-AR-II column (5 µm, 10 I.D. x 150 mm, Nacalai Tesque, Japan)) with a gradient of 2-32% MeCN and flow rate of 3 ml/min, 60ºC. The synthesized ODNs were confirmed by MALDI-TOF-MS (Table 2.1). Non-modified ODNs were purchased form Fasmac (Japan).

Table 2.1: MALDI-TOF-MS analysis of synthesized ODNs

ODN(XK), X=	Calculated Mass $(M+H)^+$	Experimental Mass $(M+H)^+$
G	4625.84	4625.90
I	4610.97	4608.23
P	4609.94	4609.55
R	4594.83	4607.69
N	4639.12	4629.48

Photo-cross-linking Reaction. The reaction mixture containing 15-mer oligonucleotides (15 µM each in 50 mM Na-cacodylate buffer (pH = 7.4) containing 100 mM NaCl) was irradiated with an LED lamp (366 nm, 1600 mW/cm^2, ZUV, Omron, Japan) at 0 ºC for 40 s. Photo-cross-linked ODNs were analysed and purified by reverse phase HPLC (PU-980, HG-980-31, DG-980-50, UV-970system (Jasco, Japan) equipped with Cosmosil® 5C18-AR-II column (5 µm, 10 I.D. x 150 mm, Nacalai Tesque, Japan)) with a gradient of 2-32% MeCN and flow rate of 3 ml/min, 60ºC.

Deamination Reaction. The solution of photo-cross-linked DNAs (4 µM in 50 mM Na-cacodylate buffer (pH = 7.4) and 100 mM NaCl) was incubated at 37 ºC. Photo-splitting

reaction was performed using 312 nm UV Trans-illuminator (Funakoshi Ltd., Japan) at 37ºC for 15min.

UPLC analysis. Samples after photo-splitting were subjected ultra-high performance liquid chromatography (UPLC) using UPLC system (Acquity, Waters, USA) equipped with BEH Shield RP18 column (1.7 µm, 2.1 x 50 mm, elution was with 0.05 M ammoniumformate containing 3-6.5% CH_3CN, linear gradient (10 min) at a flow rate of 0.2 ml/min, 60 ºC). Chromatograms were monitored at 260 nm.

UV melting analysis. UV melting curve of the duplexes ODN(XK)/cODN(C) and photo-cross-linked ODNs (5 µM in 50 mM Na-cacodylate buffer (pH 7.4) consisting of 100 mM NaCl) were measured (260 nm, 1 ºC/min) by V-630bio spectrophotometer (Jasco, Japan), equipped with temperature controller. The melting temperature (T_M) of the duplex was calculated using first derivate of the corresponding melting curve (Fig. 2.1).

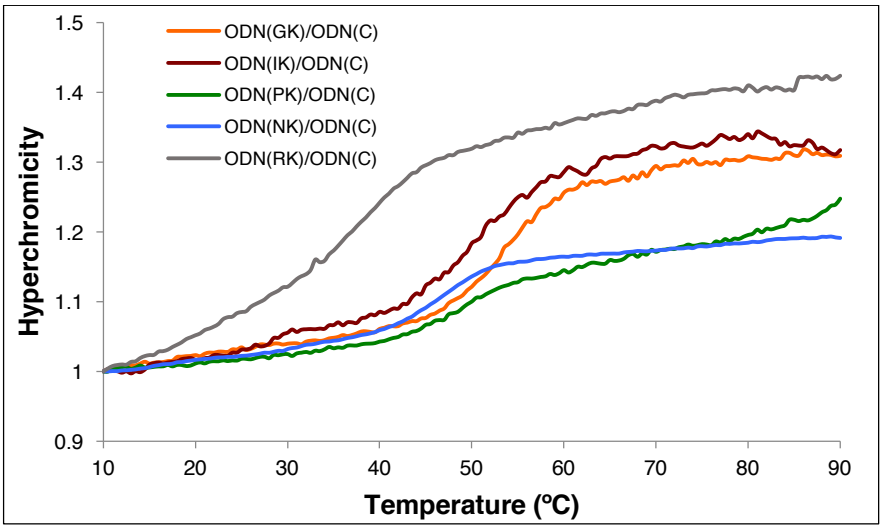


Figure 2.1: UV melting curves at 260nm of ODN(XK)/cODN(C)

Molecular modelling: Molecular modelling analysis was performed using Macromodel v8.1[16] with AMBER* force field, water as solvent, and H-bond parameter 2.5 Å constrain. Energy minimization (500 iterations), both before and after stochastic molecular dynamics, were performed to find the most stable structure. The details of the molecular modelling:

Macromodel v8.1

Maestro v5.106 ©Schrodinger LLC

Dynamics method: Stochastic, no shaking

Simulation temperature: 300 K

Time step: 1.500 fs

Equilibriation time: 1.0 ps

Simulation time: 10.0 ps

Minimization method: PRCG

Max # of iterations: 500

Converge on: Gradient

Convergence threshold: 0.0500

Force field: AMBER*

Solvent: Water

Electrostatic treatment: constant dielectric

Dielectric constant: 1.0

Charges from: force field; extended cut off

Vanderwaals: 8.0

Electrostatic: 20.0

H-bond: 4.0

Results and discussion

Photo-cross-linking

To evaluate the effect of hydrogen bonding on the rate of photo-cross-linking, photo-responsive ODNs with different counter bases (G, I, P, N, and R) and ^{CNV}K at -1 position to cytosine were prepared and subjected to photo-cross-linking reaction with cODN(C) by irradiating the samples, containing 15 uM of each ODN along with dU (75 uM) as internal standard in 50 mM sodium cacodylate buffer (pH 7.4) containing 100 mM NaCl, with 366 nm UV radiation and analyzing by ultra-high performance liquid chromatography (UPLC). The UPLC chromatograms (Fig. 2.2) shows that initially two peaks corresponding to the ODN(XK) and cODN(C) had been decreased during the course of photo-cross-linking reaction and new peak corresponding to photo-cross-linked duplex (ODN(XK<>C)) has appeared. After 2 s, the reaction has proceeded to >60% conversion in the case of ODNs with 2-aminopurine (P) and ca. 50% with nebularine (R). In the case of inosine (I), the conversion after 2 s is ca. 40% while in case of guanine (G) the conversion is ca. 10% and for 5-nitroindole (N), the conversion is even less (<8%). This suggests that the photo-reactivity of ^{CNV}K is different towards the cytosine having different counter base and the base pairing of cytosine can have either accelerating effect towards the photo-cross-linking (I, P, and R) or it can impede the photo-cross-linking (N) compared to ODN(GK). A plot to study quantitatively the effect of counter base substitution on the rate of photo-cross-linking is shown in Fig. 2.3a to depict the time course of photo-cross-linking reaction, wherein the dots represent the actual points of the reaction and the solid lines represent the fitted curves based on first order reaction kinetics and the reaction rate constants has been plotted (Fig. 2.3b) to ascertain the effect of base substitution on photo-cross-linking. Considering the rate of reaction with ODN(GK) and cODN(C) as 1, the acceleration of reaction has been observed when G is replaced with I (3.5x), R (4x), and P (7.8x) while deceleration is observed in the rate of reaction with N having the relative rate constant 0.6. These figures suggest that only the number of hydrogen bonds is not sufficient to predict the relative reaction rate for photo-cross-linking reaction but other factors are also involved which affect the ^{CNV}K mediated photo-cross-linking on changing the target cytosine's counter base. These factors include, but are not limited to, number of hydrogen bonds, stacking distance between ^{CNV}K and cytosine, and overall duplex stability.

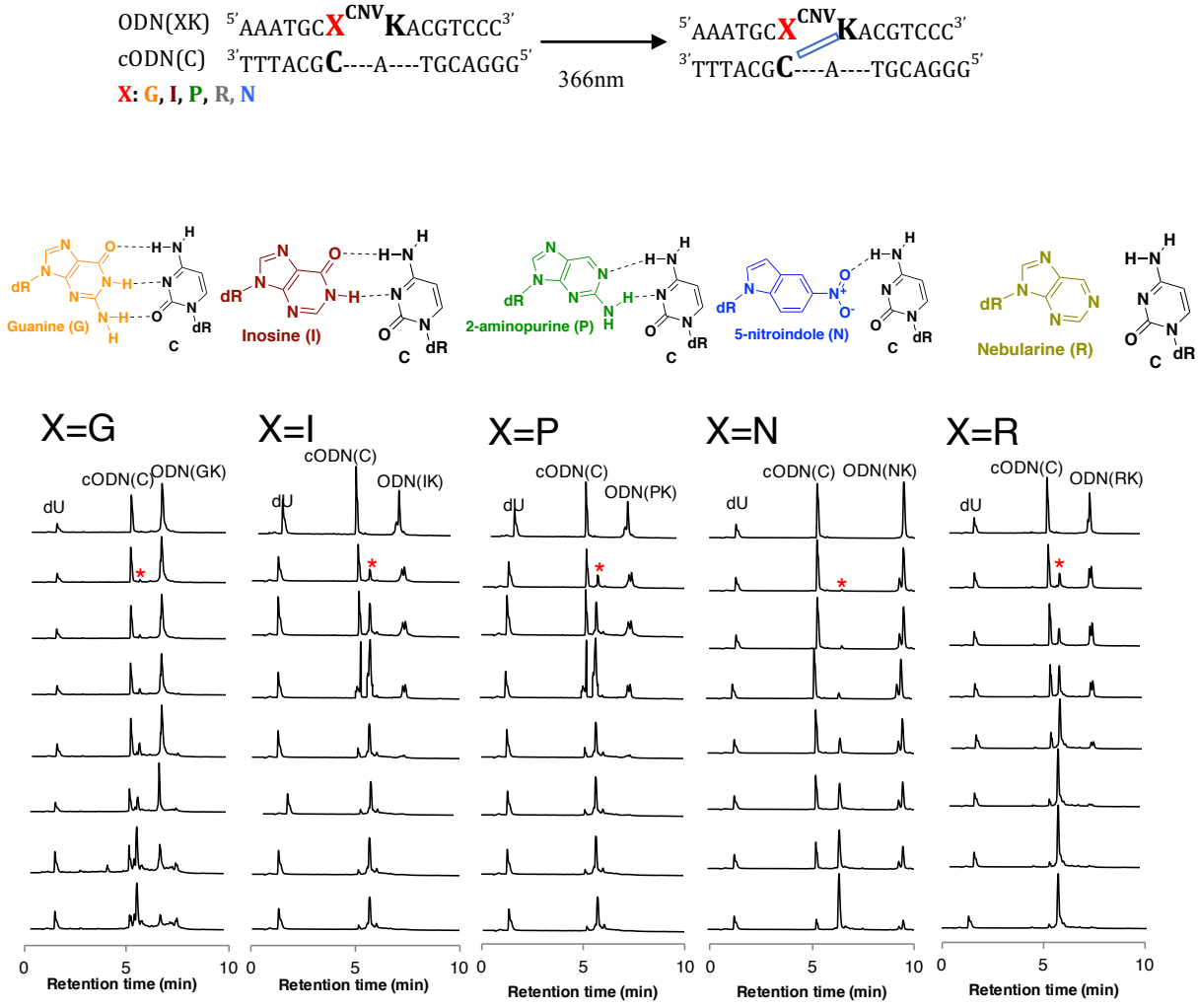
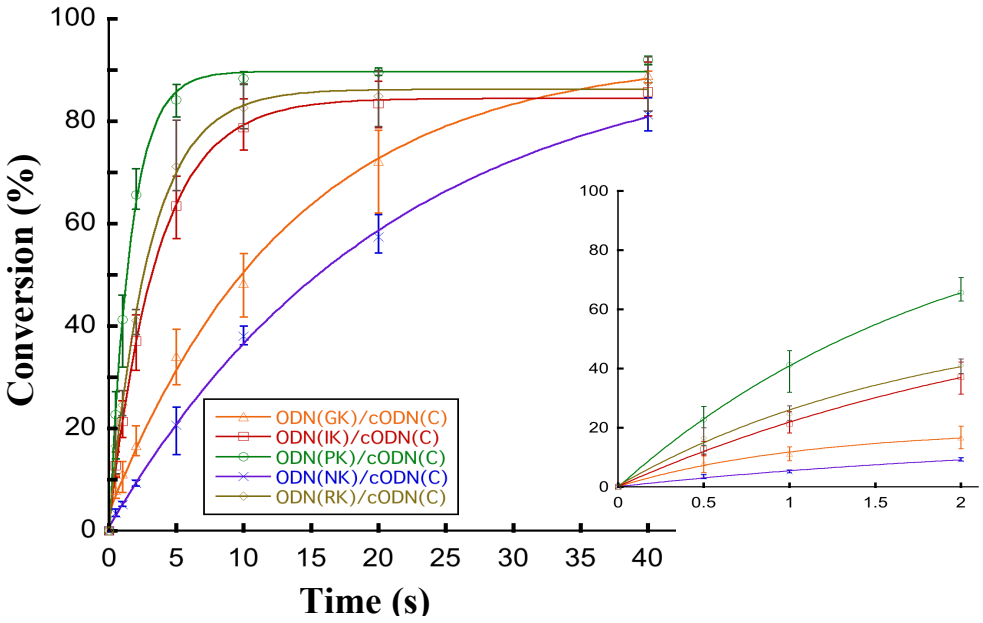


Figure 2.2: UPLC analysis of the photo-cross-linking reaction of the duplex consisting of ODN(XK) and cODN(C). [ODN(XK)]=[ODN(C)] = 15 μM in 50 mM Na-cacodylate buffer (pH 7.4) containing 100 mM NaCl with [dU]=7.5 uM as internal standard. Samples were irradiated with UV-LED (366 nm, 1600 mW/cm^2) at 0 °C. Asterisks over the new peak indicate the photo-cross-linked duplex. "dR" indicates 2-deoxyribose. Refer Table 2.2 for hydrogen bond and stacking distance.

a)



b)

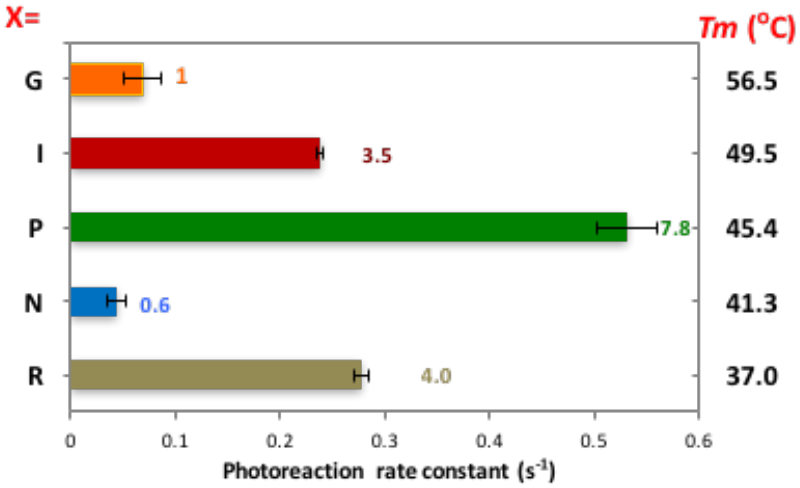


Figure 2.3: (a): Time course of the photo-cross-linking reaction of the duplex containing ODN(XK) and cODN(C), where X=G, I, P, R, and N. (b) Photo-cross-linking reaction rate constant and melting temperature (Tm) of the duplex. The numbers besides the bar indicate the relative reaction rate constant considering the reaction rate constant of photo-cross-linking of ODN(GK)/cODN(C) as base value. Refer Table 2.3 for detailed Reaction rate constant.

Table 2.3: Photo-reaction rate constants of duplex consisting of ODN(XK)/ODN(C)

ODN(XK)	k (s^{-1})	Acceleration ratio (k_X/k_G)	Stacking distance (Å) (average)
ODN(GK)	0.068 ±0.007	1	3.43
ODN(IK)	0.238 ±0.009	2.7	3.44
ODN(PK)	0.530 ±0.029	5.8	3.54
ODN(NK)	0.044 ±0.003	0.6	4.42

ODN(RK)	0.278 ±0.018	3.5	3.55

Guanine, with three hydrogen bonds to cytosine, shows the maximum stability in the duplex, with Tm=56.5 ºC. Therefore, the reaction rate of guanine is comparatively lower than other base pairs except 5-nitroindole. In 5-nitroindole, the distance between C5-C6 bond of cytosine and vinyl bond of ^{CNV}K, is larger as compared to other ODNs, 4.7 (C1' (^{CNV}K)-C6 (cytosine)) and 4.2 Å (C2' (^{CNV}K)-C5 (cytosine)) as compared to 3.6 and 3.3 Å in ODN(GK) and cODN(C), which can be based on stacking of the 5-nitroindole moiety over ^{CNV}K molecule due to non-polar van-der-waal interactions, which leads to lower reaction rate. Nebularine showing acceleration to photo-cross-linking reaction, is related to the absence of hydrogen bonding in the nebularine. In case of 2-aminopurine, the change in hydrogen bonding pattern before and after photo-cross-linking lead to highest rate of photo-cross-linking. The acceleration in the reaction rate is slightly higher in case of R (4.0) than I (3.5) due to lack of hydrogen bonding in case of ODN(RK)/cODN(C) which provide greater freedom of movement to cytosine to form cross-link with ^{CNV}K. 2-aminopurine has the largest relative acceleration towards the photo-cross-linking. It can be explained based on the fact that the number of hydrogen bonds changed before and after the cross-linking

Deamination of photo-cross-linked cytosine

To evaluate the role of hydrogen bonding in the deamination of cytosine, photo-cross-linked duplex (ODN(XK<>C)) were incubated at 37 ºC, and then irradiated (312 nm) to split the cross-linked duplex. Resulting solutions were analyzed by UPLC and the conversion of the deamination reaction was evaluated from the ratio of the peak areas identical to cODN(C) and cODN(U) generated, respectively (Fig. 2.5).

ODN(PK<>C), ODN(RK<>C) and ODN(NK<>C), when incubated at 37ºC, does not lead to deamination of cytosine even in 72 hours, whereas, ODN(IK<>C) shows 35% C→U conversion as compared to 6% in case of ODN(GK<>C) after 72 h incubation at 37ºC (Fig.

2.6a). Twelve times acceleration in the cytosine deamination rate has been observed in the case of ODN(IK<>C) with respect to ODN(GK<>C) (Fig. 2.6b).

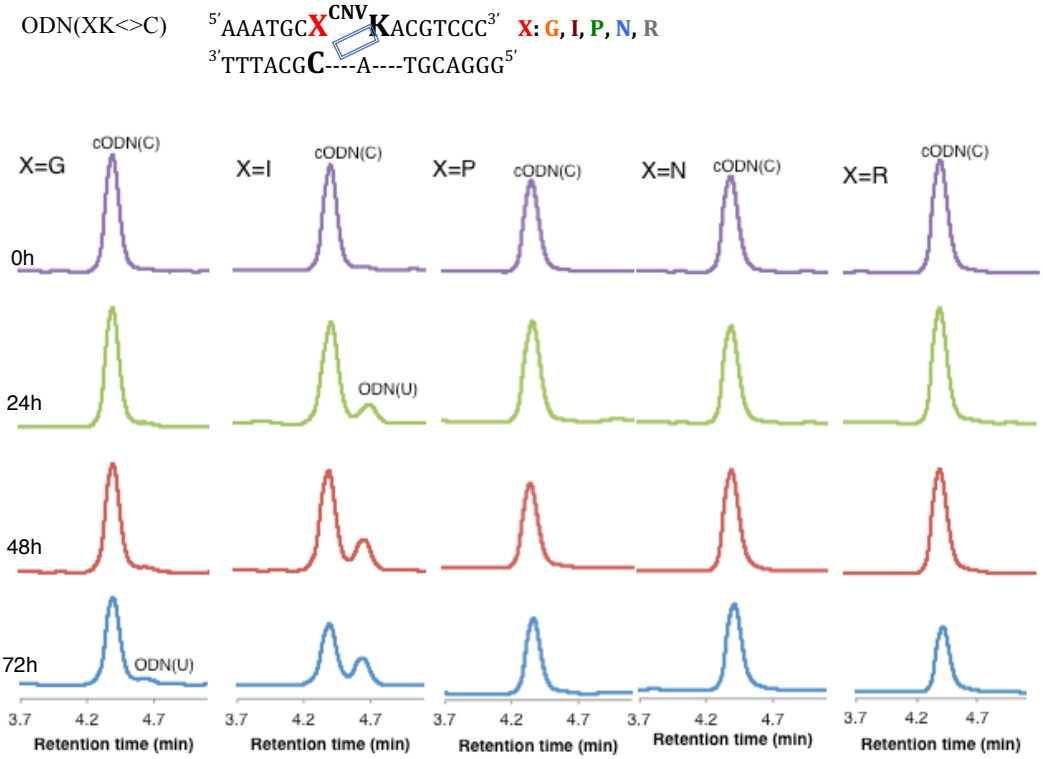
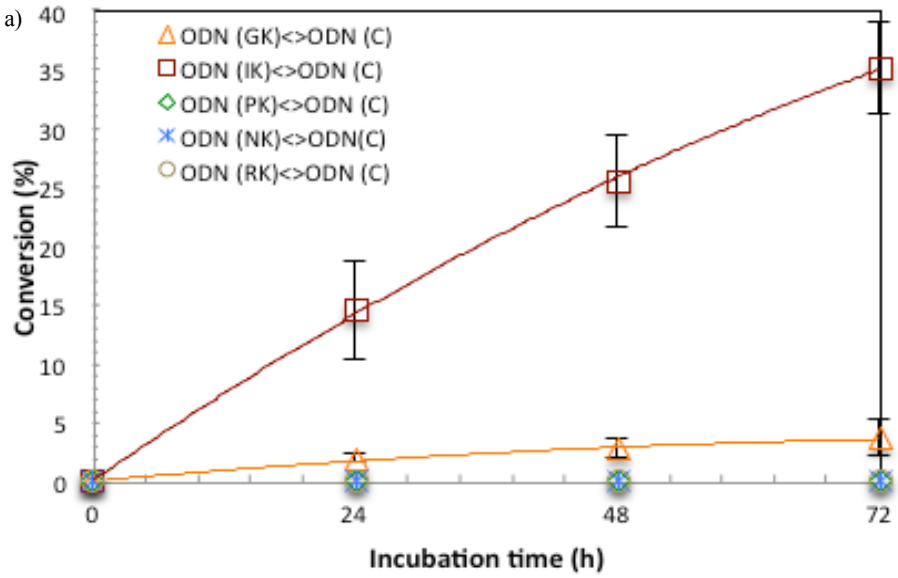


Figure 2.5: UPLC analysis of the deamination reaction of the photo-cross-linked duplex consisting of ODN(XK) and cODN(C). [ODN(XK<>C)] = 15 μM in 50 mM Na-cacodylate buffer (pH 7.4) containing 100 mM NaCl. incubated at 37ºC, photo splitting was performed with transilluminator (312 nm) at 37ºC. cODN(U) indicate the deaminated product.



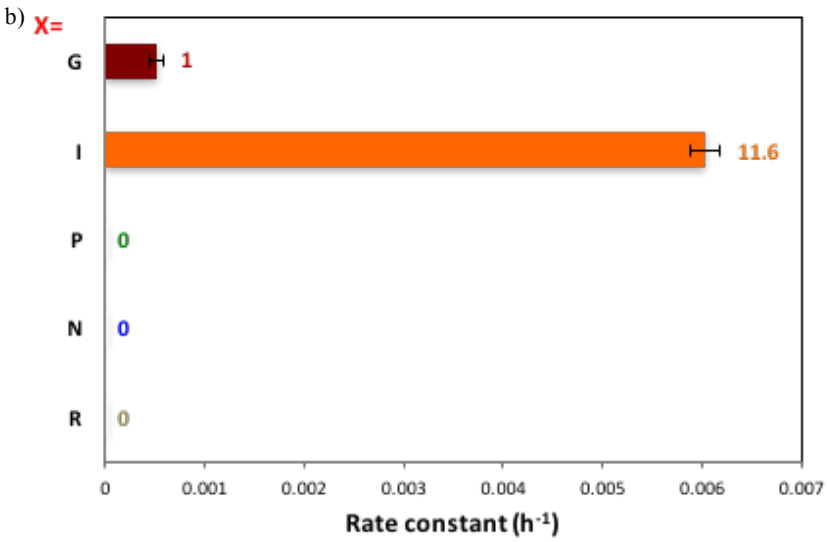


Figure 2.6: (a) Time course of the deamination reactions of the photo-adducts ODN(XK<>C). (b) Deamination reaction rate constants of the photo-adducts ODN(XK<>C). The values besides the bars indicate the acceleration ratio compared to the case of X = G. Photoreaction rate constants were estimated from the time course of the deamination reaction with an assumption of first-order reaction kinetics. Refer Table 2.4 for reaction rate constants.

Table 2.4: Reaction rate constants of deamination reaction

ODN(XK)	Rate constant, k (h^{-1})	Acceleration ratio (k_X/k_G)
ODN(GK)	$5.17 \times 10^{-4} \pm 7.500 \times 10^{-5}$	1.0
ODN(IK)	$6.02 \times 10^{-3} \pm 1.471 \times 10^{-4}$	11.6
ODN(PK)	0	0.0
ODN(NK)	0	0.0
ODN(RK)	0	0.0

These observations can be attributed to the thermodynamics involved in the formation of the pseudo reaction intermediate as per the mechanism of the deamination reaction (Fig 2.7). The difference in the energy of the photo-adduct and pseudo intermediate is a deciding factor for the rate of deamination in case of inosine containing ODN. The energy of photo-adduct in case of ODN(GK<>C) and ODN(IK<>C) is not much different, ODN(GK<>C) has slightly lower energy due to two hydrogen bonds as compared to ODN(IK<>C) with one hydrogen bond (fig 2.8), but there is large difference in the energy of pseudo intermediates thus giving rise to a large difference in ΔE values (Table 2.5). This difference in the ΔE values leads to higher rate of deamination reaction in case of ODN(IK<>C) as compared to ODN(GK<>C).

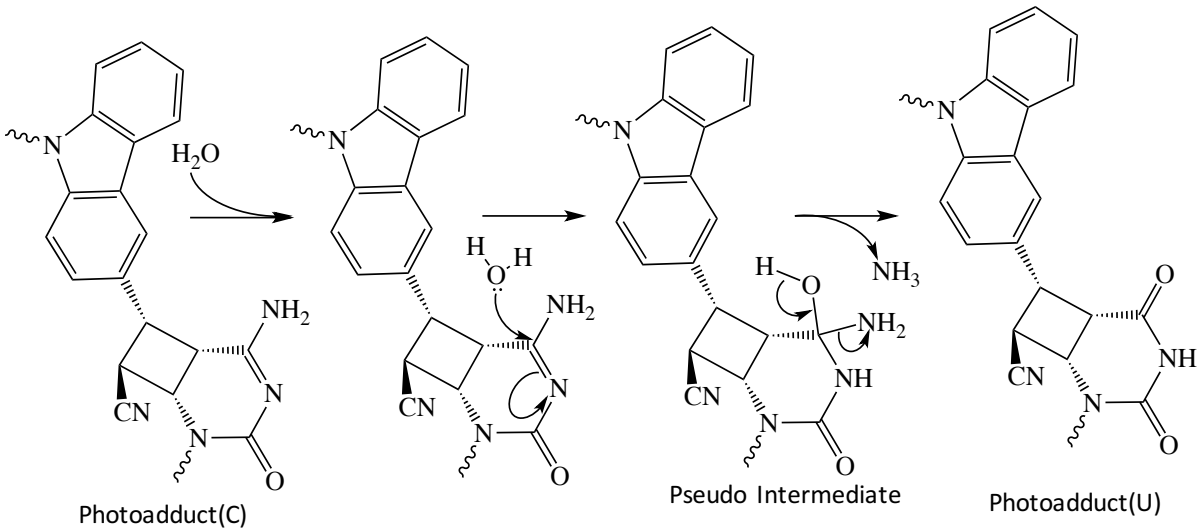


Figure 2.7: Proposed mechanism of deamination of photo-adduct

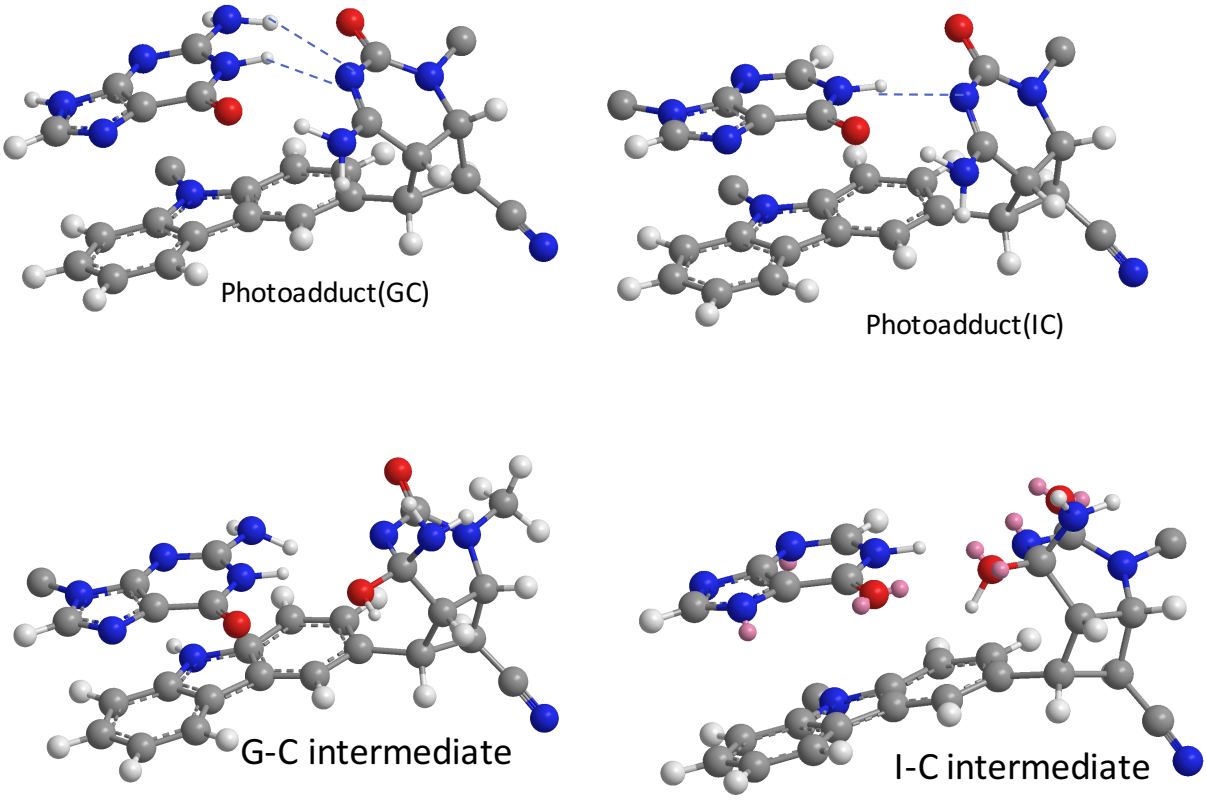


Figure 2.8: Molecular models showing number of hydrogen bonds in photo-adducts and reaction intermediates.

Table 2.5: Energy of photo-adduct and pseudo intermediated

	G-C base pair	I-C base pair
Photoadduct(C)	281	274
Intermediate*	379	335
ΔE	92	61
Reaction rate (h^{-1})	6.0×10^{-4}	5.1×10^{-4}

* Pseudo intermediate [kcal/mol]

Whereas, no deamination reaction in the case of photo-adducts with P, N, and R could be due to the higher hydrophobicity of the ODNs which might dampen the deamination reaction. Moreover, comparing 2-aminopurine and inosine for hydrogen bonding patterns, it is evident that the hydrogen bond in case of inosine with cytosine is more hydrophilic in nature as the

hydrogen bond is present between amino group of cytosine and carboxylic group of inosine giving rise to higher polarity around the amino group of cytosine whereas, in case of 2-aminopurine, the hydrogen bond of amino group is formed with endo-cyclic imine nitrogen which is part of the aromatic ring leading to higher hydrophobicity around the C4 carbon of cytosine (carbon attached to amino group) thereby making the attack of incoming nucleophile difficult (Fig 2.9). Similarly in case of nebularine, the absence of hydrophilic group and hydrogen bonding with cytosine leads to higher hydrophobicity around cytosine leading to hampering of deamination reaction. Also, in case of 5-nitroindole, which being the most hydrophobic among all counter bases, also doesn't show any deamination reaction. Upon comparing these hydrogen bond patterns with the active site of yeast cytosine deaminase (yCD)[17-20], it has been observed that in case of yCD, the cytosine from a special hydrogen bonding network in the active site of the enzyme wherein, Glu64 forms hydrogen bonding with hydrogen of amino group of cytosine leading to higher hydrophilicity around cytosine which is further complemented by attack of Zn-bound activated nucleophile (Fig 2.10).

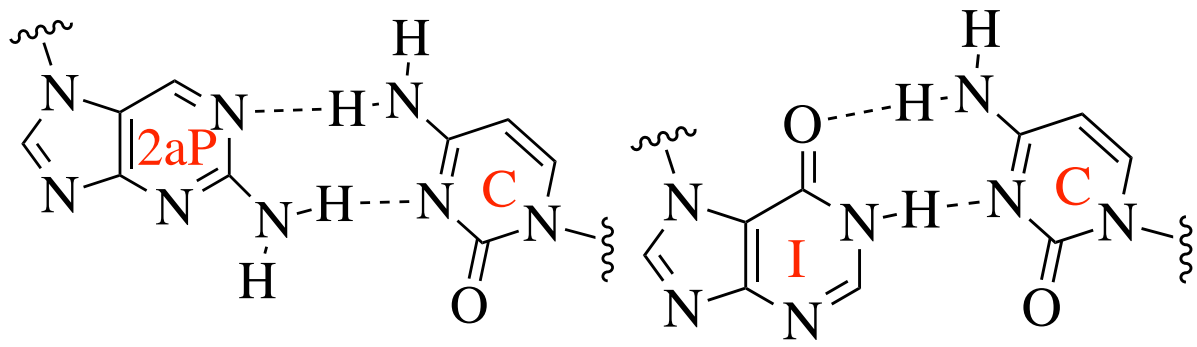


Figure 2.9: Hydrogen bonding comparison in 2-aminopurine and inosine

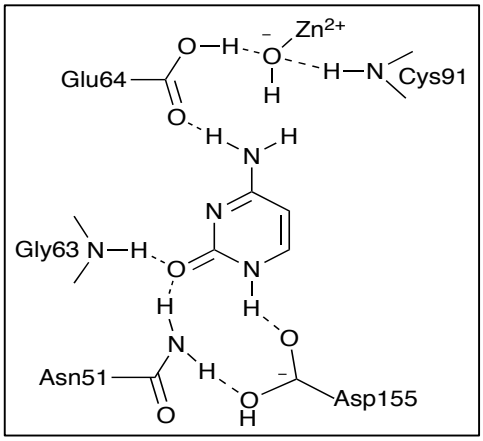


Figure 2.10: Active site representation of yeast cytosine deaminase

Inosine as complimentary base to cytosine for ^{CNV}K assisted photo-cross-linking mediated deamination of cytosine, due to proper spatial hydrogen bonding patterns lower energy requied for intermediate formation make it suitable for this reaction. The type of hydrogen bonds and spatial arrangement of cytosine are crucial factors for the deamination of cytosine. The ^{CNV}K mediated photo-cross-linking induced deamination reaction could prove to be one the revolutionary technique for *in vivo* DNA manipulation due to its simplicity and specificity.

Conclusions

Inosine has potential to be used as the counter-base of target cytosine for the ^{CNV}K mediated photo-cross-link induced deamination of cytosine in physiological conditions as it clearly accelerate the reaction to many-folds when used in the photo-responsive ODN. Hydrogen bonding of target cytosine plays a critical role in the deamination reaction along with the hydrophobicity of counter base. The major factors involved in the deamination reaction are position and number of hydrogen bonds between the counter base and target cytosine along with the spatial orientation of the photo-cross-linked cytosine.

References

1. Method of the year 2011, *Nat. Met.*, (**2012**), 9(1), 1
2. Esvelt K M, Wang H H, **(2013)** *Molecular Systems Biology,* 9(641), 1-17
3. Tan W S, Carlson D F, Walton M W, Fahrenkrug S C, Hackett P B, **(2012)** *Adv Genet*., 80, 37-97
4. Puchta H, Fauser F, **(2013)** *Int. J. Dev. Biol.,* 57, 629-637
5. Sanger J D, Joung J K, **(2014)** *Nature Biotchnology,* 32, 347-355
6. Storici F, Lewis L K, Resnick M A, **(2001)** *Nature Biotechnology,* 19(8), 773-776
7. de Souza N, **(2011)** *Nat Meth*, 9(1), 27
8. Smith J et al., **(2006)** *Nucleic Acids Res.,* 34(22), e149
9. Smith J J, Jantz D, Hellinga H W, **(2011)** US patent, US8021867 B2.
10. Baker M, **(2012)** *Nat Meth,* 9(1), 23-26
11. Hemphill J, Borchardt E K, Brown K, Asokan A, Deiters A, **(2015)** *J. Am. Chem. Soc.*, 137(17), 5642–5645
12. Kim H, Kin J, **(2014)** *Nature Reviews Genetics,* 15(5), 321-334
13. (a) Fujimoto K, Hiratsuka-konishi K, Sakamoto T, Yoshimura Y, **(2010)** *ChemBioChem,* 11, 1661-1664; (b) Fujimoto K, Hiratsuka-konishi K, Sakamoto T, Yoshimura Y, (**2010)** *Chem. Commun.*, 46, 7545-7547
14. Sakamoto T, Ooe M, Fujimoto K, **(2015)** *Bioconjugate Chem.*, 26, 1475-1478
15. Yoshimura Y, Fujimoto K, **(2008)** *Org. lett.,* 10, 3227-3230
16. MacroModel, Schrödinger LLC. **2004.** http://www.schrodinger.com.
17. Sklenak S, Yao L, Cukier R I, Yan H, **(2004)** *J. Am. Chem. Soc.*, 126(45), 14879-14889
18. Yao L, Sklenak S, Yan H, Cukier R I, **(2005)** *J. Phys. Chem. B*, 109(15), 7500-7510
19. Manta B, Raushel F M, Himo F, **(2014)** *J. Phys. Chem. B*, 118(21), 5644-5652
20. Zhang X, Zhao Y, Yan H, Cao Z, **(2016)** *J. Computational Chem.*, 37(13), 1163-1174

Chapter 3: Effect of substitution of photo-cross-linker in photochemical cytosine to uracil transition in DNA

Introduction

Current attempts to treat various genetic disorders by gene therapy are focused predominantly on DNA/RNA nucleobase editing[1,2] Most of these extensively researched techniques are based on either enzymatic[3-9] (CRISPR/Cas9 and deaminase) or chemical[10,11] (bisulphite) DNA modifications. Enzymatic techniques require the presence of a specific sequence in the genome to act. CRISPR/Cas9 is usually composed of two major components, namely, Cas9 protein and guide RNA which can recognize a specific sequence of the target DNA. Guide RNAs are complementary to the target DNA and can be engineered based on the sequence of the latter. Cas9 protein is an obligatory nuclease that recognizes protospacer adjacent motif (PAM) sequence (NGG) in the target genome and cleaves DNA 20bp away from the PAM site. Enzymatic gene editing technology is advantageous because of its gentler nature. However, there are numerous limitations associated with it, including complicated handling and limited utility *in vivo*. On the other hand, chemical methods of nucleobase editing require the use of invasive and hazardous chemicals, such as sodium bisulphite, which may cause unexpected, nonspecific reactions in the body. Therefore, noninvasive, highly specific, and enzyme-free technologies that can bring about specific changes in DNA or RNA without any nonspecific outcomes and associated side effects are urgently required to be able to use new approaches to cure genetic disorders.

We have previously reported a non-enzymatic method to accomplish cytosine to uracil (C to U) transformation in target DNA and RNA, by using photo-responsive oligodeoxyribonucleotides (ODNs) containing 3-cyanovinylcarbazole nucleoside (^{CNV}K).[13–15] The ODN containing ^{CNV}K can photo-cross-link to pyrimidine base in complementary DNA or RNA strand with a few second photoirradiation[16]. The deamination of cytosine to uracil does occur naturally, but the half-life of cytosine to uracil in single strand is approximately 200 years. However, this deamination is accelerated by photo-cross-linking reaction between cytosine and photo-cross-linker such as ^{CNV}K and 5- carboxyvinyldeoxyuridine(^{CV}U). Moreover, this photo-cross-linking reaction using ^{CNV}K can be performed in living cells[17]. Thus, the photochemical transition of C to U has the potential to edit the nucleobase in living cells. However, this method utilises high temperature (90 °C) during the deamination step, which is unsuitable for *in vivo* applications. We recently demonstrated that the base complementary to cytosine affected the rate of deamination due to specific hydrogen bonding. When inosine was the complementary base, the rate of cytosine deamination was accelerated under physiological conditions. This finding indicated that the surrounding environment of

target cytosine is an important factor in C to U transformation. In the present study, we focused on photo-cross-linkers and synthesised new photo-responsive ODNs that contain ultrafast cross-linkers, such as 3-methoxycarbonylcarbazole(^{OMeV}K), 3-carboxyvinyl-carbazole nucleoside (^{OHV}K), and 3-carbonylamidevinylcarbazole nucleoside(^{NH2V}K), which, facilitated C to U transformation in DNA, similarly to 3-cyanovinylcarbazole nucleoside (^{CNV}K). In order to accelerate deamination reaction at lower temperatures, we designed ^{CNV}K derivatives, namely, ^{OMeV}K, ^{OHV}K, and ^{NH2V}K (Figure 3.1) with different hydrophilicityies and polarities. We hypothesized that these structural changes would facilitate the attack of water on the target amino group of cross-linked cytosine. ODNs which had the photo-responsive nucleoside at -1 position with respect to target cytosine in the complimentary ODN(C) (5'-ACGGGCGCA-3') and contained modified photo-cross-linkers were synthesized. Two photo-responsive nucleosides ^{CNV}K and ^{OMeV}K were synthesized, and the other two nucleosides were prepared by the postsynthetic modification of the synthesized ODN containing ^{OMeV}K. (Figure 3.2).

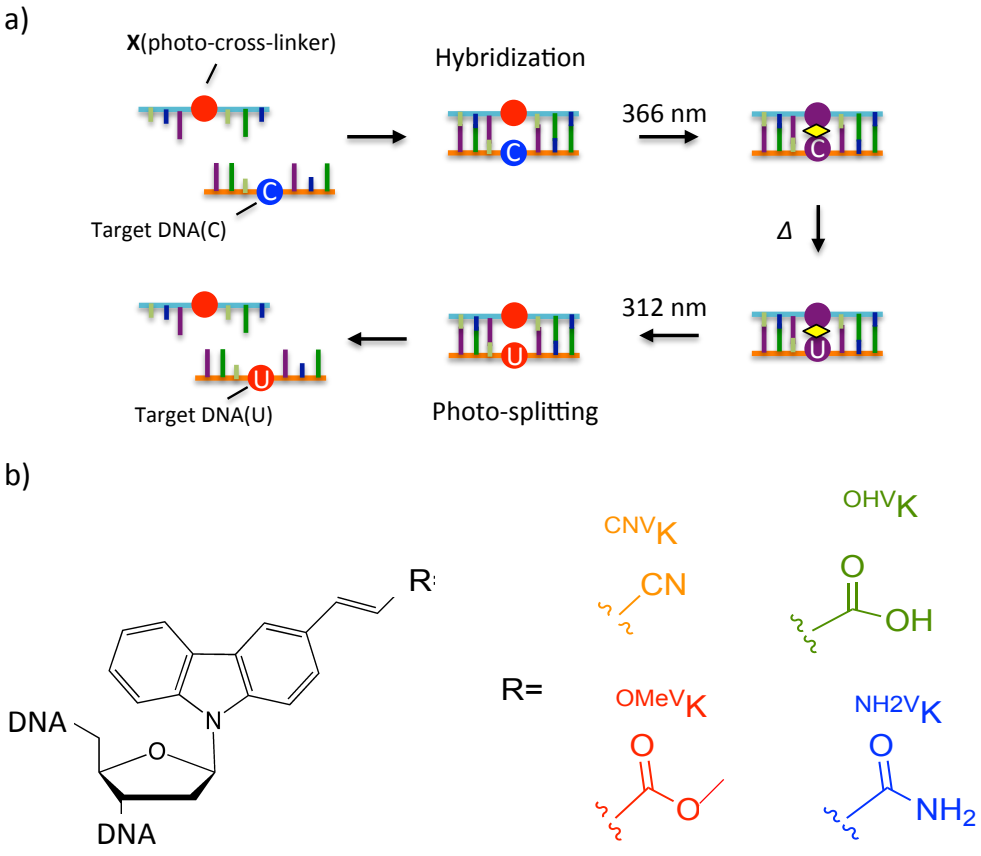


Figure 3.1: a) Photochemical Cytosine to Uracil transition in duplex. X shows photo-cross-linker. b) Structure of 3-vinylcarbazole derivatives.

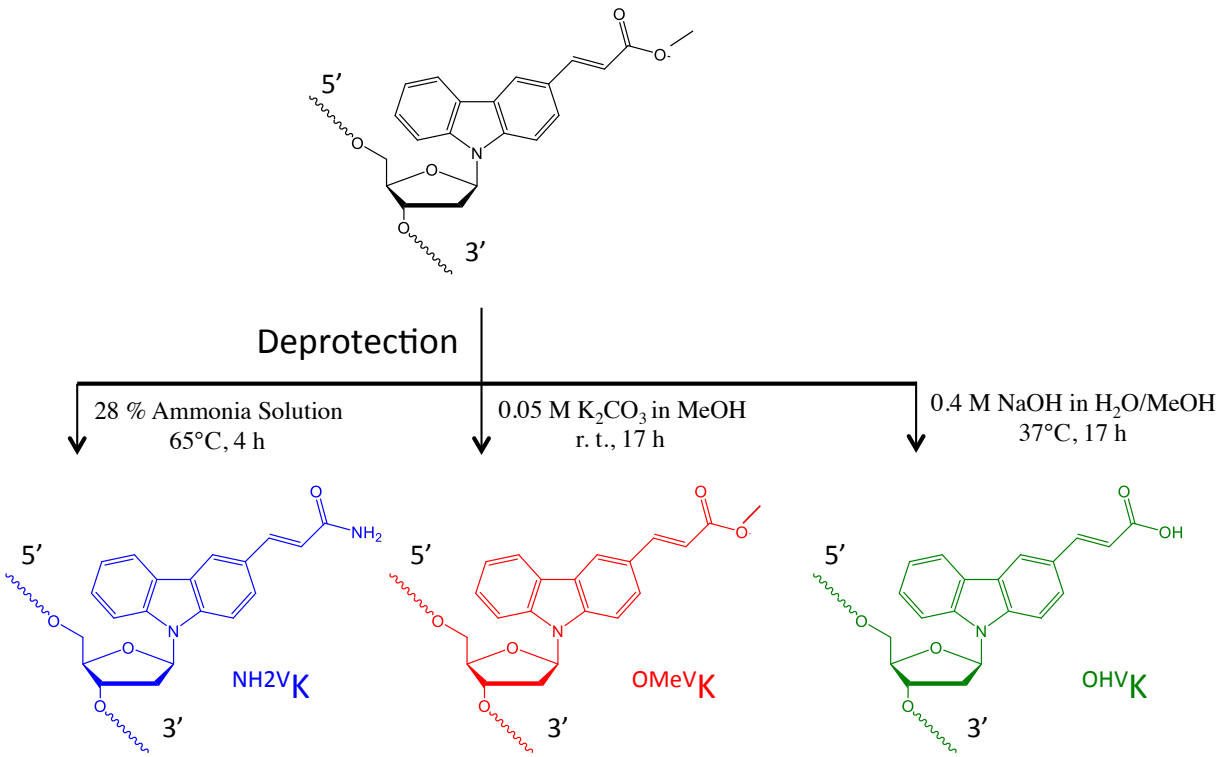


Figure 3.2: Post-modification of photo-cross-linker in ODN.

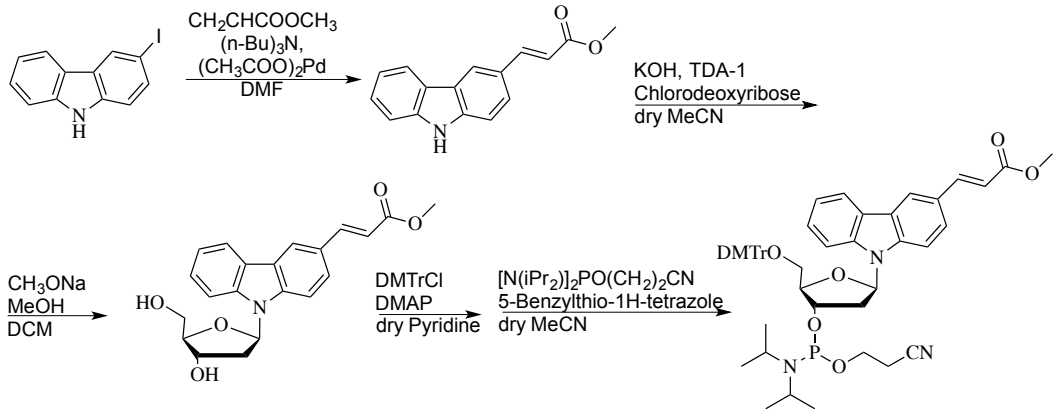
Materials and Methods

General.

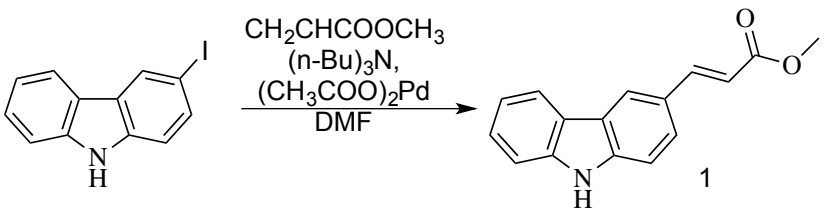
1H NMR spectrum were recorded on a Bruker AVANCE III 400 system. Mass spectra were recorded on a Voyager PRO-SF, Applied Biosystems. HPLC was performed on a Chemcosorb 5-ODS-H column with JASCO PU-980, HG-980-31, DG-980-50 system equipped with a JASCO UV 970 detector at 260 nm. Reagents for the DNA sysnthesizer such as A, G, C, T-β-cyanoethyl phosphoramidite, and CPG support were purchased form Glen research.

Synthesis and preparation of oligonucleotides

The phosphoramidite of ^{OMeV}K was prepared per the scheme (Scheme 3.1).



Scheme 3.1: Synthetic scheme for preparation of photo-responsive nucleotide



Scheme 3.2: Synthesis of 3-methoxycabonylvinyl-carbazole

3-Methoxycarbonylvinyl-carbazole (1): 3-Iodocarbazole (1.50 g, 5.12 mmol) and Palladium acetate (115 mg, 0.512 mmol) were taken in DMF (2.38 mL). Tributylamine (1.22 mL, 5.12 mmol) and Methyl acrylate (1.15 mL, 12.8 mmol) were added to the mixture. The reaction mixture was microwaved at 60W, 160ºC for 20minutes. After completion of reaction, checked by TLC (Hexane:Ethyl acetate = 4:1), Palladium acetate was filtered off. The filtrate was reduced under vacuum and subjected to silica-gel column chromatography (Hexane:Ethyl acetate = 3:1) to afford 3-Methoxycarbonylvinyl- carbazole (1.23 g, 4.89 mmol, 96%).
^{1}HNMR (400MHz, DMSO-d^{6}) 11.54(s, 1H), 8.52(s, 1H), 8.16(d, 1H, J=7.72 Hz), 7.85(d, 1H, J=15.92 Hz), 7.76(dd, 1H, J=1.44, 9.96 Hz), 7.50-7.53(m, 2H), 7.42(t, 1H, J=7.18 Hz), 7.20(t, 1H, J=7.34 Hz), 6.63(d, 1H, J=15.92 Hz), 3.74 (s, 3H)

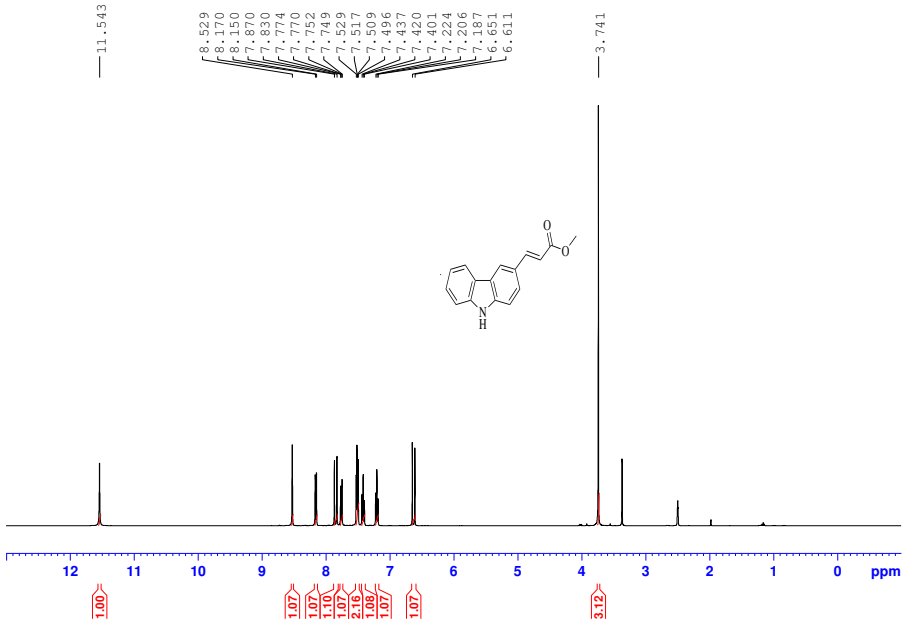
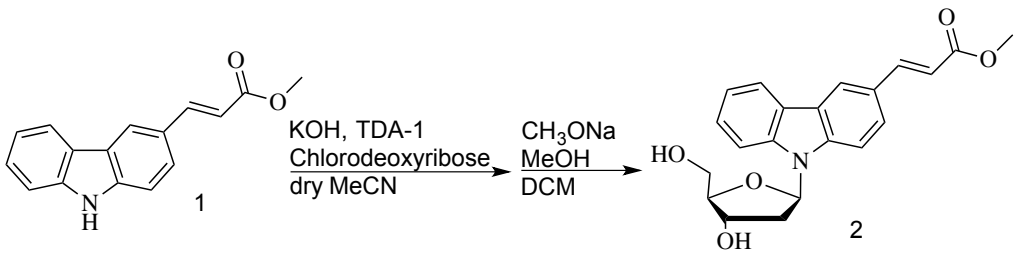


Figure 3.3: ^{1}H NMR analysis of 3-Methoxycarbonylvinyl-carbazole.



Scheme 3.3: Synthesis of 3-Methoxycarbonylvinylcarbazole-1'-β-deoxyriboside(2)

3-Methoxycarbonylvinylcarbazole-1'-β-deoxyriboside (2): 1 (1.27 g, 5.06 mmol)、KOH(0.85 g, 15.18 mmol)、TDA-1 (32 µL, 0.1012 mmol)were dissolved in dry acetonitrile (100 mL) and stirred for 5minutes. Then, Chlorodeoxyribose (2.36 g, 6.07 mmol) was added to the mixture and stirred for 1 hour at room temperature. Reaction was monitored with TLC (Hexane:Ethyl acetate = 4:1), upon completion, the mixture was subjected to vacuum and the resulting product was dried to obtain 3-Methoxycarbonylvinylcarbazole-1'-β-deoxyriboside-3',5'-di-(*p*-toluoyl)ester.

3-Methoxycarbonylvinylcarbazole-1'-β-deoxyriboside-3',5'-di-(*p*-toluoyl)ester was then dissolved in MeOH (184 mL) / CH_2Cl_2 (48 mL) mixture and NaOMe (898 mg, 16.62 mmol) / MeOH (8.31 mL)was added to it and stirred for 3hours at room temperature. TLC ($CHCl_3$:MeOH = 9:1) was checked to ascertain that the starting material has been used up. The reaction mixture was filtered and evaporated. The product was subjected to silica-gel column chromatography ($CHCl_3$:MeOH = 9:1) to get 3-Methoxycarbonylvinylcarbazole-1'-β-deoxyriboside (2) (0.643 g, 1.75 mmol, 35%).

^{1}HNMR (400MHz, $CDCl_3$) 8.16(s, 1H), 8.04(d, 1H, J=7.68 Hz), 7.84(d, 1H, J=15.92 Hz), 7.40-7.61(m, 4H), 7.27(t, 1H, J=7.34 Hz), 6.55(dt, 1H, J=30.68, 7.54 Hz), 6.44(d, 1H, J=15.92 Hz), 4.71(quin, 1H, J=3.93 Hz), 3.92-4.00(m, 2H,), 3.82 (s, 3H), 2.88(dt, 1H, J=14.04, 8.16 Hz), 2.59-2.77(m, 1H), 2.23(dq, 1H, J=3.72, 3.44)

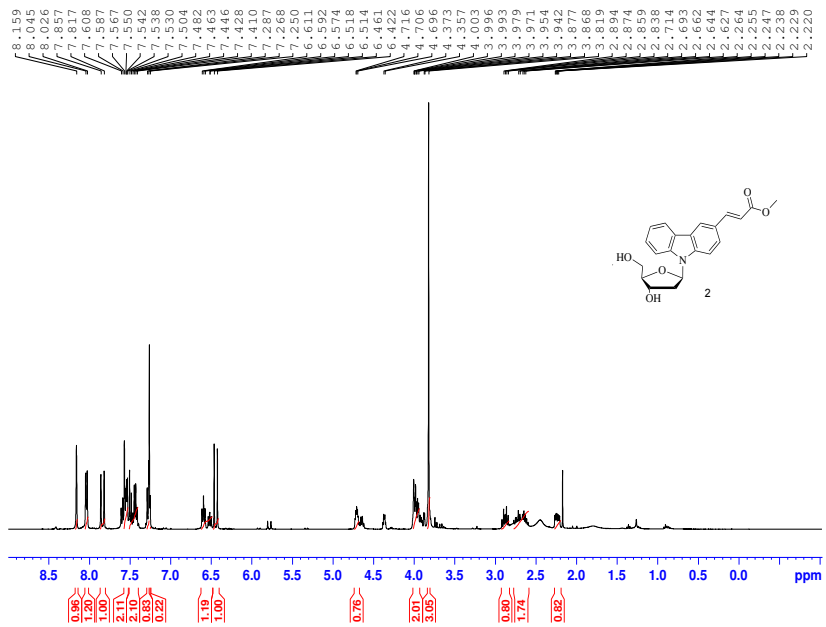
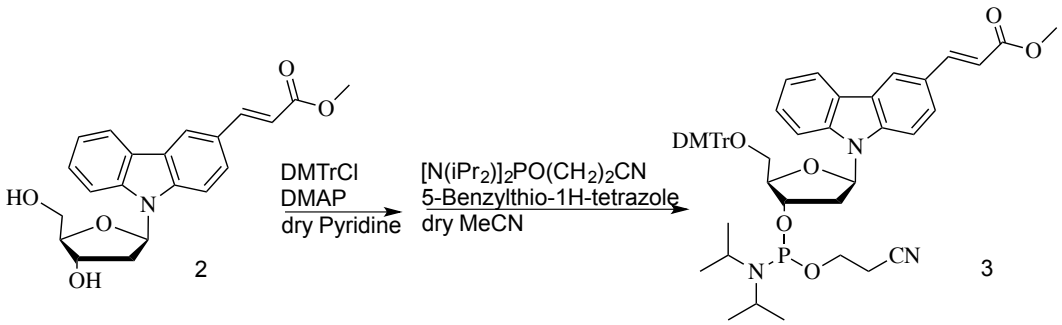


Figure 3.4: ^{1}H NMR analysis of 3-Methoxycarbonylvinylcarbazole-1'-β-deoxyriboside.



Scheme 3.4: Synthesis of 5'-O-(4,4'-dimethoxytrityl)-3-Methoxycarbonylvinylcarbazole-1'-β-deoxyriboside-3'-O-(cyanorthoxy-N,N-diisopropylamino)phosphoramidite (3)

5'-O-(4,4'-dimethoxytrityl)-3-Methoxycarbonylvinylcarbazole-1'-β-deoxyriboside-3'-O-(cyanorthoxy-N,N-diisopropylamino) phosphoramidite (3): 2 (640 mg, 1.74 mmol) was dissolved in pyridine(5.0 mL) followed by addition of DMTrCl(0.71 g, 2.09 mmol)/DMAP(43 mg, 0.35 mmol)/pyridine(8.5 mL) at 0°C under nitrogen atmosphere and stirred for 22 hours at room temperature. Completion of reaction was confirmed by TLC ($CHCl_3$:MeOH = 99:1). The product was dissolved in chloroform and washed with water, $NaHCO_3$ (aq.), and brine. Dried using sodium sulphate and the solvent was removed under vacuum and the resulting product was subjected to silica-gel column chromatography (0.1% TEA $CHCl_3$:MeOH = 99:1)

to obtain 5'-*O*-(4,4'-dimethoxytrityl)-3-Methoxycarbonylvinyl carbazole-1'-β-deoxyriboside (804 mg, 1.20 mmol, 69%).

^{1}HNMR (400MHz, $CDCl_3$) 8.19(s, 1H), 8.05-8.07(m, 1H), 7.85(d, 1H, J=15.92 Hz), 7.64-7.69(m, 2H), 7.51(s, 1H), 7.50(s, 1H), 7.39(d, 2H, J=3.60 Hz), 7.41(s, 1H), 7.40(s, 1H), 7.38(s, 1H), 7.38(s, 1H), 7.23-7.33(m, 6H), 6.82-6.84(m, 4H), 6.62-6.65(m, 1H,), 6.42 (d, 1H J=15.92 Hz), 4.81(quin, 1H, J=3.55), 4.10(q, 1H, J=3.89), 3.83(s, 3H), 3.78(s, 3H), 3.78(s, 3H), 3.59(d, 1H, J=3.80), 2.91(dt, 1H, J=14.00, 8.18 Hz), 2.24(ddd, 1H, J=4.84, 3.60, 2.68 Hz)

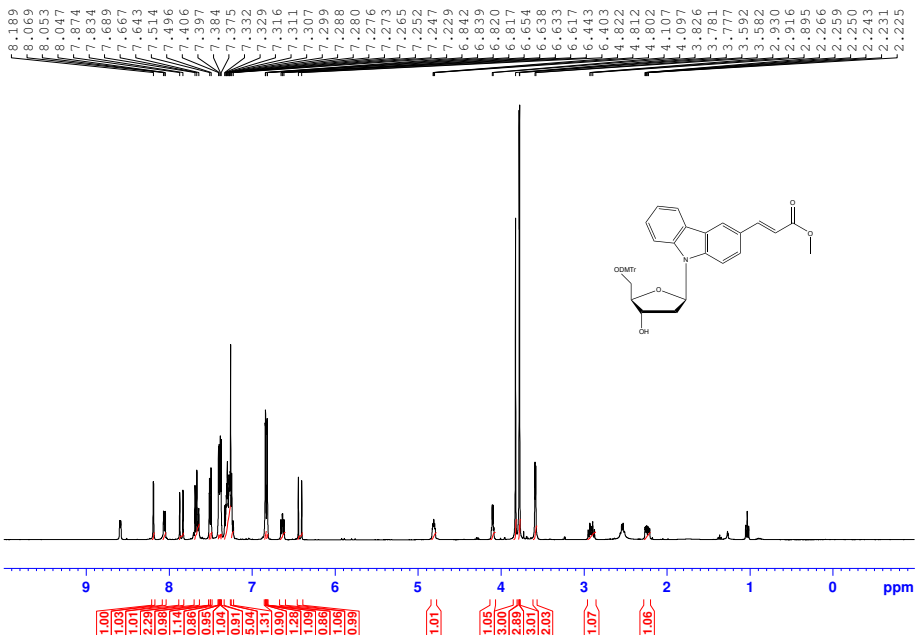
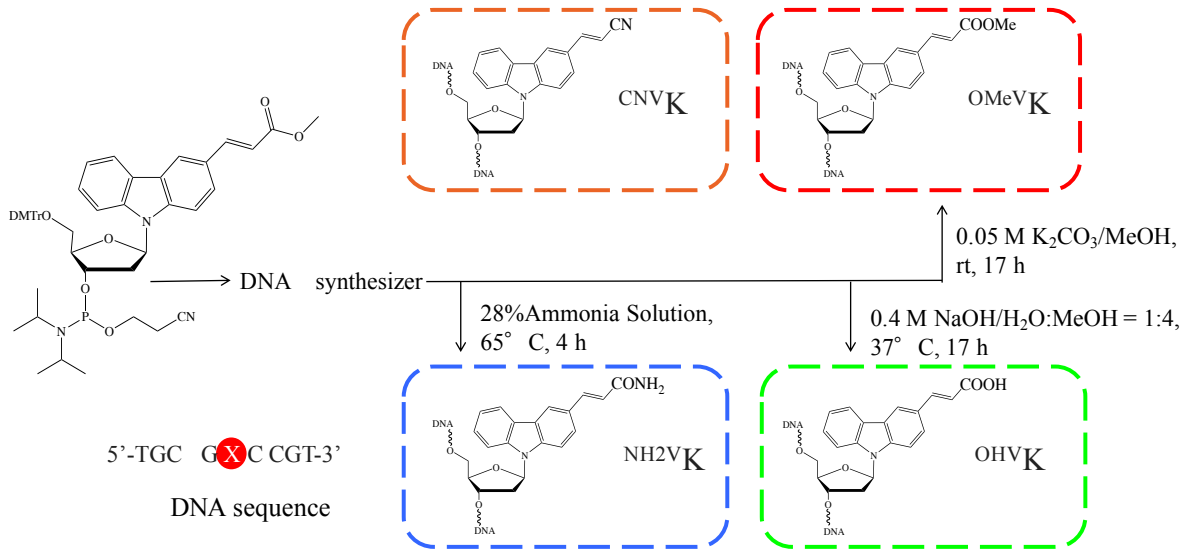


Figure 3.5: ^{1}H NMR analysis of 5'-O-(4,4'-dimethoxytrityl)-3-methoxycarbonylvinylcarbazole-1'-ß-deoxyriboside



Scheme 3.5: Post-modification of photo-cross-linker in ODN

Preparation of modified oligonucleotides.

The phosphoramidite of ^{CNV}K was prepared following to previous reports.[16] The modified oligonucleotides containing ^{CNV}K or ^{OMeV}K were prepared, according to standard phosphoramidite chemistry using DNA synthesizer (ABI 3400 DNA synthesizer, Applied Biosystems, CA). The ODN containing ^{OMeV}K was postmodified to ODN containing ^{OMeV}K, ^{OHV}K, or ^{NH2V}K in deprotection step (Scheme 3.5). Synthesized ^{CNV}K ODNs were detached from the support by soaking in concentrated aqueous ammonia for 1 h at room temperature. Deprotection was conducted by heating the concentration aqueous for 4 h at 65°C concentrated aqueous ammonia was then removing it by speedvac, and the crude oligomer was purified by reverse phase HPLC equipped with InertSustainTM C18 column Cosmosil^{TM} 5C_{18}-AR-II column (5 μm, 10 × 150 mm, Nacalai tesque, Flow rate of 3.0 mL/min, 60°C) and lyophilized. Synthesis of ODN was confirmed by MALDI-TOF-MS (Table 3.1). Other ODNs were purchased from Fasmac (Japan).

Isolation of photo-cross-linked dsDNA.

The cODN (C)-Cy3 (10 μM) and X-ODN (GK), where X is photo-cross-linker, (10 μM) in buffer solution (100 mM NaCl, 50 mM sodium cacodylate, pH 7.6) was photoirradiated at 366 nm for 300 sec using UV-LED illuminator(OMRON Inc, 1600 mW) on ice. The solution was purified by reverse phase HPLC, the concentration of photo-cross-linked dsDNA was

determined by absorbance of Cy3 molecule at 550 nm, and synthesis was confirmed by MALDI-TOF-MS analysis (Table 3.1).

Table 3.1: MALDI-TOF-MS analysis of photo-adducts.

Entry	Calcd. $[M+H]^+$	Found
^{CNV}**K- ODN(GK)/ODN(C)**	5544.26	5544.03
^{OMeV}**K- ODN(GK)/ODN(C)**	5564.96	5563.03
^{OHV}**K- ODN(GK)/ODN(C)**	5563.14	5562.05
^{NH2V}**K- ODN(GK)/ODN(C)**	5578.39	5577.04

Deamination

The solution of 5 μM ODN(GK) and ODN(C) in buffer solution was annealed and photoirradiation at 366 nm. The photoirradated solution was purified by a reversible HPLC to get the purified photo-cross-linked dsDNA. The 5 μM photo-cross-linked dsODN in a buffer solution (50 mM Na-cacodylate buffer (pH 7.4) containing 100 mM NaCl) was incubated at 37°C, 50°C, 70°C and 90°C

Polyacrylamide gel analysis

Polyacrylamide gel electrophoresis (PAGE) was performed with 15% polyacrylamide containing 8M urea. After the electrophotresis(150V, 80 min), fluorescent image was taken by luminescent image analyzer (LAS3000, Fujifilm, Japan).

Molecular modelling

Molecular modelling analysis was performed using Macromodel v8.1 with AMBER* force field, water as solvent, and H-bond parameter 2.5 Å constrain. Energy minimization (500 cycles) and stochastic molecular dynamics were performed to find the most stable structure.[18]

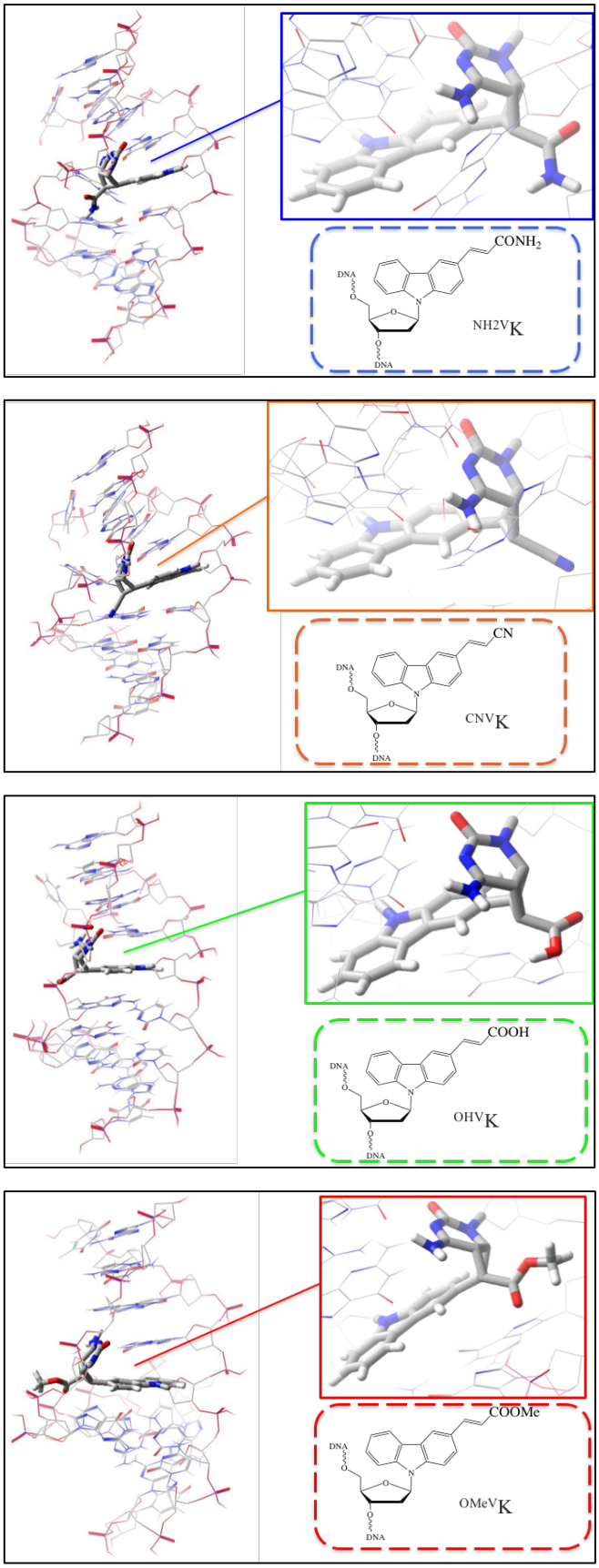
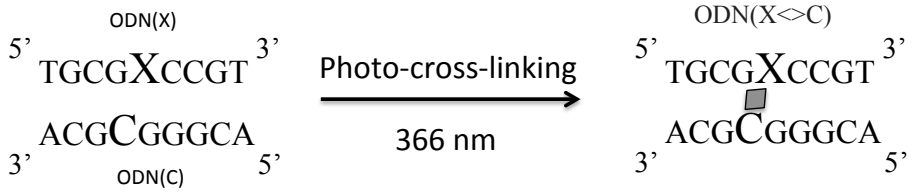


Figure 3.6: Molecular modeling images of the photo-cross-linked ODNs showing varying spatial orientation of the photo-adducts with different photo-cross-linkers.

Results and Discussion

At first, photo-responsive ODN(X) (5'-TGCGXCCGT-3') was subjected to photo-cross-linking reaction with ODN(C) by irradiating the mixture of both ODNs in 50 mM sodium cacodylate buffer (pH 7.4) containing 100 mM sodium chloride, using 366 nm UV radiation at 4 °C for 120 s (Fig 3.7), and analysed by reversed phase high performance liquid chromatography (HPLC) and ultrahigh performance liquid chromatography (UPLC). HPLC analysis mixture before and after irradiation, resulted in the appearance of a new peak and disappearance of two peaks that of the ODN mixture before and after irradiation using HPLC, a new peak was observed on HPLC chromatograms, whereas two peaks that corresponded to single-stranded ODN disappeared over the course of photoirradiation (Fig 3.7-3.10), confirming the formation of photo-adduct ODN(X<>C). The analysis of photo-cross-linking reaction at different time intervals (Figures 3.7-3.10) showed (Figure 3.11) that the maximum reaction rate was observed when ^{CNV}K (k = 0.33 × s^{-1}) was used as photo-cross-linker. Although, the new photo-cross-linkers had slightly lower reaction rate constants (^{NH2V}K: k = 0.23 × s^{-1}; ^{OMeV}K: k = 0.11 × s^{-1}; and ^{OHV}K: k = 0.08 × s^{-1}), 95% conversion was observed within 60 s for all ODNs. The photo-adducts were stable under UV irradiation even after overexposure, and no degradation was observed. The photo-adducts were purified using HPLC and analysed by MALDI-TOF-MS (Table 3.1).



Scheme 3.6: Scheme of photo-cross-linking reaction.

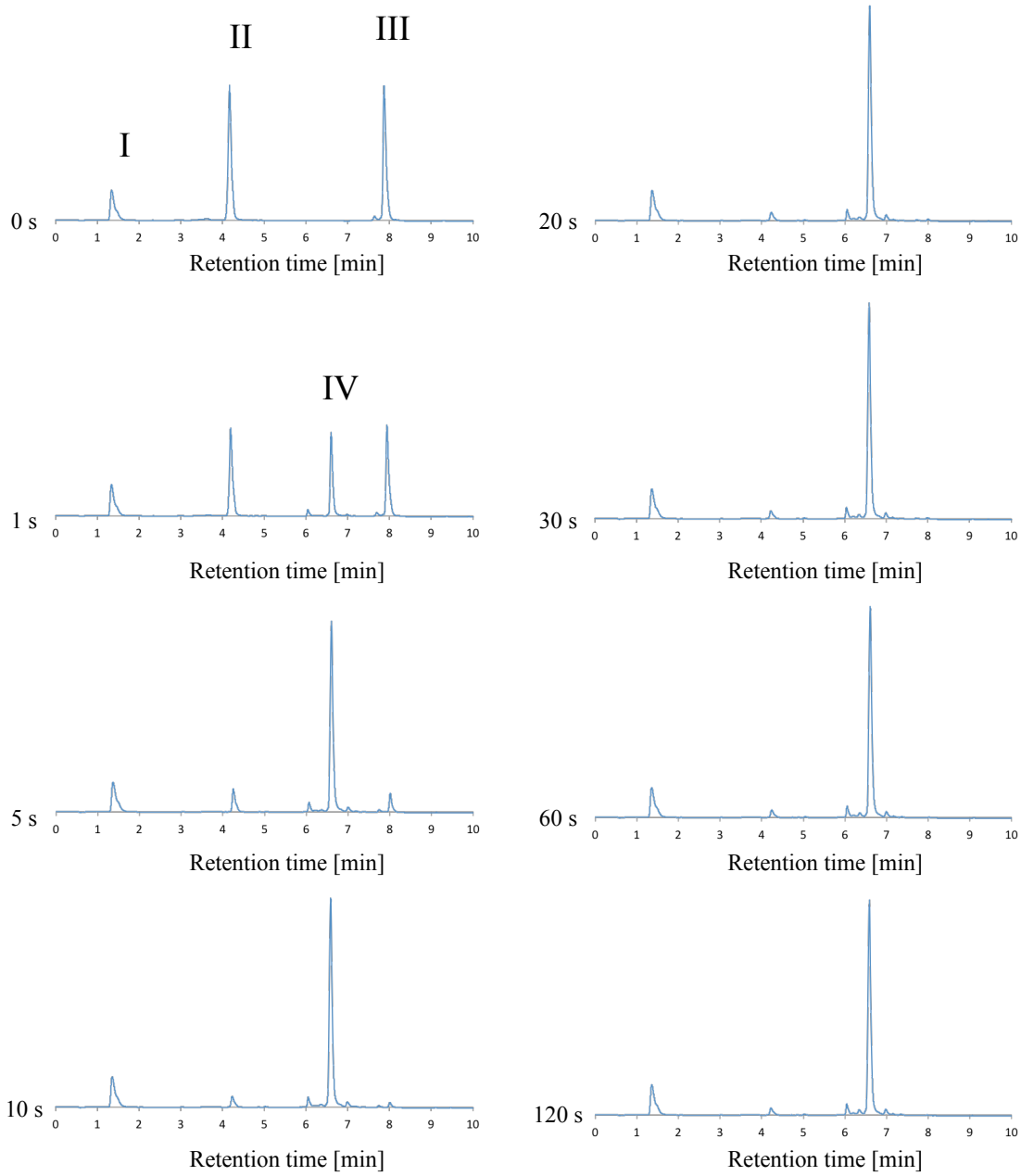


Figure 3.7: UPLC analysis of photo-cross-linking by ODN containing ^{CNV}K, I; Deoxyuridine, II; ODN(C), III; ^{CNV}K-ODN(GK), IV; Photoadduct ^{CNV}K-ODN(GK)/ODN(C). Elution: 50 mM ammmoniumformate/MeCN (99:1 to 90:10), Linear gradient (10 min) at a flow rate of 0.2 mL/min, The temperature of the column was maintained at 60 °C.

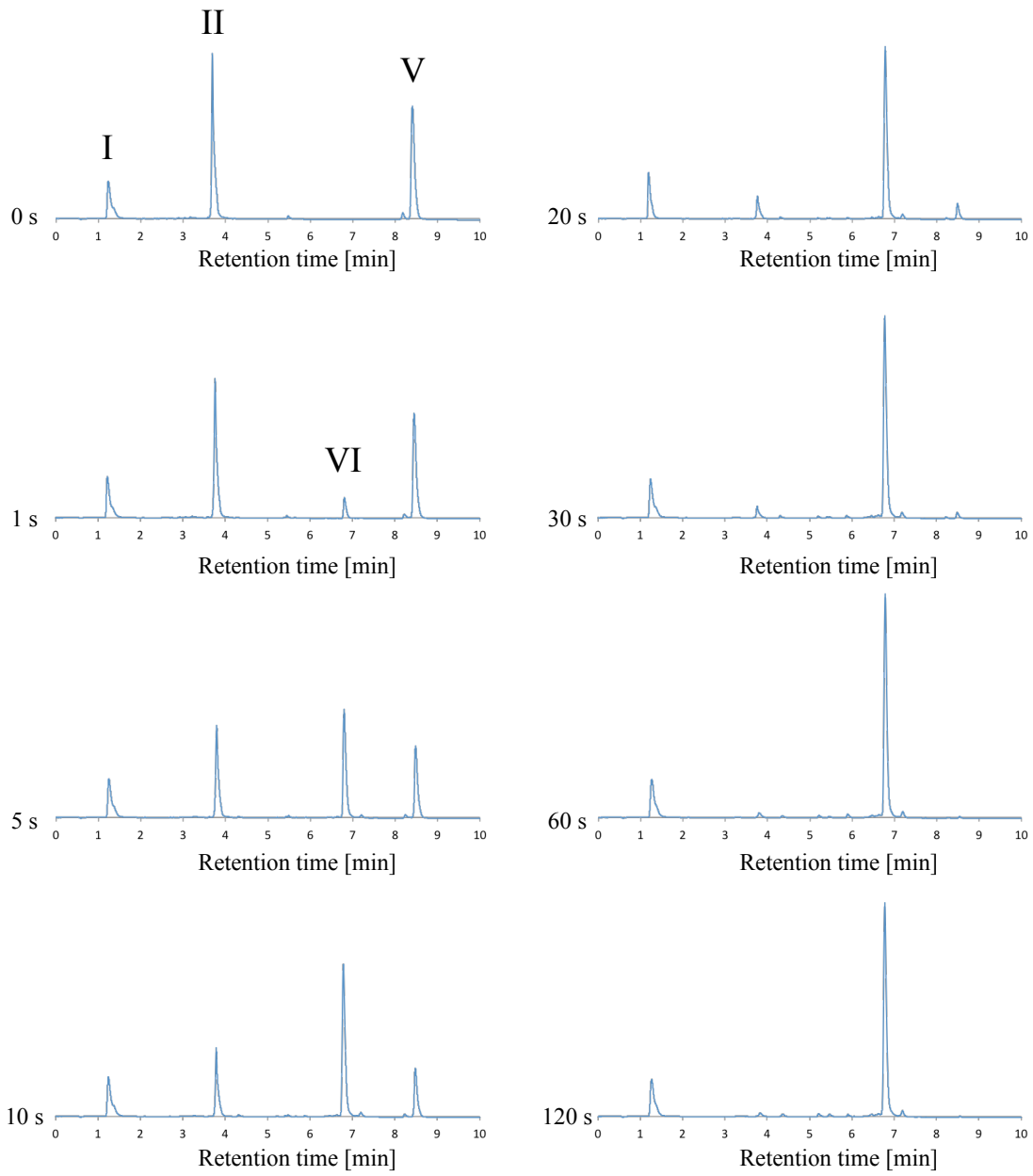


Figure 3.8: UPLC analysis of photo-cross-linking by ODN containing ^{OMeV}K, I; Deoxyuridine, II; ODN(C), V; ^{OMeV}K-ODN(GK), VI; Photoadduct ^{OMeV}K-ODN(GK)/ODN(C). Elution; 50 mM ammmoniumformate/MeCN (99:1 to 90:10), Linear gradient (10 min) at a flow rate of 0.2 mL/min. The temperature of the column was maintained at 60 °C.

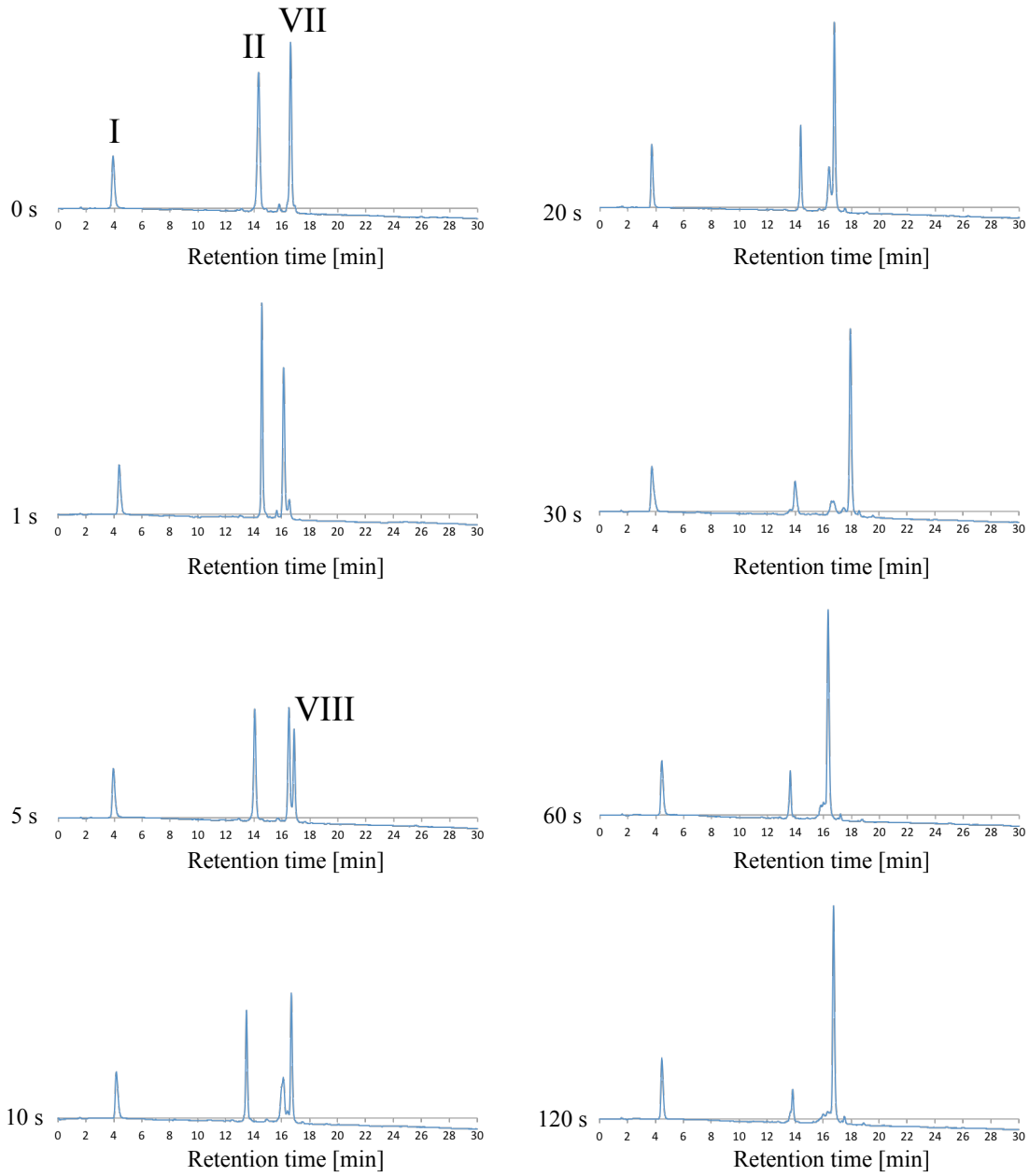


Figure 3.9: UPLC analysis of photo-cross-linking by ODN containing ^{NH2V}K, I; Deoxyuridine, II; ODN(C), IX; ^{NH2V}K-ODN(GK), X; Photoadduct ^{NH2V}K-ODN(GK) and ODN(C). Elution; 50 mM ammmoniumformate/MeCN (99:1 to 90:10), Linear gradient (10 min) at a flow rete of 0.2 mL/min, The temperature of the column was maintained at 60 °C.

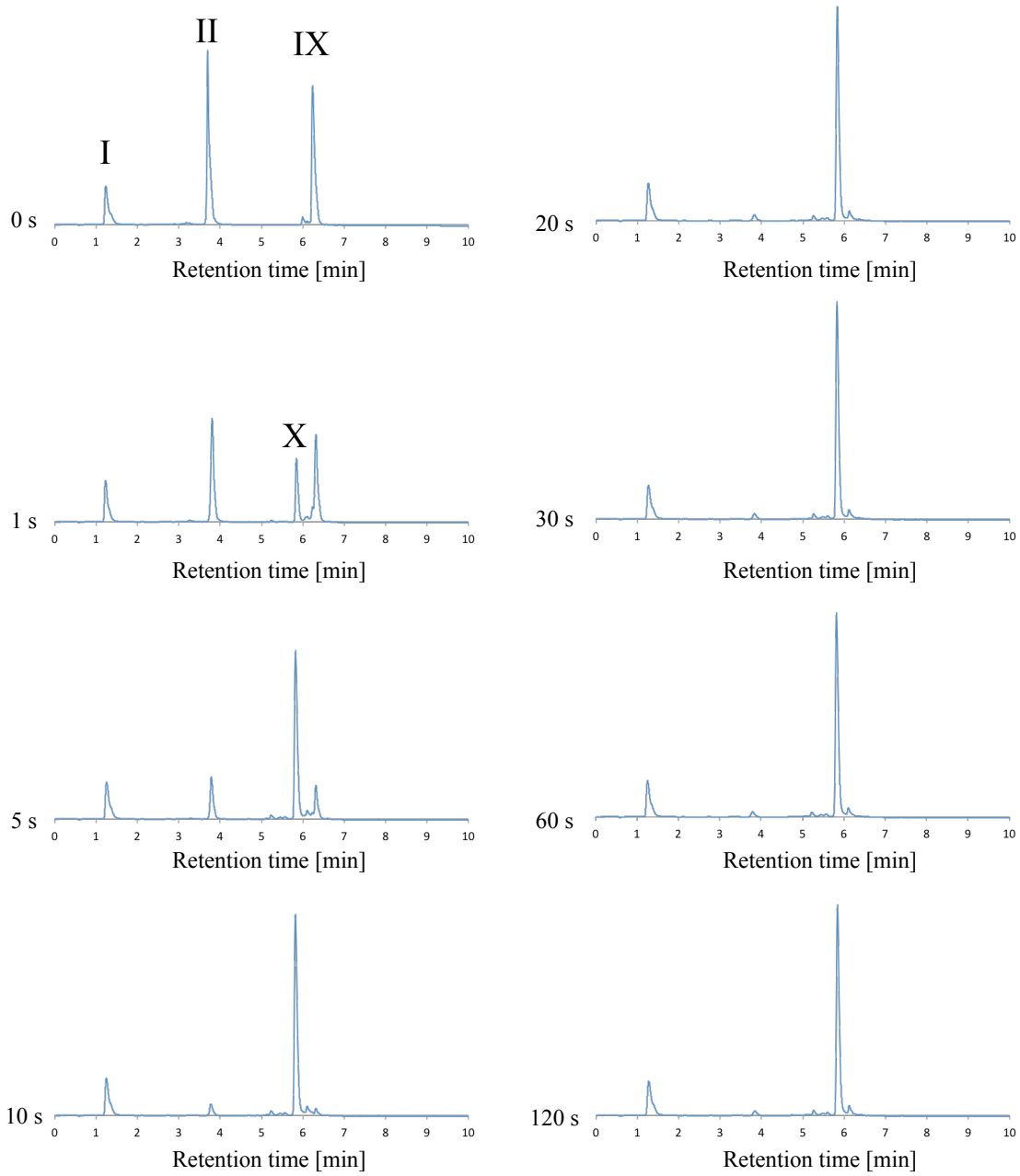


Figure 3.10: UPLC analysis of photo-cross-linking by ODN containing ^{NH2V}K, I; Deoxyuridine, II; ODN(C), IX; ^{NH2V}K-ODN(GK), X; Photoadduct ^{NH2V}K-ODN(GK) and ODN(C). Elution; 50 mM ammmoniumformate/MeCN (99:1 to 90:10), Linear gradient (10 min) at a flow rate of 0.2 mL/min, The temperature of the column was maintained at 60 °C.

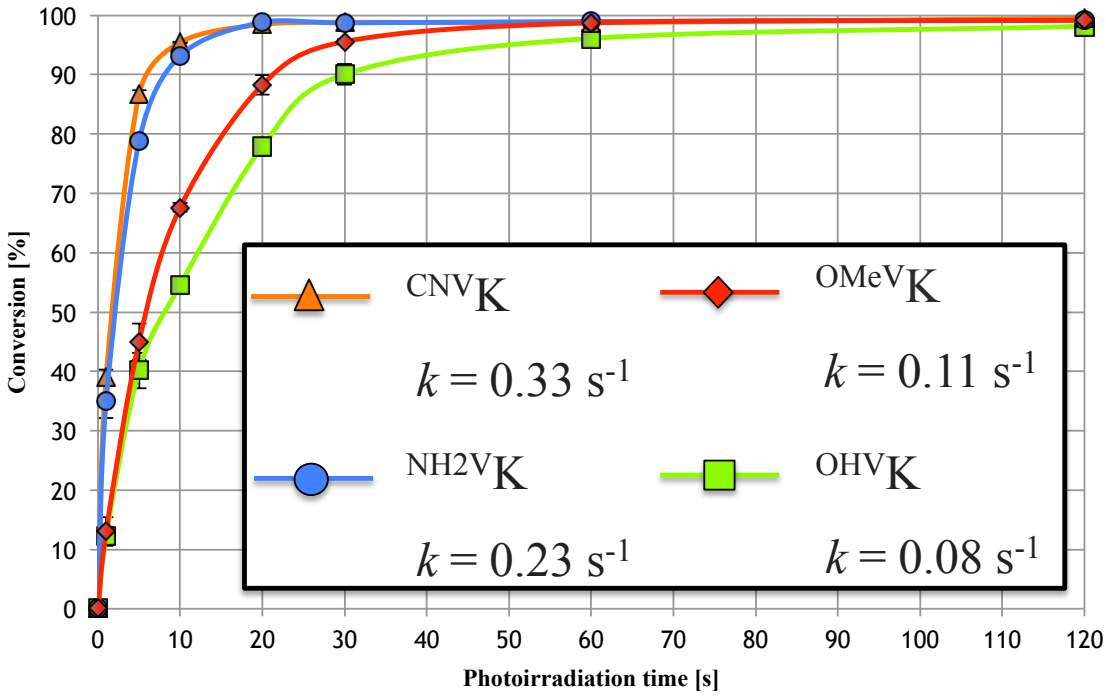


Figure 3.11: Time course of the photo-cross-linking reaction between ODN(GK) and ODN (C) and reaction rate constants with various photo-cross-linker.

To evaluate the effect of photo-responsive base substitution in C to U transformation, photo-cross-linked ODNs were incubated at 70 °C (Figure 3.12a). Upon incubation, the ODNs were analysed after incubation, by denaturing polyacrylamide gel electrophoresis after a certain interval of time until 12 h and the gels were visualised using LAS-3000 (Fujifilm). Cy3 labelled ODN(C) and ODN(U) were used in this experiment to image Cy3 fluorescence. Analysis of gels showed a clear distinction between the photo-adduct that contained C at the target site and mutated ODN that contained U at the target site (Figure 3.12b). The distinction was made by comparing ODN(X<>C) and ODN(X<>U), where in ODN(X<>U), a complementary ODN(U) was synthesized with U at the target site. We observed (Figure 3.12c) that the deamination reaction proceeded at the maximum rate ($k = 6.9 \times 10^{-5} \times s^{-1}$) in the case of ODN(X<>C) with X = ^{OHV}K. Deamination reaction rates were intermediate when ODNs containing ^{CNV}K and ^{NH2V}K were used ($4.37 \times 10^{-5} \times s^{-1}$ and $4.05 \times 10^{-5} \times s^{-1}$, respectively. The slowest rate of deamination was observed in the case of ODN that contained ^{OMeV}K ($k = 1.73 \times 10^{-5}$ s^{-1}). Thus, with ODN(^{OHV}K), the rate of deamination of cytosine to uracil was 1.5-fold faster than that with ^{CNV}K ODN.

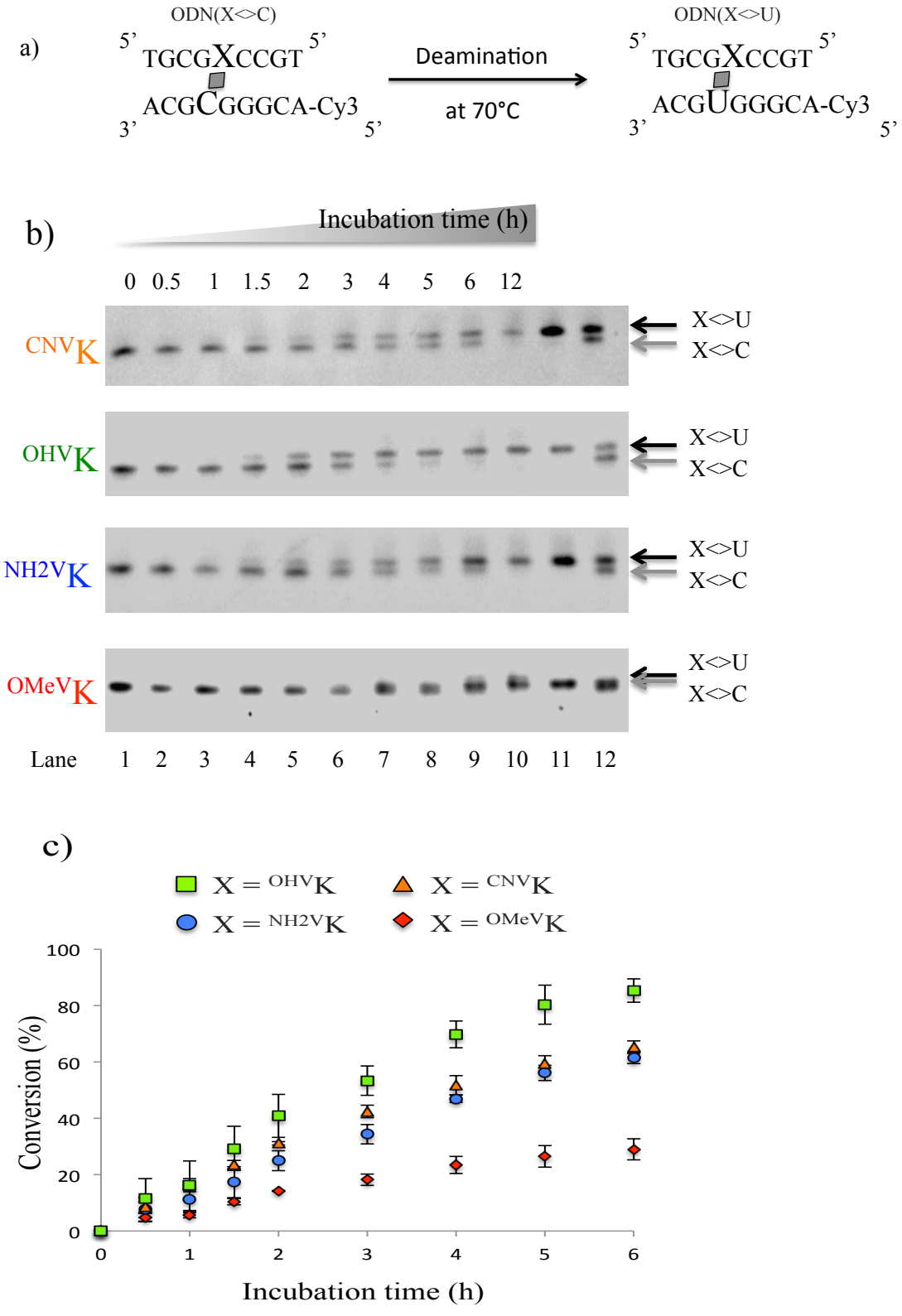


Figure 3.12: a) Scheme of deamination reaction. b) Denaturing PAGE analysis of the products after photochemical DNA editing. Lane 1: ODN(X) + Cy3-ODN(C) + UV (366nm, 5min, 0ºC); Lane 2: Lane 1 + heat (70ºC, 0.5 h); Lane 3: Lane 1 + heat (70ºC, 1 h); Lane 4: Lane 1 + heat (70ºC, 1.5 h); Lane 5: Lane 1 + heat (70ºC, 2 h); Lane 6: Lane 1 + heat (70ºC, 3 h); Lane 7: Lane 1 + heat (70ºC, 4 h); Lane 8: Lane 1 + heat (70ºC, 5 h); Lane 9: Lane 1 + heat (70ºC, 6 h); Lane 10: Lane 1 + heat (70ºC, 12 h); Lane 11: ODN(X) + Cy3-ODN(U) + UV (366nm, 5 min, 0ºC); Lane 12: Lane 1 + Lane 11. c) Time-course of deamination reaction.

To examine the relationship between temperature and the efficiency of photo-responsive nucleoside substitution in C to U transition, we carried out deamination reaction at different temperatures (90, 50, and 37 °C). We found that the acceleration effect of photoresponsive nucleobase substitution was unaltered at 50 °C and 90 °C (Figures 3.13–3.15). The highest rates of deamination reaction were observed with ODN(^{OHV}K). In case of 90 ºC deamination reaction, a new band above the band ODN(^{OMeV}K<>U) has been observed which could be due to saponification of the methyl ester. To explore whether this reaction could proceed in living cells, we also carried out experiments at 37 °C and found that deamination of cross-linked cytosine could be observed at that temperature as well, with the same relationship between reaction rates and ODN structure as those observed during experiments at higher temperatures.

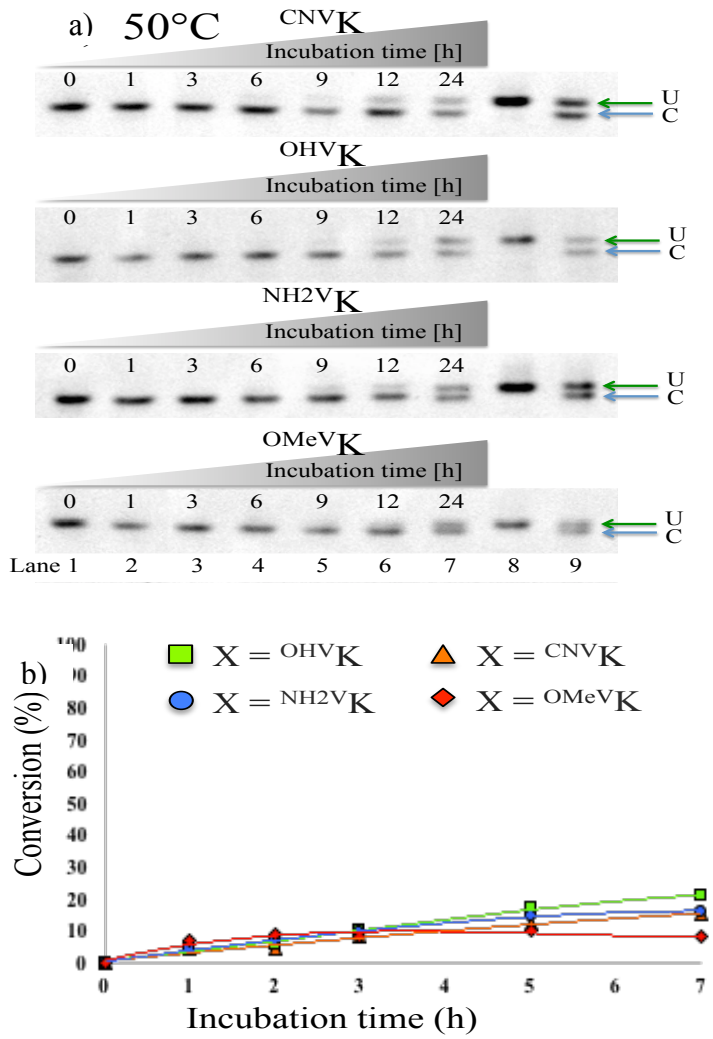


Figure 3.13: a) Denaturing PAGE analysis of the products after each reaction step of the photochemical ODN editing. Lane 1; ODN(GK) + ODN(C)-Cy3 + UV (366 nm, 2 min, 4°C), Lane 2; Lane 1 + heat (50˚C, 1.0 h), Lane 3; Lane 1 + heat (50˚C, 3.0 h), Lane 4; Lane 1 + heat (50˚C, 6.0 h), Lane 5; Lane 1 + heat (50˚C, 9.0 h), Lane 6; Lane 1 + heat (50˚C, 12.0 h), Lane 7; Lane 1 + heat (50˚C, 24.0 h), Lane 8; ODN(GK) + ODN(U)-Cy3 + UV (366 nm, 4°C), Lane 9; Lane 1 + Lane 8. Sample: ODN(GK) (5 µM), ODN(C)-Cy3 (6 µM) in 50 mM Sodium

cacodylate buffer containing 100 mM NaCl Buffer. (pH 7.4). **b)** Time course of photochemical C→U transition using 3-vinylcarbazole derivatives at 50ºC.

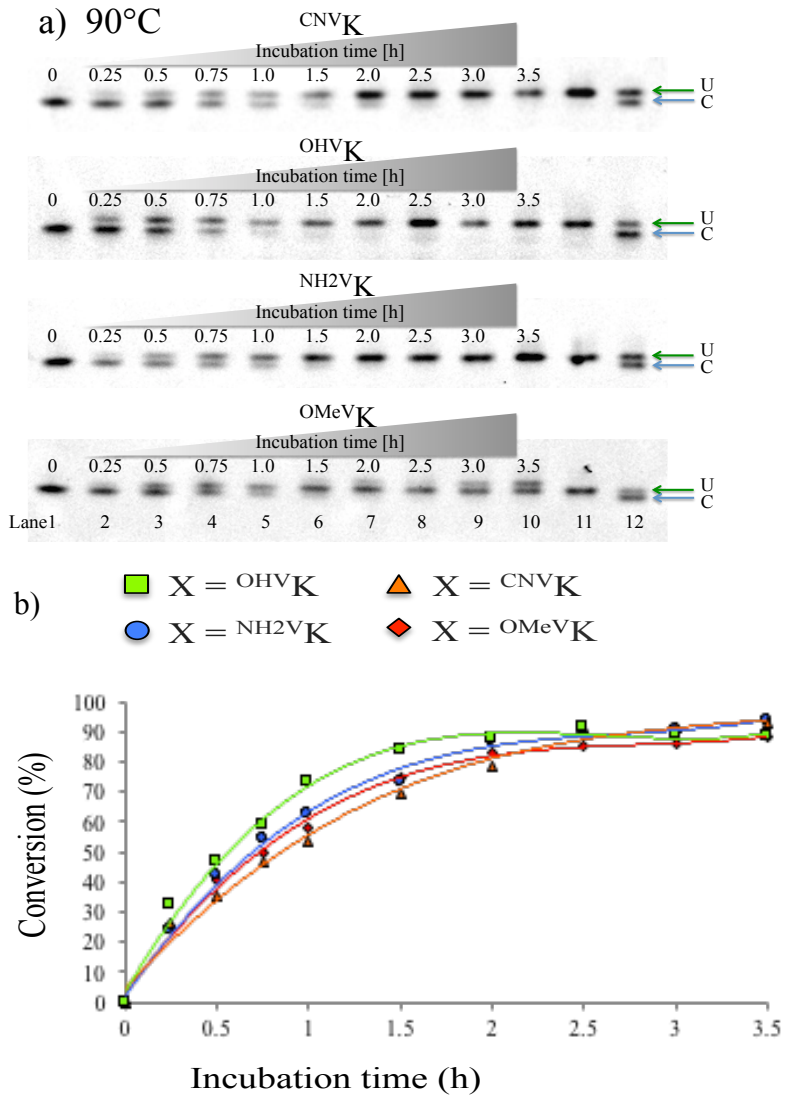


Figure 3.14: a) Denaturing PAGE analysis of the products after each reaction step of the photochemical ODN editing. Lane 1; ODN(GK) + ODN(C)-Cy3 + UV (366 nm, 2 min, 4°C), Lane 2; Lane 1 + heat (90˚C, 0.25 h), Lane 3; Lane 1 + heat (90˚C, 0.5 h), Lane 4; Lane 1 + heat (90˚C, 0.75 h), Lane 5; Lane 1 + heat (90˚C, 1.0 h), Lane 6; Lane 1 + heat (90˚C, 1.5 h), Lane 7; Lane 1 + heat (90˚C, 2.0 h), Lane 8; Lane 1 + heat (90˚C, 2.5 h), Lane 9; Lane 1 + heat (90˚C, 3.0 h), Lane 10; Lane 1 + heat (90˚C, 3.5 h), Lane 11 ODN(GK) + ODN(U)-Cy3 + UV (366 nm, 4°C), Lane 12; Lane 1 + Lane 10. Sample: ODN(GK) (5 µM) , ODN(C)-Cy3 (6 µM) in 50 mM Sodium cacodylate buffer containing 100 mM NaCl Buffer. (pH 7.4). b) Time course of photochemical C→U transition using 3-vinylcarbazole derivatives at 90ºC.

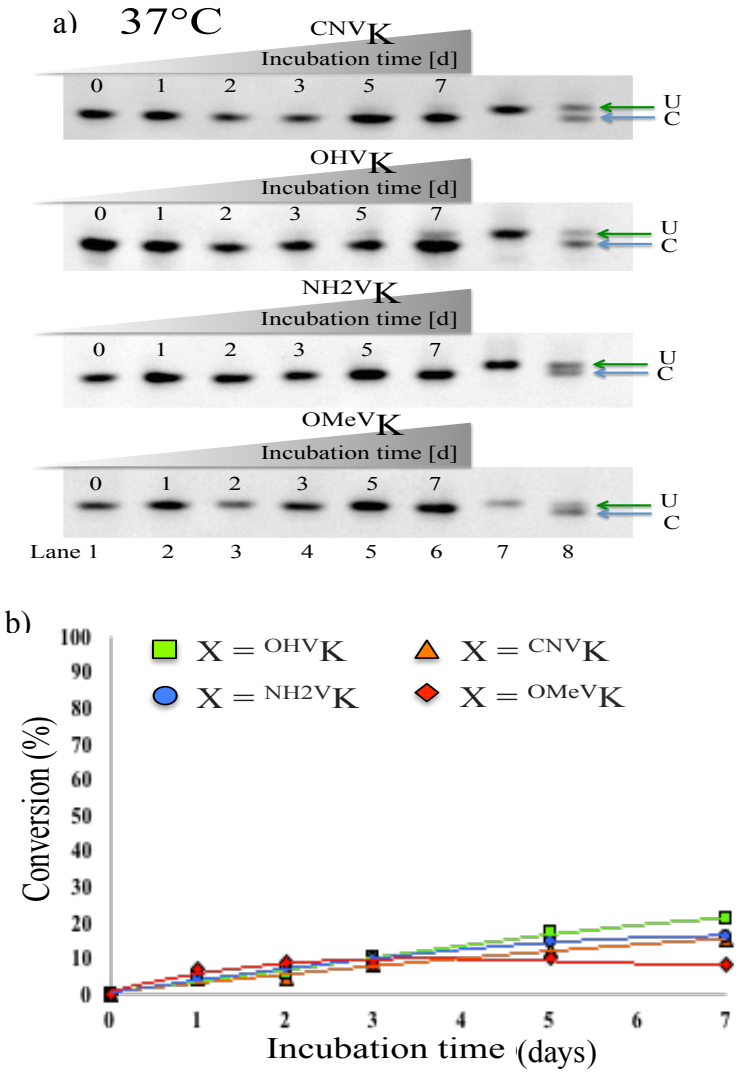


Figure 3.15: a) Denaturing PAGE analysis of the products after each reaction step of the photochemical ODN editing. Lane 1; ODN(GK) + ODN(C)-Cy3 + UV (366 nm, 2 min, 4°C), Lane 2; Lane 1 + heat (37˚C, 1 day), Lane 3; Lane 1 + heat (37˚C, 2 days), Lane 4; Lane 1 + heat (37˚C, 3 days), Lane 5; Lane 1 + heat (37˚C, 5 days), Lane 6; Lane 1 + heat (37˚C, 7 days), Lane 7; ODN(GK) + ODN(U)-Cy3 + UV (366 nm, 4°C), Lane 8; Lane 1 + Lane 8. Sample: ODN(GK) (5 µM) , ODN(C)-Cy3 (6 µM) in 50 mM Sodium cacodylate buffer containing 100 mM NaCl Buffer. (pH 7.4). b) Time course of photochemical C→U transition using 3-vinylcarbazole derivatives at 50ºC

To confirm C to U transition, ODNs photo-split by 312 nm UV irradiation at 60 °C after deamination reaction were subjected to enzymatic digestion using nuclease P1 and snake venom. The digested samples were analysed by UPLC. UPLC chromatograms showed that after deamination, a new peak corresponding to uracil nucleoside appeared in ODNs (Figure 3.16). This confirmed that C to U transformation indeed occurred after photo-cross-linking.

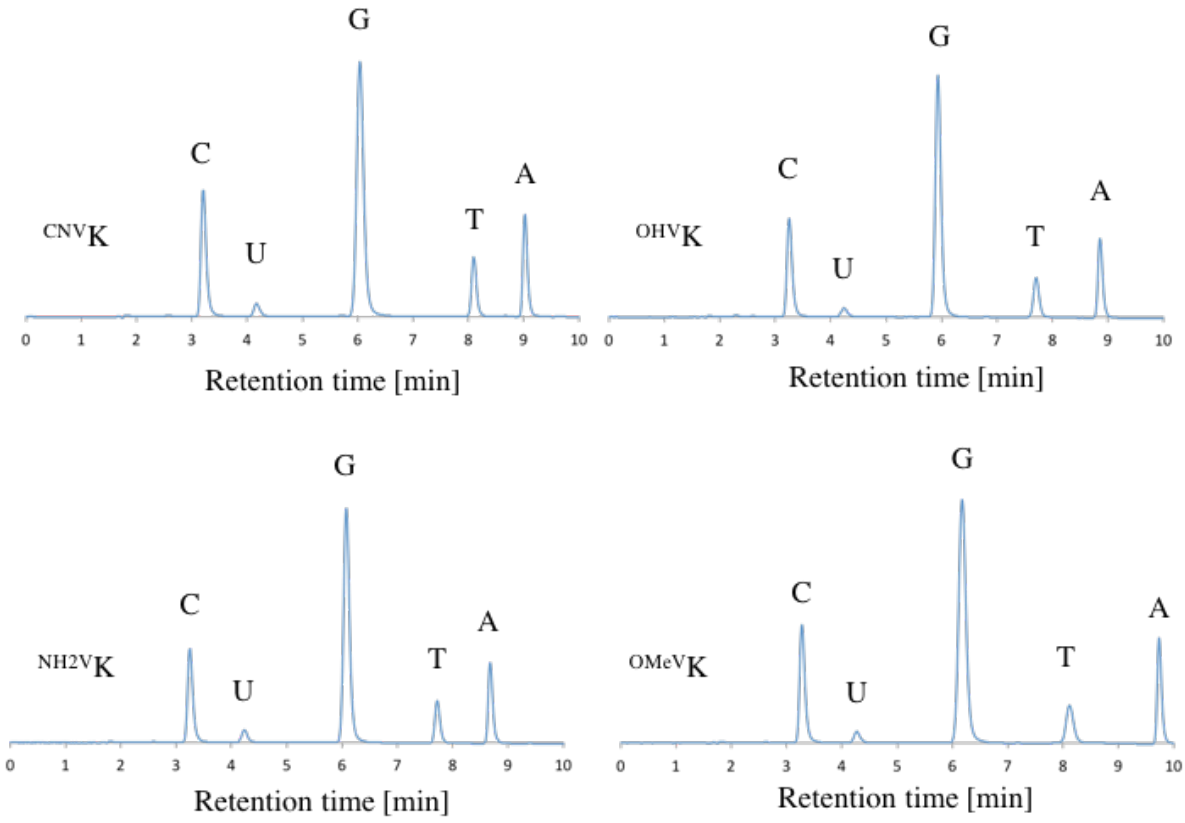


Figure 3.16: Enzymatic digestion of photo-splitted ODNs after deamination reaction. The peak ratio indicated the amount of nucleoside in the ODN. Peak of uracil is observed in the UPLC.

To assess the correlation between the hydrophilicity of ODN and rate of deamination reaction, we calculated partition coefficient (LogP)[19-20] of photo-responsive ODN(X) using their HPLC retention times (Figure 3.17). We found that hydrophilicity of ODNs positively correlated with their ability to convert cytosine to uracil in a photo-responsive manner (Figure 3.18, Table 3.2). ODN(^{OHV}K) had the highest hydrophilicity value (LogP = 0.53), followed by ODN(^{CNV}X) and ODN(^{NH2V}K) (LogP = 0.55 and 0.61, respectively), whereas ODN(^{OMeV}K) was most hydrophobic (LogP = 0.82). These results were similar to those obtained in experiments with yeast cytosine deaminase. In these experiments, Zn was shown to bind to a water molecule, which acted as a nucleophile, and to attack cytosine residue to affect deamination.[21-24] Similarly, in our experiments, the most hydrophilic moiety, ^{OHV}K, was associated with the maximum acceleration of deamination reaction, which indicates that photo-cross-linkers might bind to water molecule to facilitate nucleophilic attack. To determine the mechanism of acceleration of the deamination reaction and to understand the nature of the relationship between ODN structure and reaction rate, we calculated the activation energy values of photo-responsive ODN(X) using the reaction rates values at 50, 70, and 90 °C. Activation energy was inversely proportional to the rate of deamination reaction (Table 3.2). The highest value was observed for ODN(^{OMeV}K) and the lowest value for ODN(^{OHV}K). These results indicated that the hydrophilic substituent decreased the activation energy in photochemical transition of C to U.

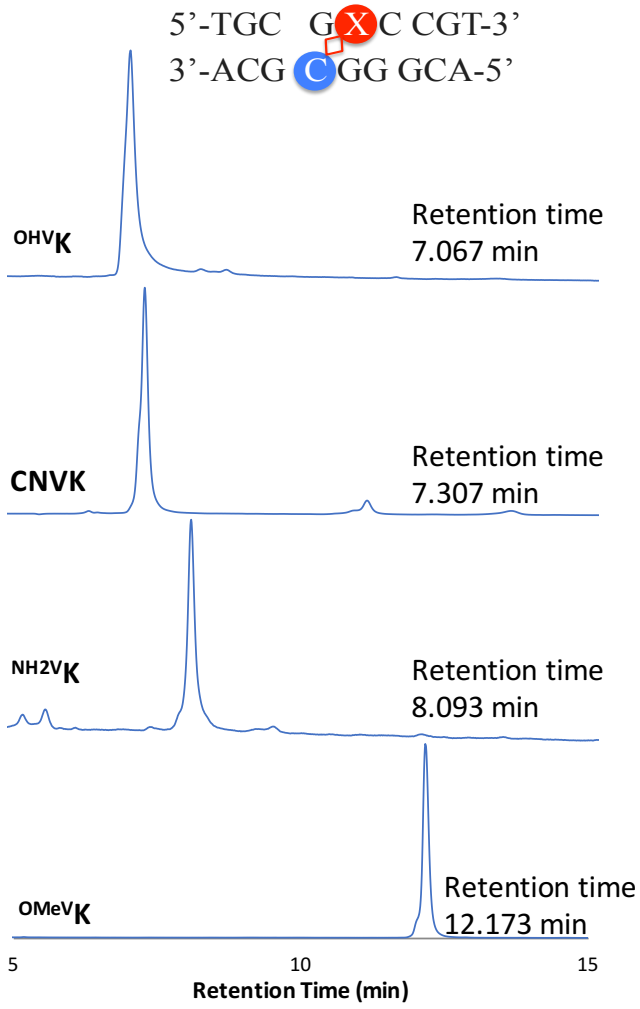


Figure 3.17: Comparison test of LogP by the HPLC method. HPLC chromatograms of the photo-cross-linking duplexes consisting of ODN(GK) and ODN(C). Elution; 50 mM ammmoniumformate/MeCN (100:0 to 60:40), Linear gradient (60 min) at a flow rate of 1.0 mL/min. The temperature of the column was maintained at 60 °C.

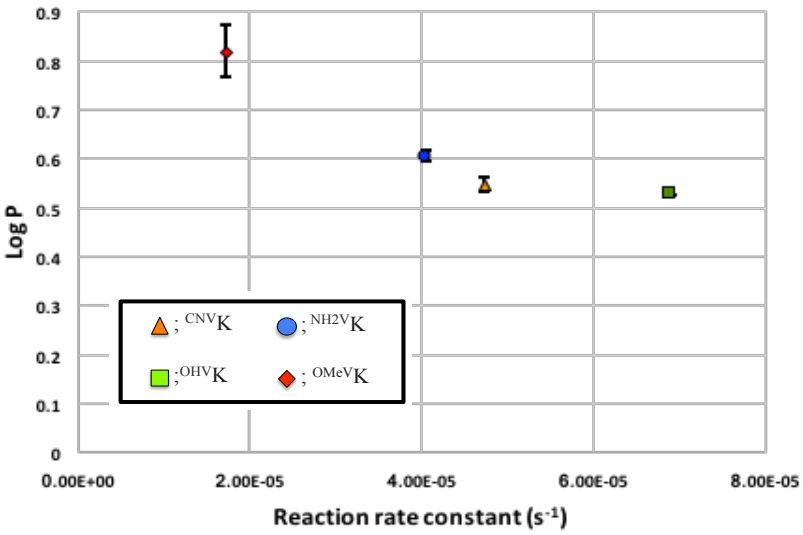


Figure 3.18: Relation between partition coefficient and rate of photo-cross-linking reaction.

Table 3.2: Partition coefficient, rate constant of deamination and activation energy

Entry	Retention time [min]	LogP		Deamination at 70°C [s^{-1}]		Activation energy [kcal/mol]	
^{OMeV}K	10.82	0.53	hydrophobicity	1.73E-5	slow	-21.94	high
^{NH2V}K	9.98	0.55		4.05E-5		-20.56	
^{CNV}K	9.78	0.61		4.37E-5		-20.35	
^{OHV}K	8.72	0.82	hydrophilicity	6.90E-5	fast	-20.20	low

Conclusion

In conclusion, we demonstrated that efficient C to U transformation can be achieved non-enzymatically at relatively low temperatures by using 3-cyanovinylcarbazole derivatives. The rate of deamination reaction is primarily dependent on polarity, hydrophilicity, and activation energy of the oligodeoxyribonucleotides that contain photo-responsive nucleosides. This study is the first step towards the synthesis of novel ultra-fast photo-cross-linkers that would mediate photochemical C→U transformations at physiological conditions, so that this technology can be applied to *in vivo* studies.

References

1. Ling R, Yoshida M, Mariano P S, **(1996)** *J. Org. Chem. 61,* 4439.
2. Lamont P J, Surtees R, Woodward C E, Leonard J V, Wood N W, Harding A E, **(1998)** *Arch. Dis. Child 79*, 22.
3. Watanabe W, Nobuta A, Tanaka J, Asaka M, **(1996)** *Int. J. Cancer 67*, 264.
4. Mali P, Yang L, Esvelt K M, Aach J, Guell M, DiCarlo J E, Norville J E, Church G M, **(2013)** *Science 339*, 823.
5. Jiang W, Bikard D, Cox D, Zhang F, Marraffini L A, **(2013)** *Nat. Biotechnol. 31*, 233
6. Sander J D, Joung J K, **(2014)** *Nat. Biotechnol. 32*, 347.
7. Javier N D, Michael N S, **(2002)** *Nature 419*, 43.
8. Nishida K, Arazoe T, Yachie N, Banno S, Kakimoto M, Tabata M, Mochizuki M, Miyabe A, Araki M, Hara K Y, Shimatani A, Kondo A, **(2016)** *Science 353*, 6305.
9. Pham P, Bransteitter R, Petruska J, Goodman M F, **(2003)** *Nature 424*, 103.
10. Bransteitter R, Pham P, Schardd M D, Goodman M F, **(2003)** *Proc. Natl. Acad. Sci. U. S. A. 100*, 4102.
11. Frommer M, McDonald L E, Millar D S, Collis C M, Watt F, Grigg G W, Molloy P L, Paul C L, **(1992)** *Proc. Natl. Acad. Sci. U. S. A 89*, 1827.
12. Tanabe K, Okamoto A, **(2007)** *Bioorg. Med. Chem. Lett. 17*, 1912.
13. Fujimoto K, Konishi-Hiratsuka K, Sakamoto T, Yoshimura Y, **(2010)** *Chem. Commun. 46*, 7545.
14. Fujimoto K, Konishi-Hiratsuka K, Sakamoto T, Yoshimura Y, **(2010)** *ChemBioChem. 11*, 1661.
15. Fujimoto K, Futamura D, Sakamoto T, **(2013)** *Chem. Lett. 11*, 5065.
16. Yoshimura Y, Fujimoto K, **(2008)** *Org. Lett.* 10, 3227.
17. Sakamoto T, Shigeno A, Ohtaki Y, Fujimoto K, **(2014)** *Biomater. Sci. 2*, 1154.
18. MacroModel, Schrödinger LLC. 2004. http://www.schrodinger.com.
19. OECD GUIDLINE FOR TESTING OF CHEMICALS 117. http://www.oecd.org/chemicalsafety/risk-assessment/1948177.pdf
20. Borges N M, Kenny P W, Montanari C A, Prokopozyk I M, Ribeiro J F, Rocha J R, Sartori G R, **(2017)** *J. Comput. Aided. Mo. Des. 31*, 163.
21. Sklenak S, Yao L, Cukier R I, Yan H, **(2005)** *J. Phys. Chem. B 109*, 7500.
22. Manta B, Raushel, F M, Himo F, **(2014)** *J. Phys. Chem. B 118*, 5644.

23. Zhang X, Zhao Y, Yan H, Cao X, **(2016)** *J. Computational Chem. 37*, 1163.

Chapter 4: Study of Photochemical Cytosine to Uracil Transition via Ultrafast Photo-Cross-Linking Using Vinylcarbazole Derivatives in Duplex DNA

Introduction

Nucleic acid chemistry and its biological applications have a great scope in the development of futuristic drugs and cure for many diseases.[1] Enzymatic methods for the nucleic acid manipulation, like PCR-plasmid based DNA manipulation has been used for many years to create directed mutations, recombination, deletion, and insertion of desired gene in the genome, but these methods have limited applications when *in vivo* applications are considered.[2] Recent developments in the field of genome editing has given us methods for genome editing using proteins, RNA, and other chemical agents. Currently, the major enzymatic techniques used for genome editing are meganuclease,[3] zincfinger nuclease (ZNF),[4] transcription activator-like effector-based nuclease (TALEN),[5] and clustered regularly interspaced short palindromic repeats associated system (CRISPR-Cas system).[6-9]

Currently, CRISPR-Cas system is considered one of the most advanced technique for specific genetic manipulation.[10] CRISPR-Cas9 is a two-component based system, wherein, one of the component is an endonuclease (Cas9) and a guide RNA (gRNA), a 20-nucleotide sequence that determines target specificity.[11] Despite high specificity of CRISPR system, unintended mutations at sites other than target sites is one of the major drawbacks of the system.[12]

Along with CRISPR, ZNF and TALEN have also been used for genome editing by using enzymes which cleaves the DNA at specific sites to induce double stranded breaks which lead to deletion and frame-shift mutations.[13,14] Although useful, these systems have shown cytotoxicity and cost of producing the enzymes is high.[15]

Besides enzymatic methods for DNA manipulation, chemical methods, like sodium bisulphite nucleobase editing have been developed.[16] Bisulphite based nucleobase editing is considered internecine technique as it requires deleterious chemical and the technique is highly non-specific.

Due to limitation in the chemical methods, certain photo-chemical methods were developed which use azbenzene and psolaren incorporated in the DNA strand and could undergo either photo-isomerization or photo-cross-linking by UV irradiation to form a stable cross-linked duplex.[17-20] These photo-active compounds have been used in the DNA to induce anti-sense effect and formation of stable DNA structures.[21-23]

Presence of limitations and drawbacks of the current techniques available for nucleic acid editing has led to new non-enzymatic and non-invasive techniques for manipulation of nucleic acids. In 2010, we reported a new enzyme free technique for nucleic acid editing which takes advantage of cytosine deamination to induce single point mutation in the DNA/RNA sequence

using photo-reactive 3-cyanovinylcarbazole nucleotide (^{CNV}K).[24-25] Upon incorporation of ^{CNV}K in oligodeoxyribonucleic acid (ODN), the ODN become responsive towards UV radiation and can undergo cross-linking with cytosine at -1 position with respect to ^{CNV}K in a sequence specific manner. Upon cross-linking, the aromaticity of cytosine is lost which makes it prone to nucleophilic attack by water to undergo deamination. The major drawback of this reaction is requirement of high temperature to accomplish deamination.

In 2015, we showed the effect of hydrogen bonding on rate of photo-cross-linking reaction by changing counter-base of cytosine,[26] which led to the idea that hydrogen bonding could also play role in deamination reaction. In 2017, we further proved our hypothesis by showing that certain counter-bases of cytosine are better for deamination than other bases due to a specific type of hydrogen bonding in the cross-linked cytosine.[27] Furthermore, we also reported that not only hydrogen bonding, but polarity around cytosine-photo-adduct is also responsible for the velocity of deamination reaction at lower temperatures using the derivatives of vinylcarbazole.[28]

In this report, we have shown the combined effect of hydrogen bonding and polarity by taking varying combinations of counter-base (inosine, guanine, and cytosine) and vinylcarbazole derivatives (3-cyanovinyl carbazole, 3-amidovinylcarbazole, and 3-carboxyl vinyl carbazole) and incorporating them in the ODN to accomplish deamination at physiological conditions (figure 4.1).

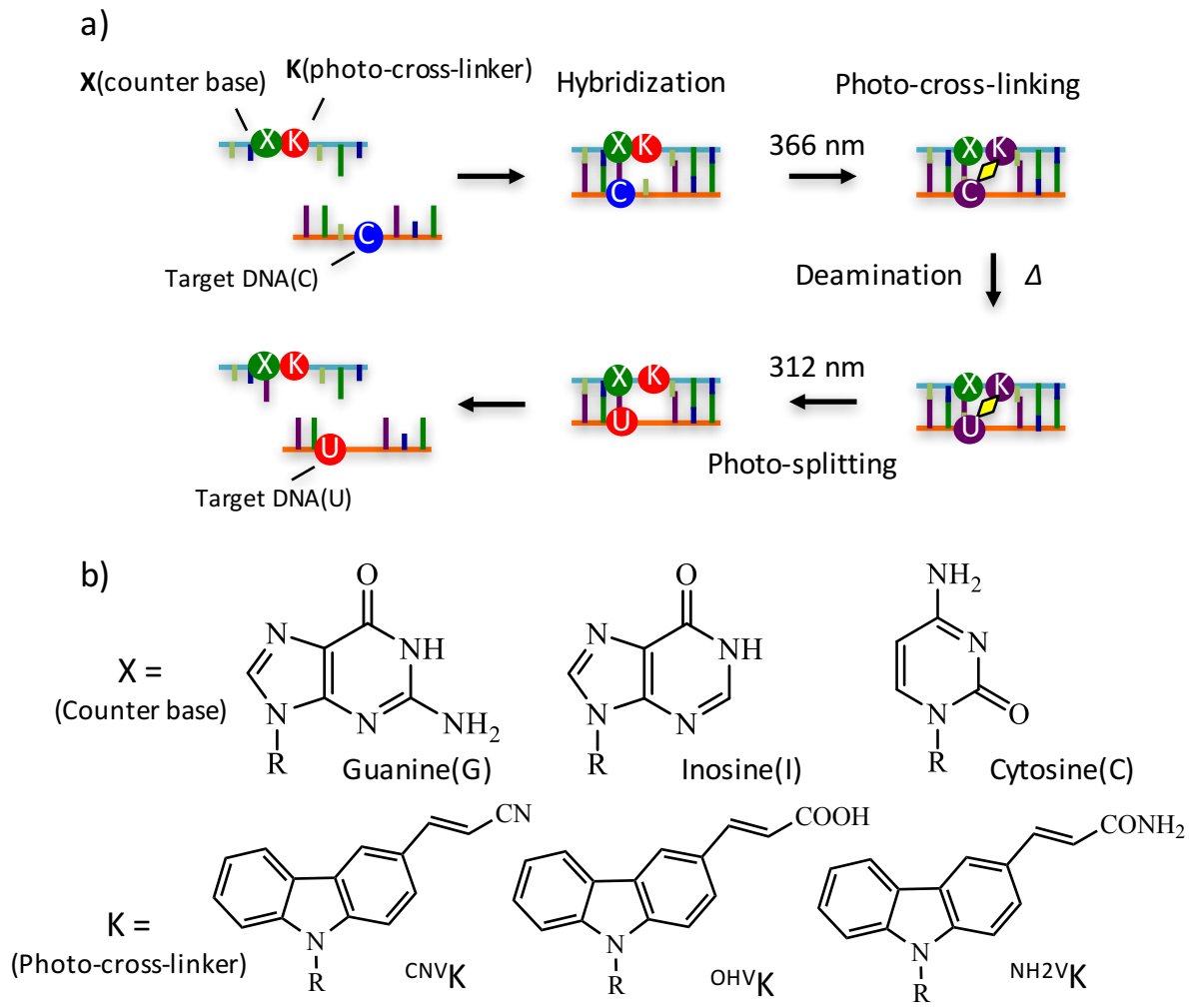


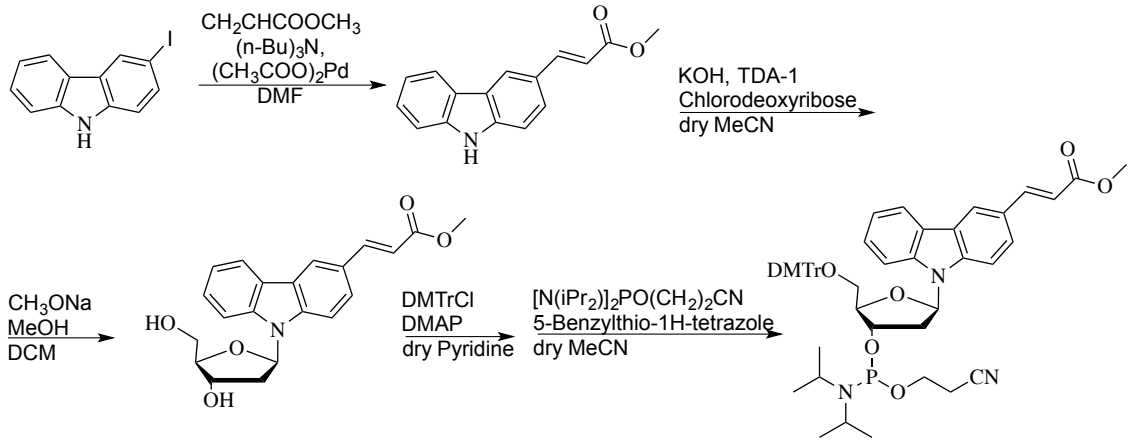
Figure 4.1: a) Overview of photo-cross-link based cytosine deamination. b) chemical structures of counter bases and photo-cross-linkers.

Materials and Methods

General.

1H NMR spectrum were recorded on a Bruker AVANCE III 400 system. Mass spectra were recorded on a Voyager PRO-SF, Applied Biosystems. HPLC was performed on a Chemcosorb 5-ODS-H column with JASCO PU-980, HG-980-31, DG-980-50 system equipped with a JASCO UV 970 detector at 260 nm. Reagents for the DNA synthesizer such as A, G, C, T-β-cyanoethyl phosphoramidite, and CPG support were purchased form Glen research.

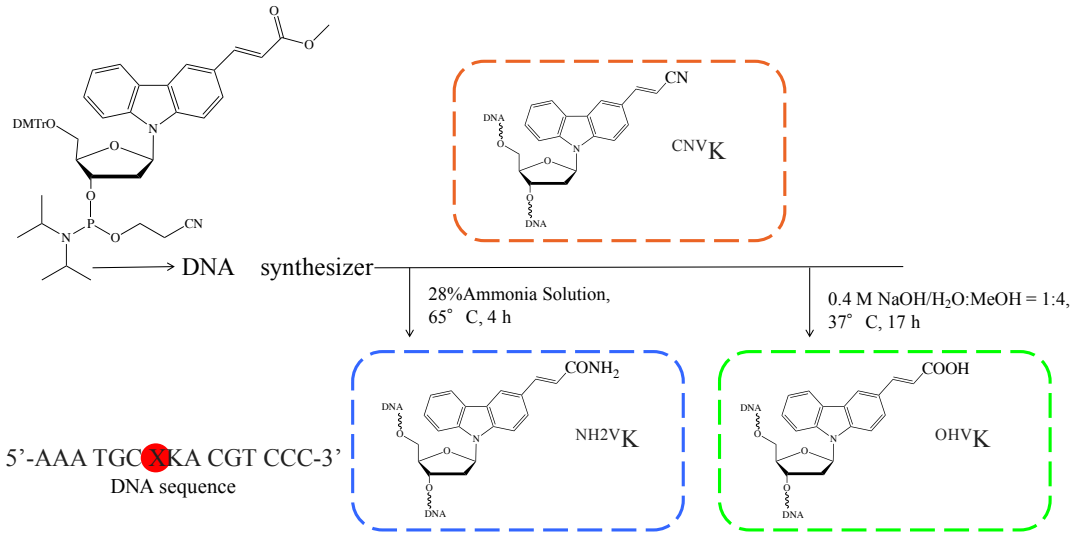
Synthesis and preparation of oligonucleotides



Scheme 4.1: Synthetic scheme for preparation of photo-responsive nucleotide

The phosphoramidite of ^{OMeV}K and ^{CNV}K was prepared per the scheme (Scheme 4.1) and previous literature.[26]

Post-modification of Oligonucleotides



Scheme 4.2: Post-modification of photo-cross-linker in ODN

Preparation of modified oligonucleotides.

The phosphoramidite of ^{CNV}K was prepared following to previous reports.[33] The modified oligonucleotides containing ^{CNV}K or ^{OMeV}K were prepared, per standard phosphoramidite chemistry using DNA synthesizer (ABI 3400 DNA synthesizer, Applied Biosystems, CA). The ODN containing ^{OMeV}K was post modified to ODN containing ^{OHV}K or ^{NH2V}K in deprotection step (Scheme 4.2). Synthesized ^{CNV}K ODNs were detached from the support by soaking in concentrated aqueous ammonia for 1 h at room temperature. Deprotection was conducted by heating the concentration aqueous for 4 h at 65°C concentrated aqueous ammonia was then removing it by speedvac, and the crude oligomer was purified by reverse phase HPLC equipped with InertSustainTM C18 column Cosmosil^{TM} 5C$_{18}$-AR-II column (5 µm, 10 × 150 mm, Nacalai tesque, Flow rate of 3.0 mL/min, 60°C) and lyophilized. Synthesis of ODN was confirmed by MALDI-TOF-MS (Table 4.1). Other ODNs were purchased from Fasmac (Japan).

Table 4.1: MALDI-TOF-MS analysis of ODNs.

ODN(XK)	Calculated Mass $(M+H)^+$	Experimental Mass $(M+H)^+$
$C^{CNV}K$	4587.84	4586.43
$C^{NH2V}K$	4605.22	4607.21
$C^{OHV}K$	4592.87	4592.00
$G^{CNV}K$	4627.84	4625.90
$G^{NH2V}K$	4645.33	4644.25
$G^{OHV}K$	4632.08	4630.44
$I^{CNV}K$	4610.97	4608.23
$I^{NH2V}K$	4629.65	4631.02
$I^{OHV}K$	4617.17	4616.19

Isolation of photo-cross-linked dsDNA.

The cODN (C) (15 μM) and ODN (XK), where X is the counter base and K is the photo-cross-linker, (15 μM) in buffer solution (100 mM NaCl, 50 mM sodium cacodylate, pH 7.6) was photoirradiated at 366 nm for 120 sec using UV-LED illuminator(OMRON Inc., 1600 mW) at 37 ºC. And the solution was purified by reverse phase HPLC, and the concentration of photo-cross-linked dsDNA was determined by absorbance at 260 nm.

Deamination

The solution of 15 μM ODN(XK) and ODN(C) in buffer solution was annealed and photoirradiated at 366 nm. The photoirradated solution was purified by a reversible HPLC to get the purified photo-cross-linked dsDNA. The 5 μM photo-cross-linked dsODN in a buffer solution (50 mM Na-cacodylate buffer (pH 7.4) containing 100 mM NaCl) was incubated at 37°C.

Photo-splitting and UPLC analysis

After the photo-splitting with irradiation of 312 nm (15 min at 37°C, transillluminator, Funakoshi), reaction mixture was analyzed with UPLC system (Aquity, Waters); elution was

with 0.05 M ammonium formate containing 3-6.5% CH_3CN, linear gradient (10 min) at a flow rate of 0.2 mL/min.

Partition coefficient (LogP)

LogP of photo-cross-linked DNA was measurement by their retention time following OECD protocol. 10 μM photo-cross-linked dsDNA was analyzed with HPLC system; elution was with 0.05 M ammonium formate containing 98-50% CH_3CN, linear gradient (60 min) at flow rate of 1 mL/min. 4-Acetylpyridine, Aniline, Acetanilide, Phenol, Benzonitrile, and Acetophenone were used as control compound to create calibration curve.[29-30]

Enzymatic digestion

ODNs after photo-splitting were incubated with 0.2 unit/μl Nuclease P1 for 24 h at 37ºC. Further, the solution was treated with alkaline phosphatase (0.1 unit/μl) for 24 h at 37ºC and analyzed with HPLC system. Elution was with 1-20% MeCN in 50 mM ammonium formate buffer for 30 min at linear gradient with column temperature 60ºC and flow rate of 1 ml/min.

Results and Discussion

Deamination reaction

In order to find the best combination of photo-active nucleobase and counter-base of the target, various combinations were studied for deamination. For counter base: inosine, guanine, and cytosine were chosen. Whereas, for photo-active nucleobase, 3-cyanovinyl carbazole (^{CNV}K), along with two derivatives of ^{CNV}K: 3-amidevinyl carbazole and 3-carboxyl vinyl carbazole were chosen. In our previous reports, these bases have shown promising results when studied singularly, therefore, we chose the combinations based on the previous singular studies.

Firstly, 15mer ODNs containing the various combinations, namely: G^{CNV}K, G^{NH2V}K, G^{OHV}K, C^{CNV}K, C^{NH2V}K, C^{OHV}K, I^{CNV}K, I^{NH2V}K, and I^{OHV}K (table 4.2), were synthesized using DNA synthesizer (AB 3600, USA), The ODNs were designed in such a way that the counter base gets placed complimentary to the target cytosine and photo-active nucleotide gets placed at -1 position with respect to target cytosine present in the complimentary strand.

The two ODNs, cODN(C) and ODN(XK) (15 µM each), where X=G, C, and I; and K=^{CNV}K, ^{NH2V}K, and ^{OHV}K, in 50 mM sodium cacodylate buffer having 100 mM sodium chloride (pH 7.4) were irradiated with 366 nm UV radiation for 120 seconds at 37 ºC using Omron LED. The ODNs were cross-linked in 60 seconds with a yield of >90% as analyzed by ultra-high performance liquid chromatography (Fig 4.2). These photo-adducts were purified using high performance liquid chromatography (HPLC).

Table 4.2: Sequence of photo-active ODNs

ODN(XK)	**ODN sequence (5'→3')**
C^{CNV}K	AAATGCC^{CNV}KACGTCCC
C^{NH2V}K	AAATGCC^{NH2V}KACGTCCC
C^{OHV}K	AAATGCC^{OHV}KACGTCCC
G^{CNV}K	AAATGCG^{CNV}KACGTCCC
G^{NH2V}K	AAATGCG^{NH2V}KACGTCCC
G^{OHV}K	AAATGCG^{OHV}KACGTCCC
I^{CNV}K	AAATGCI^{CNV}KACGTCCC
I^{NH2V}K	AAATGCI^{NH2V}KACGTCCC
I^{OHV}K	AAATGCI^{OHV}KACGTCCC

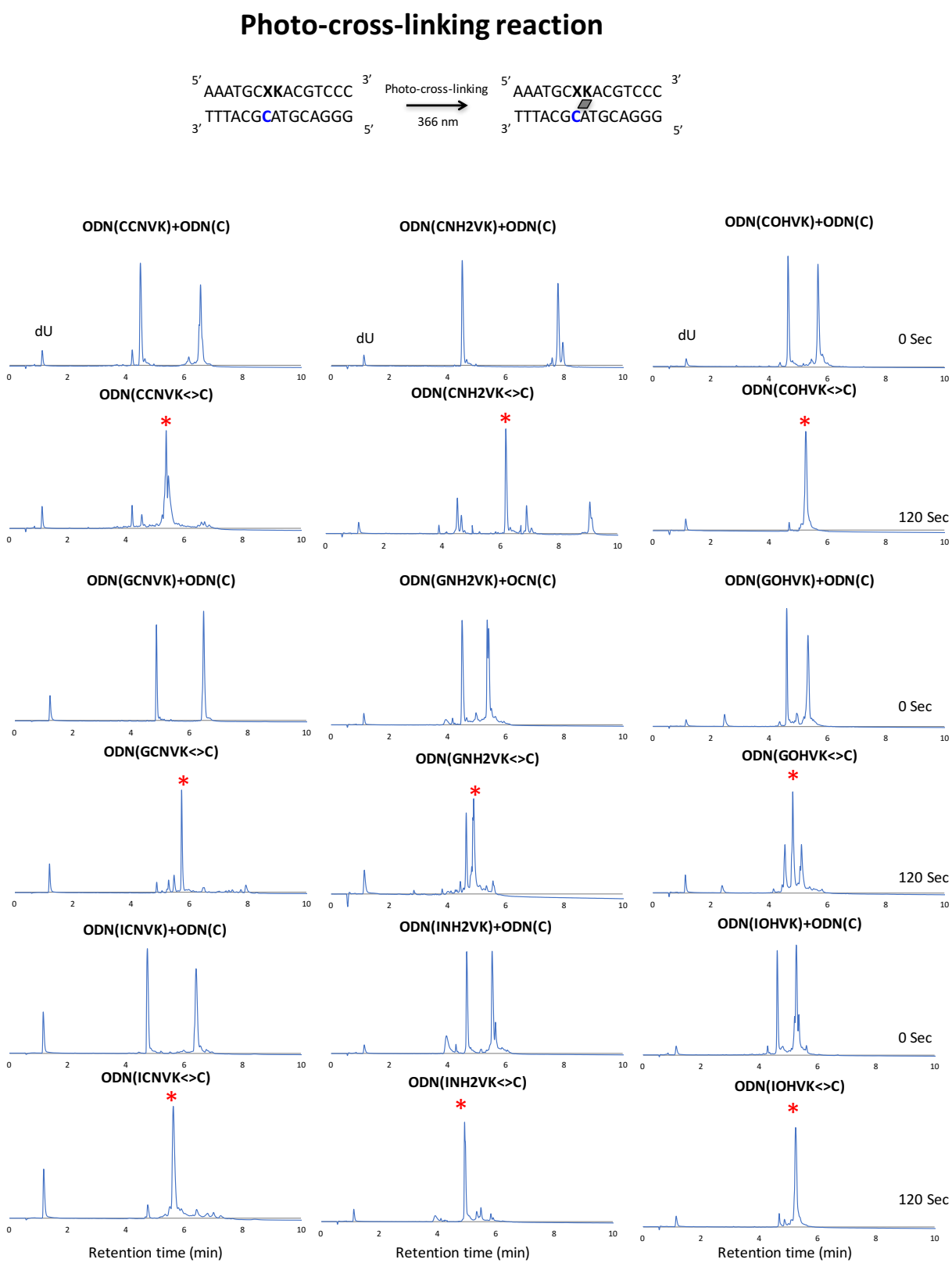


Figure 4.2. UPLC chromatograms of the mixture of OGN(XK) and cODN(C). [ODN] = 10 µM in 50 mM sodium cacodylate buffer (pH 7.4) containing 100 mM NaCl. Photoirradiation at 366 nm was performed at 4ºC. 50 µM 2'-deoxyuridine (dU) was used as an internal standard. Peaks marked with asterisk (*) are the newly formed photo-adducts. X= G, C, I. K= ^{OHV}K and ^{NH2V}K.

The photo-adducts ODN(XK<>C) (5 µM) in 50 mM sodium cacodylate buffer and 100 mM sodium chloride (pH 7.4) were incubated for 7 days at 37 ºC. Samples were taken out at interval of 1 day and the ODNs were photo-split using 312 nm UV radiation for 15 min at 37 ºC. The photo-split ODNs were then analyzed by UPLC. The conversion of C→U was studied by comparing peak areas under the curves of UPLC depicting the individual cODN(C) and cODN(U) shown in the figure 4.3-4.5. The photo-split ODNs were subjected to enzymtic digestion using nuclease P1 and alakaline phosphatase (Figure 4.7).

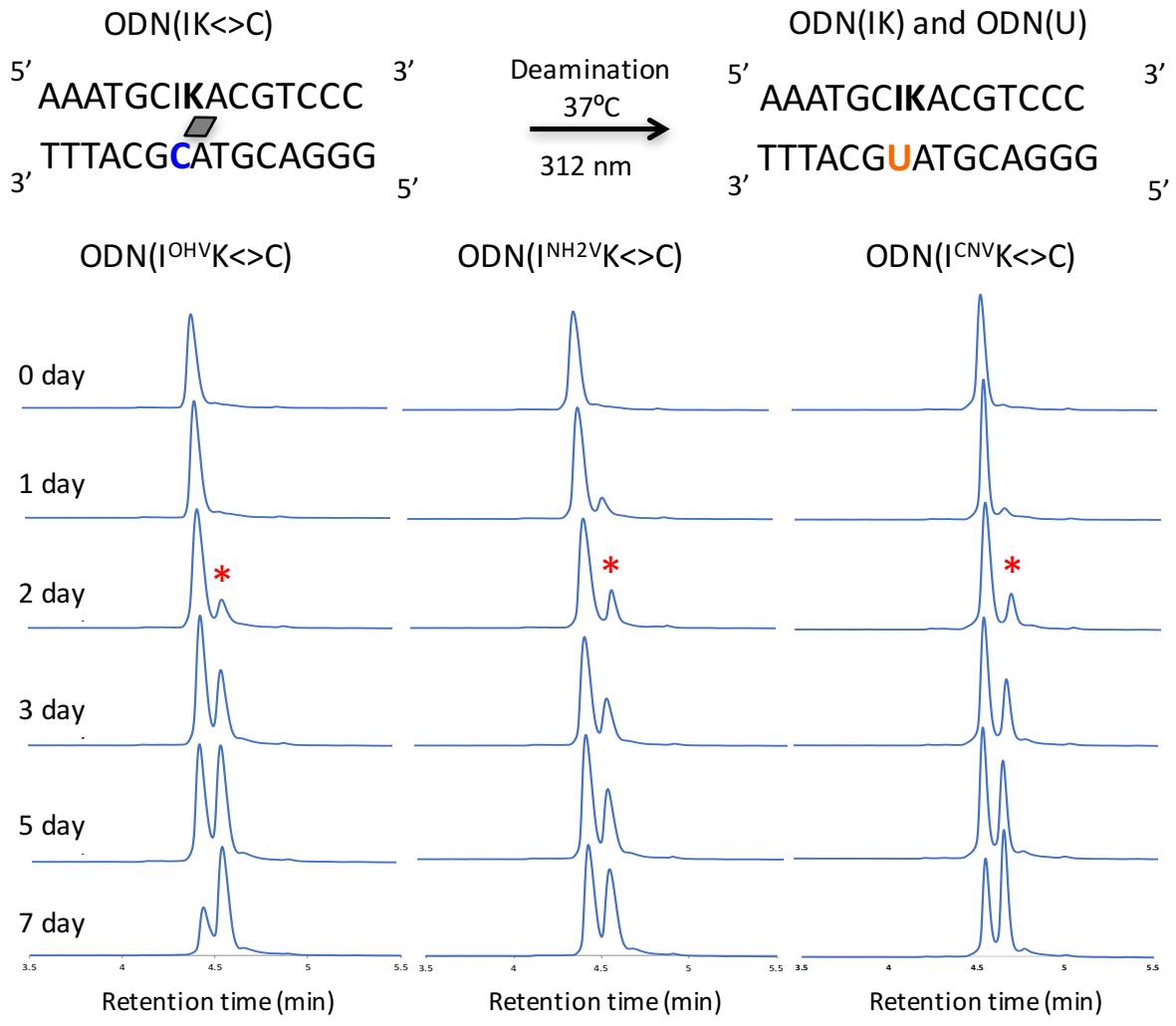


Figure 4.3. UPLC analysis of the deamination reaction of the photo-cross-linked duplex consisting of ODN(IK) and cODN(C). [ODN(IK<>C)] = 5 µM in 50 mM Na-cacodylate buffer (pH 7.4) containing 100 mM NaCl. incubated at 37ºC, photo splitting was performed with transilluminator (312 nm) at 37ºC. Peak marked with asterisk indicate the newly formed ODN(U).

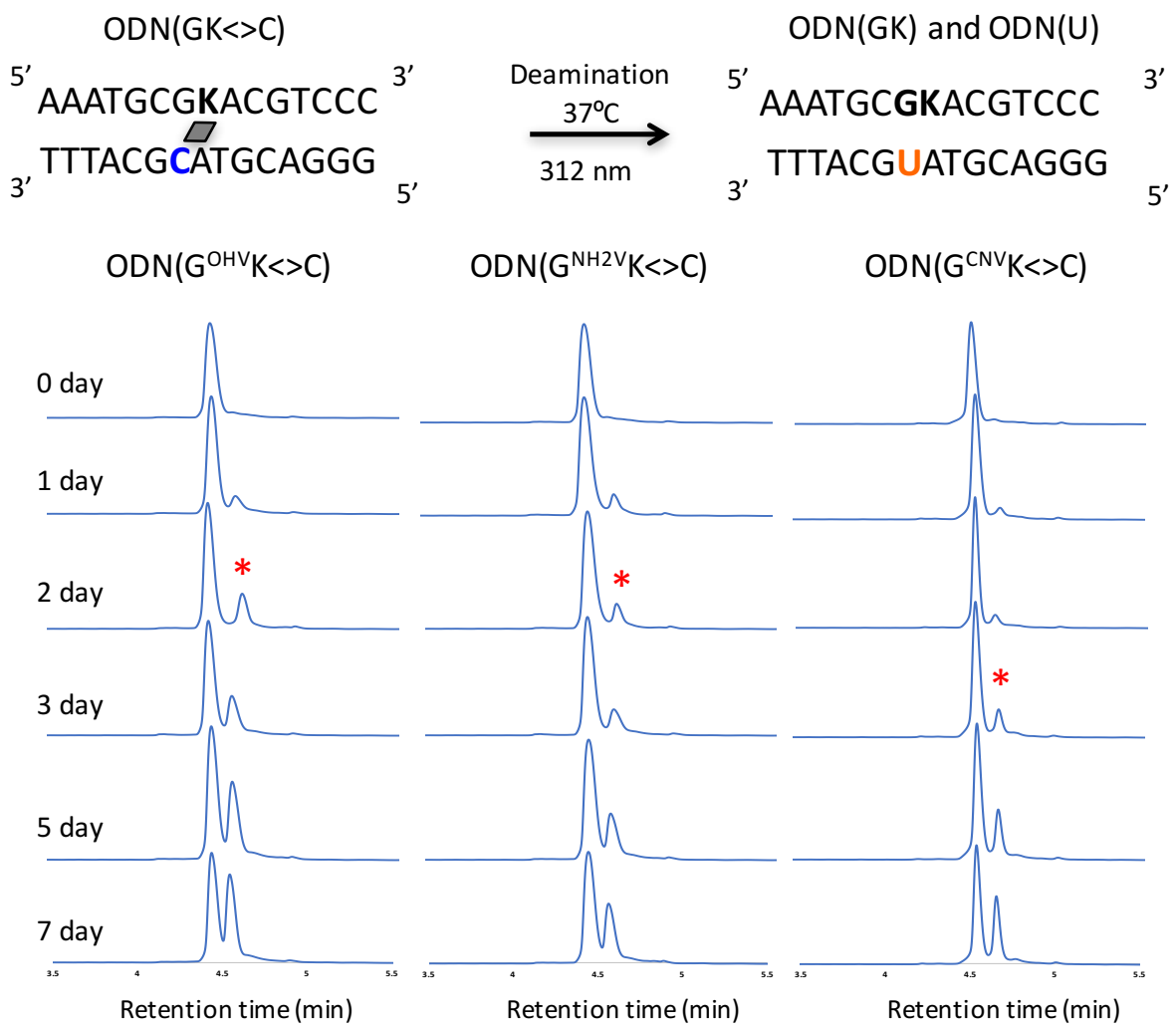


Figure 4.4. UPLC analysis of the deamination reaction of the photo-cross-linked duplex consisting of ODN(GK) and cODN(C). [ODN(GK<>C)] = 5 μM in 50 mM Na-cacodylate buffer (pH 7.4) containing 100 mM NaCl. incubated at 37ºC, photo splitting was performed with transilluminator (312 nm) at 37ºC. Peak marked with asterisk indicate the newly formed ODN(U).

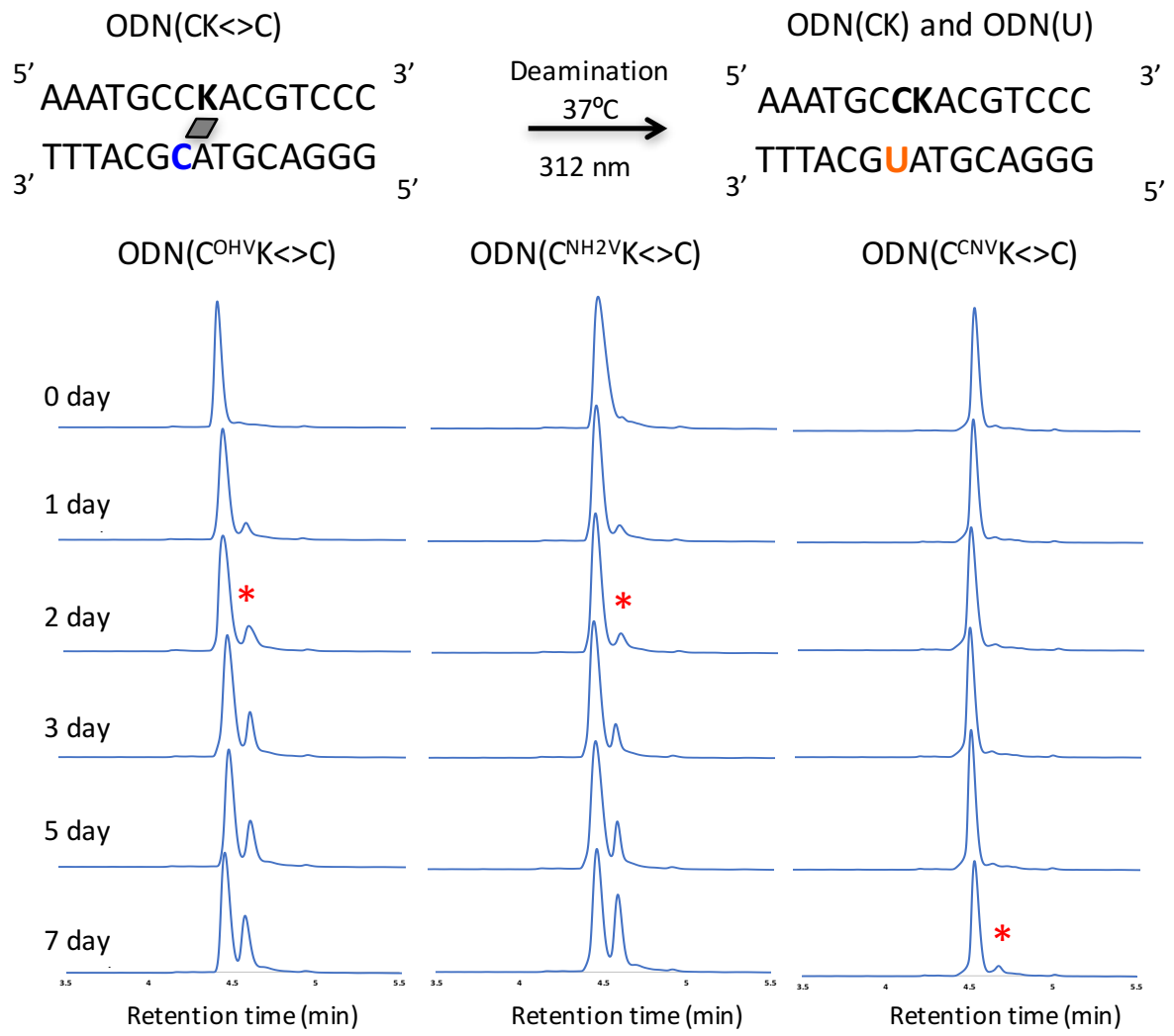


Figure 4.5. UPLC analysis of the deamination reaction of the photo-cross-linked duplex consisting of ODN(CK) and cODN(C). [ODN(CK<>C)] = 5 µM in 50 mM Na-cacodylate buffer (pH 7.4) containing 100 mM NaCl. incubated at 37ºC, photo splitting was performed with transilluminator (312 nm) at 37ºC. Peak marked with asterisk indicate the newly formed ODN(U).with asterisk indicate the newly formed ODN(U).

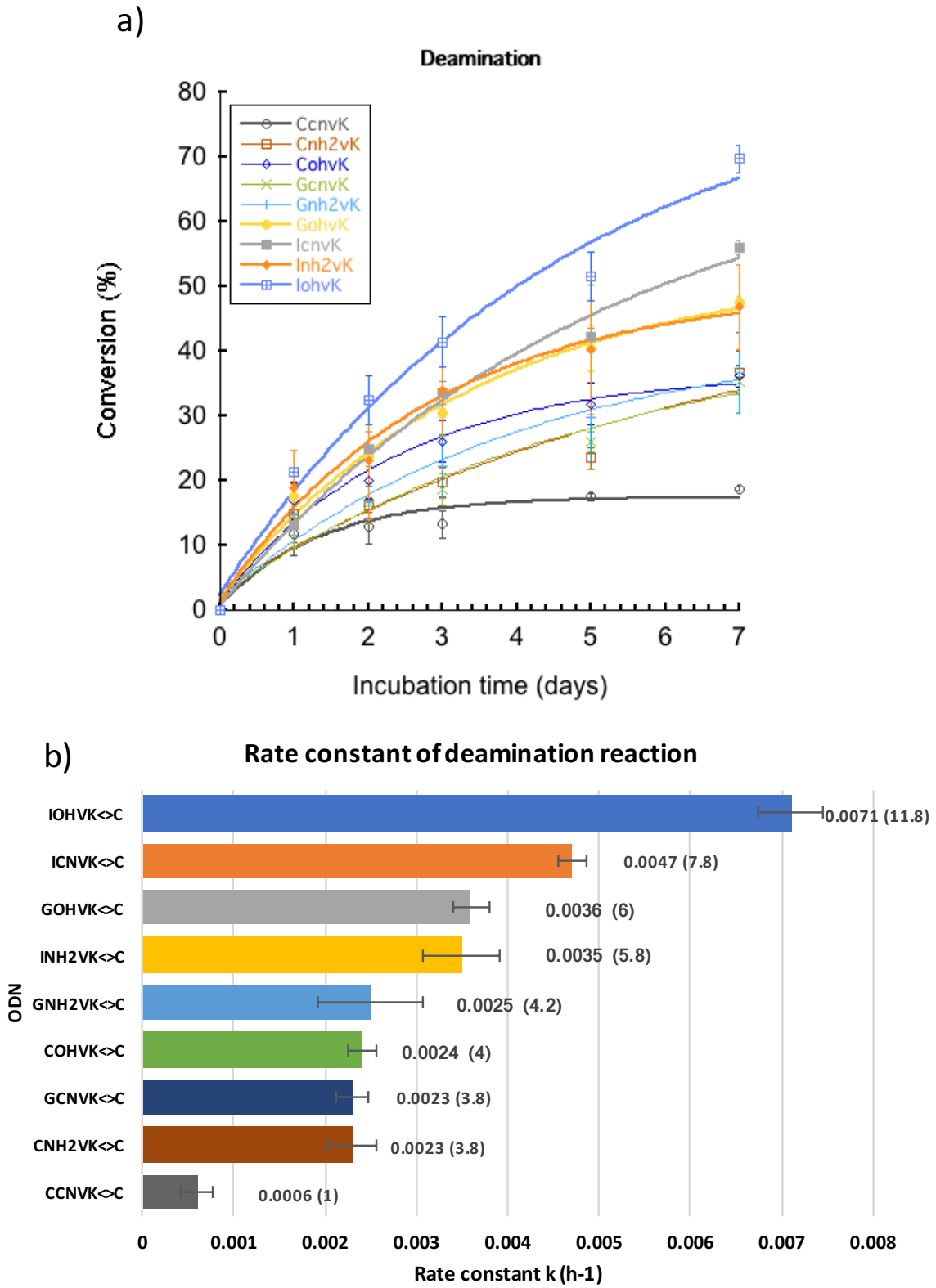


Figure 4.6. (a) Time course of the deamination reactions of the photo-adducts ODN(XK<>C). (b) Deamination reaction rate constants of the photo-adducts ODN(XK<>C). The values besides the bars are reaction rate constants and the values in braces indicate the acceleration ratio compared to the case of $G^{CNV}K$. Photoreaction rate constants were estimated from the time course of the deamination reaction with an assumption of first-order reaction kinetics.

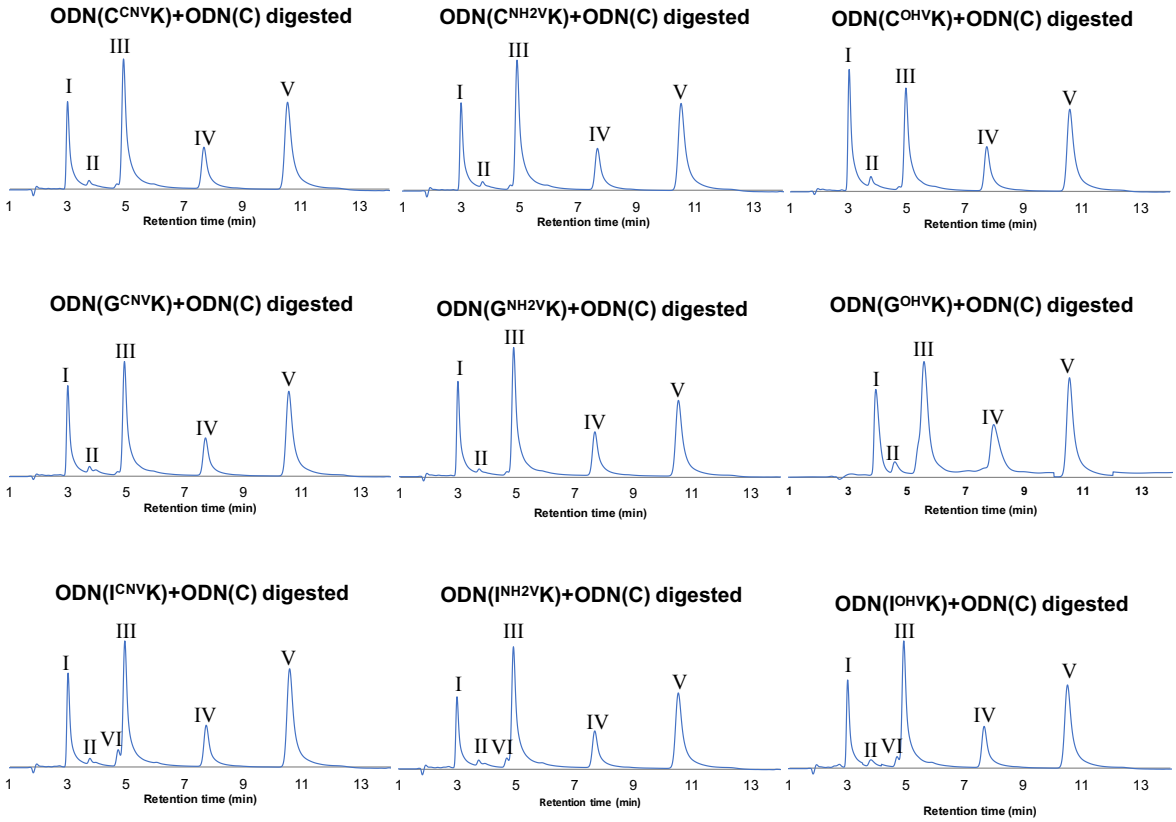


Figure 4.7: Enzymatic digestion of photo-splitted ODNs after deamination reaction using nuclease P1 and alkaline phosphatase. The peak ratio indicated the amount of nucleoside in the ODN. I=dC, II=dU, III=dG, IV=dT, V=dA, VI=dI. HPLC conditions: 1-20% MeCN in 50 mM ammonium formate, liner gradient, for 30 min, column temperature 60ºC.

In case of Inosine as counter base, the deamination occurred by incubation at 37 ºC using ODN($I^{OHV}K$, $I^{NH2V}K$, and $I^{CNV}K$). The peak identical to ODN(C) was observed before incubation. The peak was decreased on incubation and the new peak identical to ODN(U) was appeared. Surprisingly, in case of ODN($I^{OHV}K$), the deamination occurred approximately 70% upon incubation at 37℃ for a week (figure 4.6). We observed that the deamination proceeded at the maximum rate (k = 7.1 x 10^{-3} h^{-1}) in case of ODN($I^{OHV}K$). Deamination reaction rates were intermediate when ODN($I^{NH2V}K$) and ODN($I^{CNV}K$) were used k = 3.5 x 10^{-3} h^{-1} and 4.7 x 10^{-3} h^{-1}, respectively. Thus, with ODN($I^{OHV}K$), the rate of deamination of cytosine to uracil was 2-fold faster than that with ODN($I^{NH2V}K$) (figure 4.6).

In case of guanine as counter base, the deamination occurred by incubation at 37℃ using ODN ($G^{OHV}K$, $G^{NH2V}K$, and $G^{CNV}K$). The peak identical to ODN(C) was observed before incubation. The peak was decreased upon incubation and the new peak identical to ODN(U) was appeared.

In the case of cytosine as counter base, deamination occurred by incubation at 37ºC using ODN(C^{CNV}K<>C), ODN(C^{OHV}K<>C), and ODN(C^{NH2V}K<>C). A peak identical to that of ODN(C) was observed before incubation. A peak decreased on incubation and a new peak identical to ODN(U) appeared. It can be noted that the highest conversion of C→U was achieved by ODN(I^{OHV}K<>C) (~70%) in 7 days. In general, among all the counter-bases, inosine containing ODNs showed maximum C→U conversion in combination with the photo-cross-linkers. ODNs containing cytosine and guanine showed almost similar reaction rates for all the photo-cross-linkers. The difference in the rate of deamination in the photo-cross-linked ODNs can be attributed to hydrogen bonding pattern. Due to cross-linking, the overall planarity of cytosine is perturbed and a tilt appears in the non-planar part of cytosine towards the central axis pertaining to change in the hydrogen bonding pattern. These changes lead to a hydrogen bond between the adjacent base of cytosine and the amino group of cytosine, specifically observed in case of ODN with inosine. These hydrogen bonding patterns are comparable to hydrogen bonding observed near the active site of yeast cytosine deaminase (yCD) wherein a special network of hydrogen bonds has been observed in the enzyme.[32-33] This special network of hydrogen bonding is critical for the protonation of N3 of cytosine through the Glu64 and conformational changes of cytosine giving accessibility to water attack facilitated by zinc-bound water on C4 of cytosine. The patterns observed in the cross-linked cytosine of hydrogen bonding are equivalent to these patterns observed in yCD. Prior to cross-linking, the hydrogen bonding is according to Watson-Crick base-pairing, whereas, after cross-linking, due to change in planarity of cytosine, the hydrogen is with adjacent base facilitating the nucleophile attack on C4. Furthermore, the hydrogen bonding N3 is also crucial for proton shuttling.

Therefore, the number of hydrogen bonds in the photoadduct directly affected the rate of deamination, wherein, fewer H-bonds between the target cytosine and counter-base facilitated the deamination reaction. Upon comparing the photo-cross-linkers, the same pattern was observed. In this photochemical deamination, the hydrophilicity around the C4 in the target cytosine is the important factor

because of cytosine deamination requires water molecules. To assess the correlation between the hydrophilicity of photo-cross-linked dsDNA and rate of deamination reaction, we calculated the partition coefficient (LogP) of photo-cross-linked dsDNA using their HPLC retention times (Table 2). The hydrophilicity of ODNs positively correlated with their ability to convert cytosine to uracil in a photo-responsive manner. Photo-cross-linked ODN(I^{OHV}K<>C), ODN(G^{OHV}K<>C), and ODN(C^{OHV}K<>C) have the highest hydrophilicity

value (LogP = 0.52, 0.54, and 0.57 respectively) among all the ODNs due to presence of highly hydrophilic photo-cross-linker ^{OHV}K (table 4.2 and Figure 4.8). Among these ODNs, it has been observed that the highest rate of cytosine to uracil conversion has been observed in ODN(I^{OHV}K<>C) having the highest hydrophilicity. Similar pattern for deamination of cytosine has been observed in the ODNs containing ^{NH2V}K. The highest rate has been observed in ODN(I^{NH2V}K<>C) followed by ODN(G^{NH2V}K<>C) and ODN(C^{NH2V}K<>C), having their LogP in the similar order, i.e. highest hydrophilicity is shown by ODN(I^{NH2V}K<>C) among these three. Even in the ODNs with ^{CNV}K, the relation between LogP (hydrophilicity) and deamination reaction rate constant is positive correlation (Table 4.3, Figure 4.9, figure 4.10). ODNs containing ^{OHV}K are the most hydrophilic therefore shows highest rate of cytosine to uracil conversion. Thus, upon careful analysis of the results it is evident that the two major factors required for high rate of deamination in the vinyl carbazole based photo-cross-link assisted cytosine deamination are hydrogen bonding in the target cytosine and the hydrophilicity of the photo-cross-linker. Upon combining these two factors, the maximum C→U conversion can be achieved by using ODN containing inosine and ^{OHV}K at physiological conditions, therefore ODN(I^{OHV}K<>C) having the highest hydrophilicity and optimum hydrogen bonding has shown highest rate of deamination among all the other ODNs. Thus, ODN(I^{OHV}K<>C) was then further incubated for 20 days at 37ºC to determine the extent of deamination that can be achieved in this period. It has been observed from the UPLC chromatogram that the deamination reached a stationary phase, and the reaction proceeded only minimally in the later stage of deamination (Figure 4.11).

Table 4.2: Partition coefficients (Log P)

Entry	**Retention time(min)**	**Log P**	
I^{OHV}K	4.49	0.56	**Hydrophilicity (decreasing)**
G^{OHV}K	4.96	0.64	
C^{OHV}K	5.31	0.70	
I^{NH2V}K	5.38	0.71	
G^{NH2V}K	5.66	0.76	
I^{CNV}K	6.00	0.820	
G^{CNV}K	6.04	0.83	

$C^{CNV}K$	6.47	0.90	
$C^{NH2V}K$	7.78	1.13	

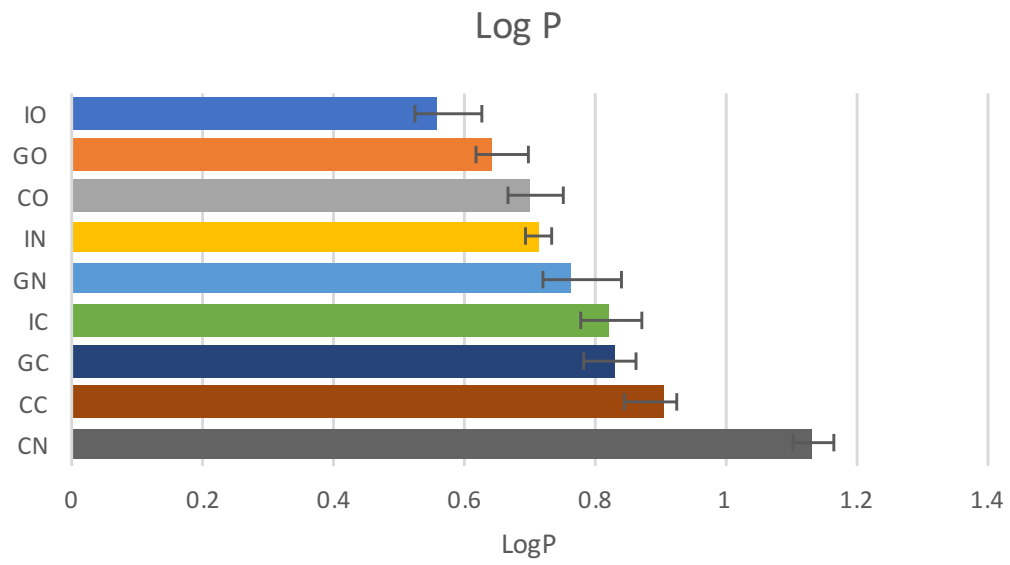


Figure 4.8: Bar graph showing correlation between ODN and Log P values in increasing order of hydrophobicity.

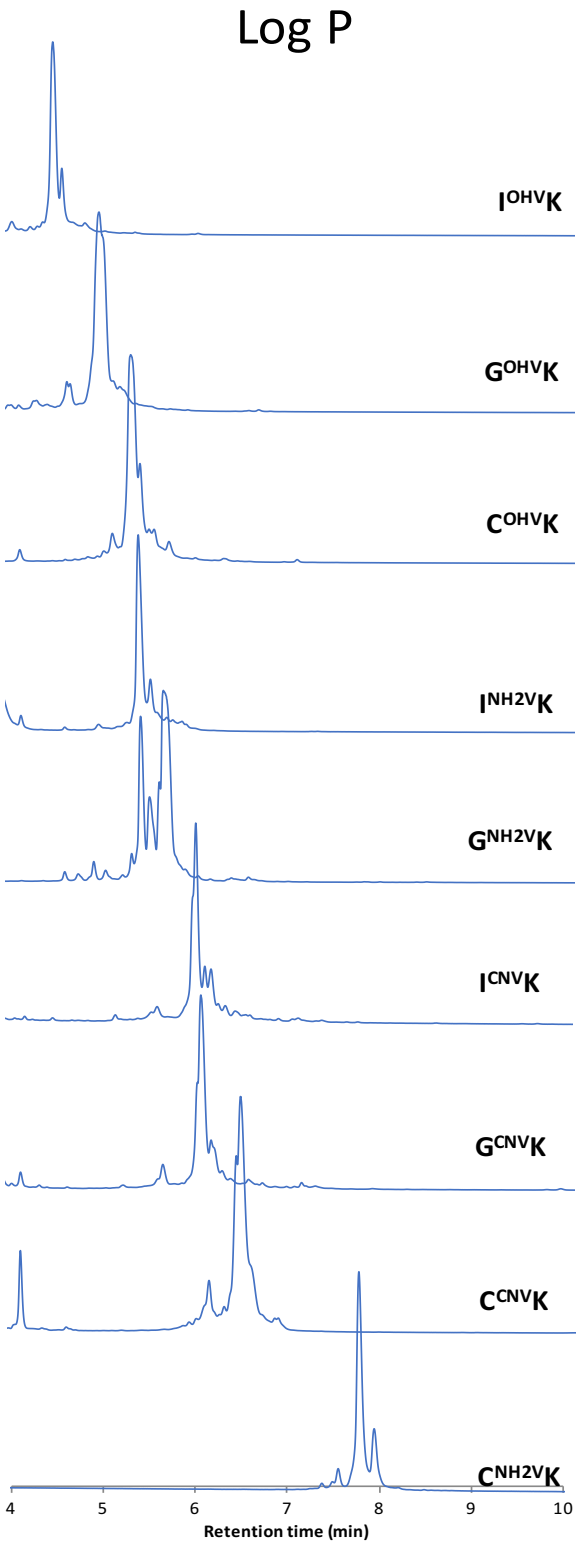


Figure 4.9: UPLC chromatograms showing retention time of individual ODN in 1 to 15 % MeCN gradient over 12 min linearly in ammonium formate (50 mM).

Table 4.3: Log P and reaction rate constant correlation.

Entry	Retention time (min)	Log P	Deamination rate constant (h^{-1})
$I^{OHV}K$	4.49	0.56	0.0071
$G^{OHV}K$	4.96	0.64	0.0036
$C^{OHV}K$	5.31	0.70	0.0024
$I^{NH2V}K$	5.38	0.71	0.0035
$G^{NH2V}K$	5.66	0.76	0.0025
$I^{CNV}K$	6	0.82	0.0047
$G^{CNV}K$	6.04	0.83	0.0023
$C^{CNV}K$	6.47	0.90	0.0006
$C^{NH2V}K$	7.78	1.13	0.0023

Hydrophilicity (decreasing)

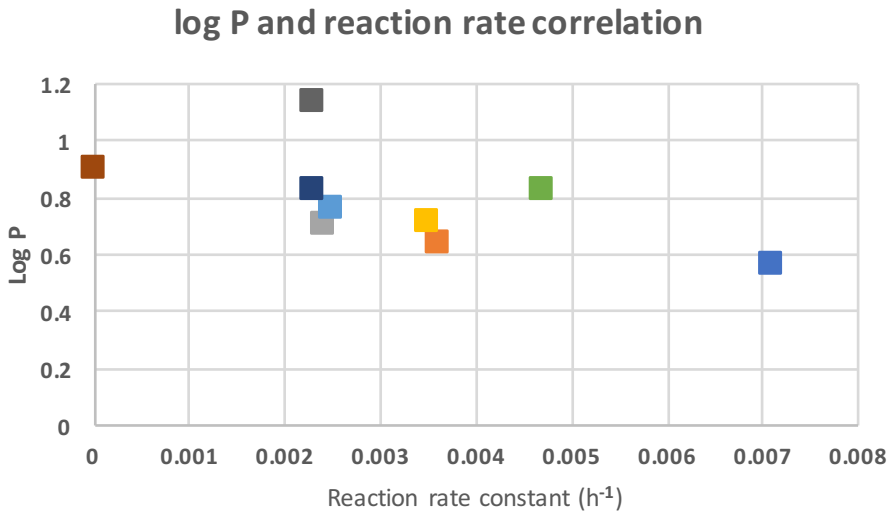


Figure 4.10: Correlation of LogP and reaction rate constant.

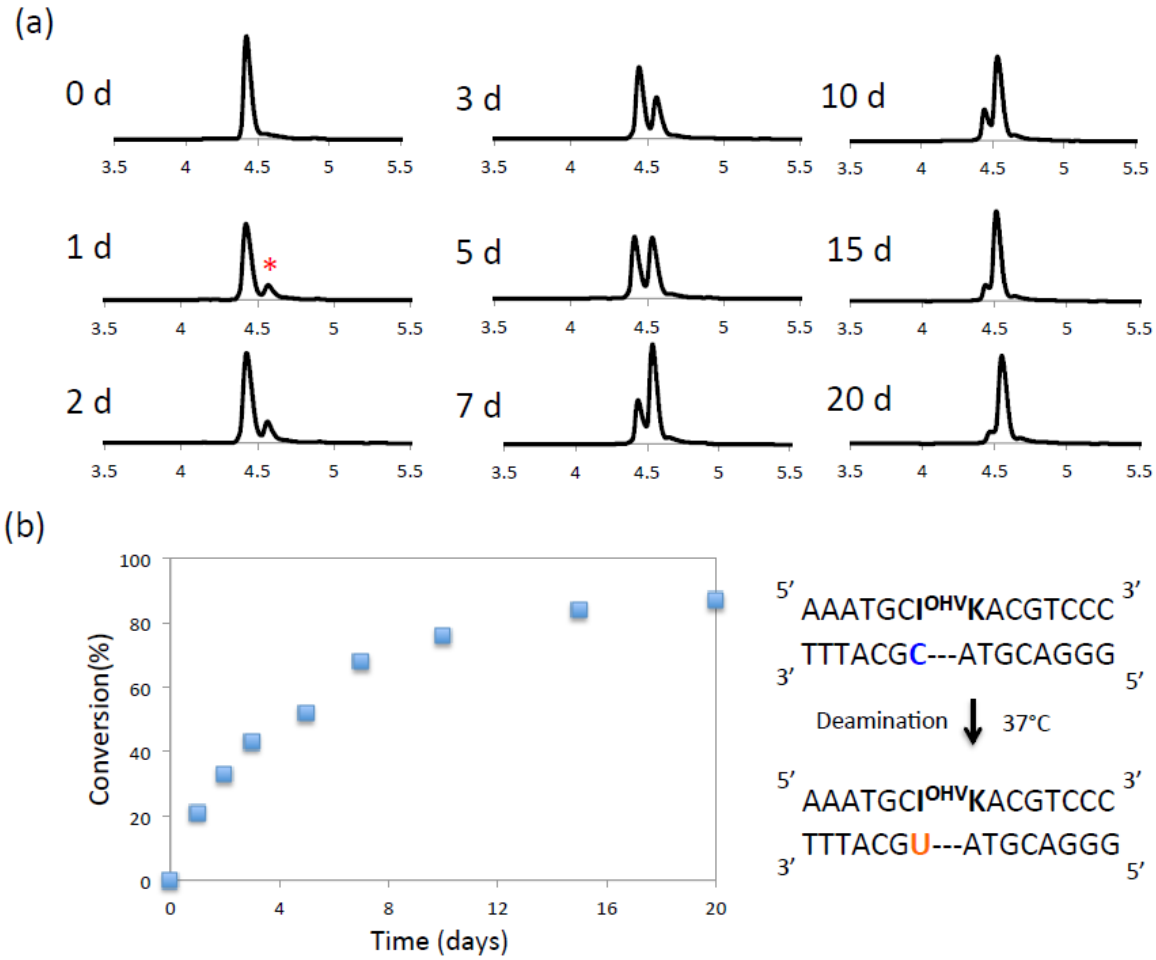


Figure 4.11: a) UPLC analysis of the deamination reaction of the photo-cross-linked duplex consisting of ODN(IK) and cODN(C) for a duration of 20 days. [ODN(IK<>C)] = 5 µM in 50 mM Na-cacodylate buffer (pH 7.4) containing 100 mM NaCl. incubated at 37ºC, photo splitting was performed with transilluminator (312 nm) at 37ºC. Peak marked with asterisk indicate the deaminated product. b) Scheme and sequence of deamination, and time course of deamination reaction.

Therefore, 7 days reaction which gives ~70% conversion is most optimum for *in vivo* applications. In the past, the deamination reaction used to occur at 90ºC using ^{CNV}K cross-linked cytosine. It has come all the way from that point to the reaction now taking place at physiological conditions, which will lead to realization of this technique in the *in vivo* application and further in the clinical scenario. This technique will be used in the future for the treatment of skin and other cancers along with treatment of disorders arising due to genetic mutations, e.g. Leigh's syndrome. We will now be focusing on using 366 nm UV radiation instead of 312 nm due to slightly invasive nature of 312 nm UV radiation by employing DNA-strand displacement technique for photo-splitting.

Conclusions

The major factors responsible for the rate of 3-vinyl carbazole based photo-cross-link assisted cytosine deamination at the physiological conditions are hydrophilicity of the photo-cross-linker and the hydrogen bonding of the target cytosine. Thus, the best combination of counter-base and photo-cross-linker can give the highest conversion of C→U in the least time at physiological conditions is when inosine is used as counter-base and ^{OHV}K is used as photo-cross-linker.

References

10. Cascalho M, **(2004)** *The Journal of Immunology*, 172(11), 6513-6518
11. Storici F, Lewis L K, Resnick M A, **(2001)** *Nature Biotechnology,* 19(8), 773-776
12. Esvelt K M, Wang H H, **(2013)** *Molecular Systems Biology,* 9(641), 1-17
13. Tan W S, Carlson D F, Walton M W, Fahrenkrug S C, Hackett P B, **(2012)** *Adv Genet*., 80, 37-97
14. Puchta H, Fauser F, **(2013)** *Int. J. Dev. Biol.,* 57, 629-637
15. Sanger J D, Joung J K, **(2014)** *Nature Biotchnology,* 32, 347-355
16. Mali P, Yang L, Esvelt K M, Aach J, Guell M, DiCarlo J E, Norville J E, Church G M, **(2013)** *Science* 339, 823.
17. Jiang W, Bikard D, Cox D, Zhang F, Marraffini L A, **(2013)** *Nat. Biotechnol.* 31, 233
18. Javier N D, Michael N S, **(2002)** *Nature* 419, 43.
19. White M K, Kaminski R, Young W B, Roehm P C, Khalili K, **(2017)** *J. Cellular Biochemistry,* 117(11), 3586-3594
20. Cho S W, Kim S, Kim J M, Kim J S, **(2013)** *Nat. Biotech.,* 31, 230-232
21. Zhang X H, Tee L Y, Wang X G, Huang Q S, Yang S H, **(2015)** *Mol. Ther. Nucleic Acids,* 4, 264.
22. de Souza N, **(2011)** *Nat Meth*, 9(1), 27
23. Smith J, et al., **(2006)** *Nucleic Acids Res.,* 34(22), e149
24. Kim H, Kin J, **(2014)** *Nature Reviews Genetics,* 15(5), 321-334
25. Frommer M, McDonald L E, Millar D S, Collis C M, Watt F, Grigg G W, Molloy P L, Paul C L, **(1992)** *Proc. Natl. Acad. Sci. U. S. A* 89, 1827.
26. Yamana K, Yoshikawa A, Nakano H, **(1996)** *Tetrahedron Lett*, 37, 637-640.
27. Lee B L, Blake K R, Miller P S, **(1988)** *Nucleic Acdis Res*, 16, 10681-10697.
28. Montes C. V.; Memczak, H.; Gyssels, E.; Torres, T.; Madder, A.; Schneider, R. J.; **(2017)** *Langmuir*, *33* (5), 1197-1201
29. Kurz, M.; Gu, K.; Lohse, P. A.; **(2000),** *Nucleic Acids Research*, 28(18), 83.
30. Liang X, Wakuda R, Fujioka K, Asanuma, H, **(2010)** *FEBS J*, 277, 1551-1561.
31. Liu J, Geng Y, Pound E, Gyawall S, Ashton J R, Hickey J, Woolley A T, Harb J N, **(2011)** *ACS Nano*, 5, 2240-2247.
32. Iwase, R.; Namba, M.; Yamaoka, T.; Murakami, A, **(1997),** *Nucleic Acids Symposium Series,* 37, 203-204.

33. Fujimoto K, Konishi-Hiratsuka K, Sakamoto T, Yoshimura Y, **(2010)** *Chem. Commun.*, 46, 7545.
34. Fujimoto K, Konishi-Hiratsuka K, Sakamoto T, Yoshimura Y, **(2010)** *ChemBioChem.*, 20 11, 1661
35. Sakamoto T, Ooe M, Fujimoto K, **(2015)** *Bioconjugate Chem.*, 26, 1475-1478
36. Sethi S, Ooe M, Sakamoto T, Fujimoto K, **(2017)** *Molecular BioSys.,* 13(6), 1152-1156
37. Sethi S, Yasuharu T, Nakamura S, Fujimoto K, **(2017)** *Bioorg Med Chem Lett.*, 27(16), 3905-3908
38. OECD GUIDLINE FOR TESTING OF CHEMICALS 117. http://www.oecd.org/chemicalsafety/risk-assessment/1948177.pdf
39. Borges, N. M.; Kenny, P. W., **(2017)** *J. Comput. Aided. Mo. Des*. 31, 163-181
40. Sklenak, S.; Yao, L.; Cukier, R. I.; Yan, H, **(2004)** *J. Am. Chem. Soc.* 126, 14879-14889
41. Zhang X, Zhao Y, Yan H, Cao Z, **(2016)** *J. Computational Chem.*, 37(13), 1163-1174
42. Yoshimura, Y.; Fujimoto, K., **(2008)** *Org. Lett.* 10, 3227-3230.

Chapter 5: Super-fast deamination of cytosine in dsDNA using phosphate modified 3-cyanovinylcarbazole assisted photo-crosslinking

Introduction

Genes are often considered as blue prints for the designing of living body.[1] Proteins are transcribed from double-stranded DNA to single-stranded RNA, further translated from RNA into proteins with specific functionality and become part of the body.[2] Hence, if DNA with a normal design map is transcribed, protein with normal functionality will be created. However, if there is some abnormality in DNA, it gives a translated protein without original functionality which leads to genetic disorders.

Genetic diseases are diseases caused by deficiencies or mutations in the genes. Several genetic disorders usually arise from a single base mutation.[3-5] For example, Leigh's syndrome, having symptoms such as mental development delay and difficulty in swallowing, can be seen when base 8993 of mitochondrial DNA is mutated to C from T.[6-7] However, the treatments present are symptomatic like diet therapy and drugs but it does not constitute fundamental treatments.[8]

For fundamental treatments of genetic diseases, nucleic acids need to be artificially edited.

Currently, enzymatic nucleic acid editing method and chemical nucleic acid editing method have been reported in the literature.[9-10]

One of the enzymatic editing methods, CRISPR / Cas 9 is called genome editing technology of the third generation. With CRISPR technology, it is possible to cleave arbitrary genomic sequence with guide RNA containing a sequence complementary to the target nucleotide sequence of the DNA to be cleaved and cas9 act as cleaving enzyme.[11-12] However, there are reports that the base sequence outside the target which is not 100% complementary have been cleaved, leading to problem in specificity.[13]

Another class of nucleic acid editing methods is using chemical methods to induce mutations. By reacting cytosine of the DNA strand with sulfurous acid under the condition of pH = 5, hydrogen enters the 3rd position. If pH is kept below 13 in that state, ammonia escapes, and the amino group at 4th position gets replaced with hydroxyl group, and get converted to uracil.[14] However, there are major problems in biological applications such as extreme pH conditions and non-specificity.[14]

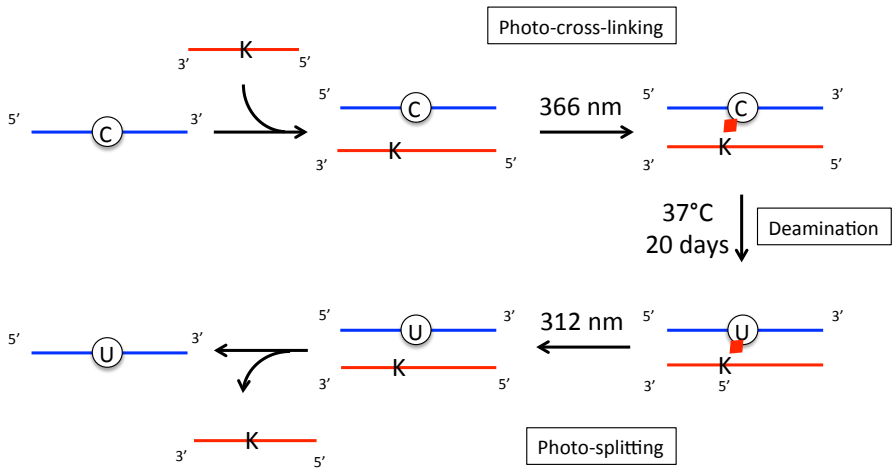
Due to inability of the current methods to cope with site-specificity, we had reported a photo-chemical method to induce sequence selective and site specific changes in the DNA and RNA in which target cytosine can be converted to uracil by photo-cross-linking the said cytosine with photo-active nucleotide, 3-cyanovinylcarbazole (^{CNV}K), using 366 nm UV radiation in only a few seconds. The crosslink assist the deamination of cytosine upon heating at 90 ºC. The photo-crosslink can be cleaved using 312 nm UV light.[15]

However, since the procedure require heating at 90 °C during deamination, its application is not realized in the cellular system.

To overcome the heating step, we reported that deamination can be promoted under the physiological condition by changing the complementary base of cytosine from guanine to inosine.[16] In addition, we also reported that the deamination reaction could be promoted by replacing the vinyl carbazole derivative substituent with a highly hydrophilic substituent.[17] Furthermore, when both these components were combined in a single ODN, it was found that upon using ODN in which inosine, as counter base to target cytosine, when coupled with ^{OHV}K, shows a drastic acceleration of deamination reaction.[18]

From these facts, we found that the base editing efficiency is promoted even at physiological conditions (37 °C) due to changes in the surrounding environment (scheme 5.1). It also suggested that complementary bases to cytosine affect reaction efficiency. The only drawback of this method was the efficiency of the reaction. The deamination reaction was slow and could only give 80% yield even after 20 days of incubation.[18]

Therefore, in this study, we will comprehensively examine the micro-environment around cytosine by using phosphate modified 3-cyanovinylcarbazole and counter base of cytosine to aim at further accelerating deamination reaction at physiological conditions.



Scheme 5.1: Schematic drawing of photochemical cytosine to uracil transition.

Materials and Methods

Preparation of oligonucleotides

The phosphoramidite of ^{CNV}K was prepared according to the literature method.[ref] The modified oligonucleotides containing ^{CNV}K (ODN(K); 5'-^{CNV}KGTACTA-3' and the phosphate group-modified oligonucleotides containing ^{CNV}K, p^{CNV}K or ps^{CNV}K (ODN(Kp); 5'-p^{CNV}KGTACTA-3'; ODN(psK); 5'-ps^{CNV}KGTACTA-3') and the third complementary oligonucleotides (ODN(X); 5'-CGTATGCATX-3' (X=I, C, G, A, T, Ip, Cp, Gp, Ap, T)) and target oligonucleotide (ODN(tC); 5'-TAGTACGCATGCATACG-3') were prepared according to standard β-cyanoethyl phosphoramidite chemistry on DNA synthesizer using the phosphoramidite of ^{CNV}K. All oligonucleotides were purified by reversed-phase HPLC and identificationed by MALDI-TOF-MS.

Table 5.1: MALDI-TOF-MS analysis of photo-active ODNs

ODN(X)	**Sequence**	**Calculated Mass (M+H)**	**Observed value (M+H)**
ODN(K)	5'-^{CNV}KGTACTA-3'	2188.56	2192.92
ODN(pK)	5'-p^{CNV}KGTACTA-3'	2272.50	2272.52
ODN(psK)	5'-ps^{CNV}KGTACTA-3'	2288.48	2290.12

Photoirradiation

Photo-cross-linking reaction of ^{CNV}K in oligonucleotide duplex (5 μM in 50 mM Na-Cacodylate buffer (pH 7.4) containing 100 mM NaCl) was performed at 4ºC using UV-LED (366 nm, 9500 mW/cm^2, OmniCure® LX405S) and photo-splitting was performed at 37°C using transilluminator (312 nm, 2.0 mW/cm^2, Funakoshi).

Deamination reaction

The solution of ODN(K) and ODN(C) in buffer solution was annealed and photoirradiation at 366 nm. The photoirradated solution was purified by a reversible HPLC to get the purified photo-cross-linked dsDNA. Photo-cross-linked DNAs (5 μM in 50 mM Na-cacodylate buffer (pH7.4) containing 100 mM NaCl) were incubated at 37ºC.

UPLC analysis

The photoirradiated samples were analyzed with analyzed by ultra-high performance liquid chromatography (UPLC). UPLC system (Aquity, Waters) equipped with BEH Shield RP18 column (1.7 μm, 2.1 × 50 mm, elution was with 0.05 M ammonium formate containing 1-13% CH_3CN, linear gradient (10 min) at a flow rate of 0.4 mL/min, 60ºC). The deamination sample were analyzed with above equipment. (Elution was with 0.05M ammoniumformate containing 1-13%CH_3CN).

HPLC analysis for RNA samples

The deaminated (and photo-split) RNA samples were analyzed by a reversed phase HPLC ((PU-980, HG-980-50, UV-970) system equipped with Cosmosil 5C18-AR-II column (4.6ID x 150 mm) with a gradient of 2-20% MeCN and flow rate of 1 ml/min with column temperature of 60ºC).

Partition coefficient (LogP)

LogP of photo-cross-linked DNA was measurement by their retention time following OECD protocol. 10 μM photo-cross-linked dsDNA was analyzed with HPLC system; elution was with 0.05 M ammonium formate containing 98-50% CH_3CN, linear gradient (60 min) at flow rate of 1 mL/min. 4-Acetylpyridine, Aniline, Acetanilide, Phenol, Benzonitrile, and Acetophenone were used as control compound to create calibration curve.[2-3]

CD Spectra

Circular dichroism (CD) spectra were measured using a Jasco model J-820 CD spectrophotometer. The spectra were recorded using a 0.1 cm path length cell. Samples were prepared according to each step of the DNA editing reaction. Solutions for CD spectra were prepared as 0.3 ml samples at a 10 μM strand concentration in the presence of 50 mM sodium cacodylate buffer containing 100 mM NaCl (pH 7.4).

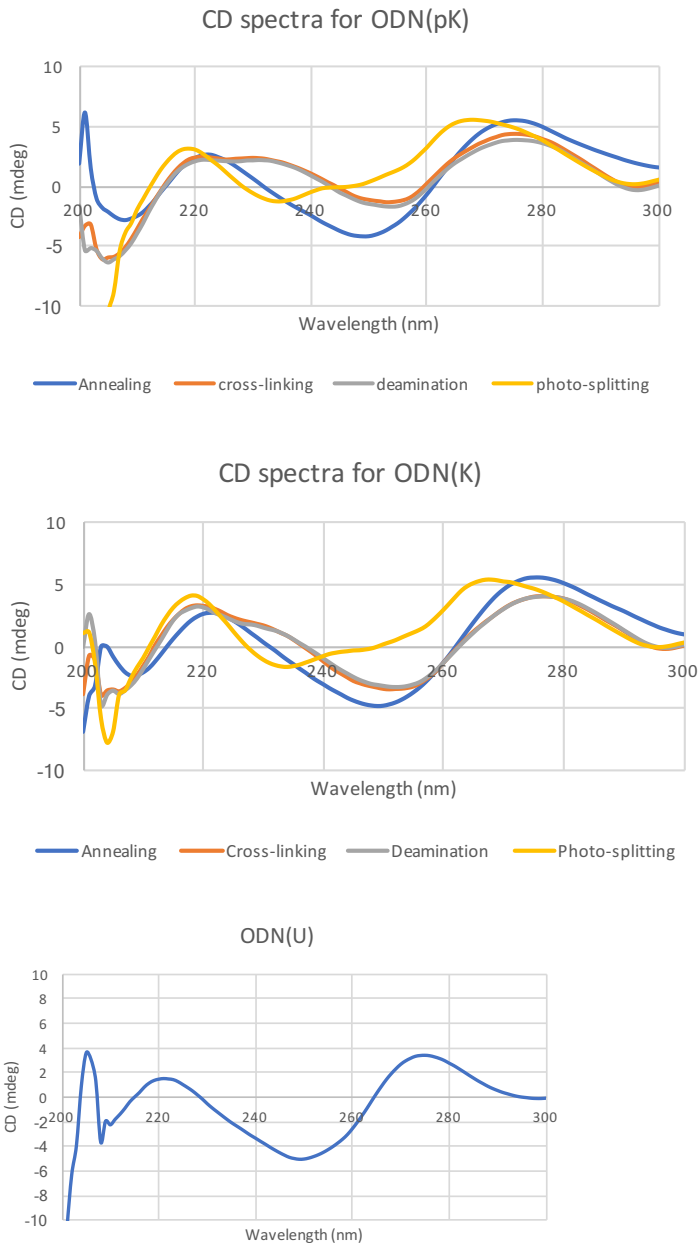
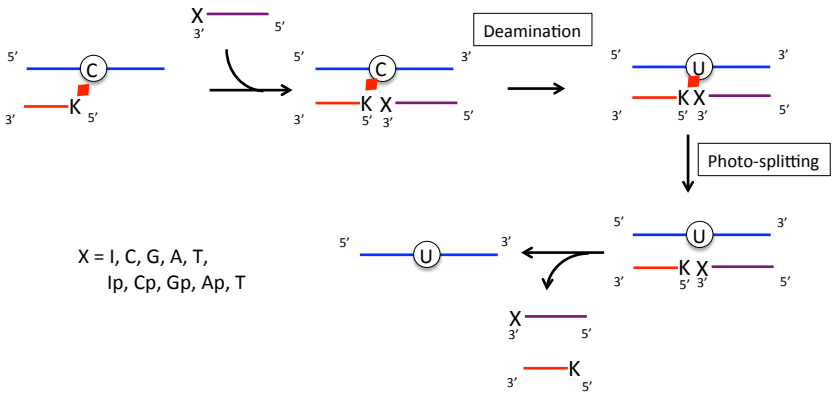


Figure 5.1: Circular Dichorims (CD) spectra for each stage of photo-chemical DNA editing using a) ODN(pK) and ODN(tC); b) ODN(K) and ODN(tC). C) CD spectra of ODN(U).

Results and Discussion

In order to approach the acceleration of deamination, based on the previous reports, the hydrophilicity around the target cytosine has been increased by using two disjoint ODNs complementary to each half of the target ODN. ODN(K) having photo-active 3-cyanovinylcarbazole (^{CNV}K) at the 5' end of the ODN. Second ODN, ODN(X) having complementary base to cytosine at 3' end with a free phosphate group at the 3' end of the complementary base (scheme 5.2 and figure 5.2). The complementary bases used are guanosine, inosine, adenine, thymine, and cytosine. The sequence of the ODNs have been shown in the table 5.2.



Scheme 5.2: Schematic representation of the photo-reaction for deamination of cytosine with various counter bases to study the effect of micro-environment around the target cytosine.

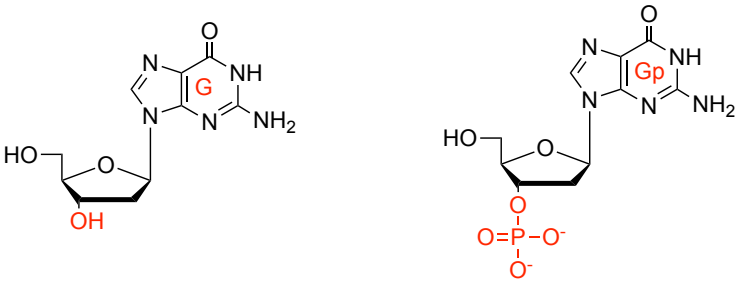


Figure 5.2: Structure of modified counter base.

Table 5.2: Sequence of ODNs used

ODN(X)	Sequence
temp(C)	5'-TAGTACGCATGCATACG-3'
temp(U)	5'-TAGTACGUATGCATACG-3'
ODN(K)	5'-^{CNV}KGTACTA-3'
ODN(pK)	5'-p^{CNV}KGTACTA-3'
ODN(psK)	5'-ps^{CNV}KGTACTA-3'
ODN(G)	5'-CGTATGCATG-3'
ODN(C)	5'-CGTATGCATC-3'
ODN(T)	5'-CGTATGCATT-3'
ODN(A)	5'-CGTATGCATA-3'
ODN(I)	5'-CGTATGCATI-3'
ODN(Gp)	5'-CGTATGCATGp-3'
ODN(Cp)	5'-CGTATGCATCp-3'
ODN(Tp)	5'-CGTATGCATTp-3'
ODN(Ap)	5'-CGTATGCATAp-3'
ODN(Ip)	5'-CGTATGCATIp-3'

Firstly, deamination reaction was carried out using photo-adduct of ODN(K) and temp(C) with ODN(G) and ODN(Gp) as complementary strand of temp(C) ((5 μM each in 50 mM Na-cacodylate buffer (pH7.4) containing 100 mM NaCl) was incubated for 0 h～24 h at 37 ºC). The deaminated sample was then photo-split using 312 nm radiation (2.0 mW/cm^2, 37 ºC) and analyzed by UPLC. The results have been shown in figure 5.3. In case where G is non-modified, i.e ODN(G) has shown around 2-6% conversion of cytosine to uracil in 24h while when modified ODN(Gp) has been used, the conversion is 77% in 24h.

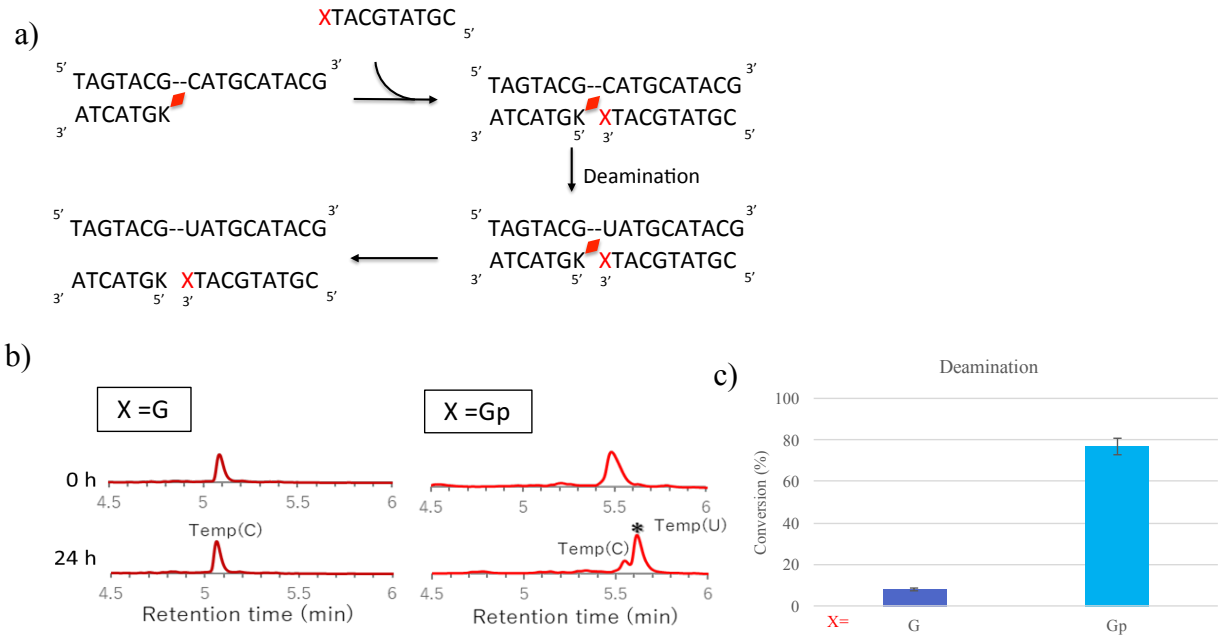
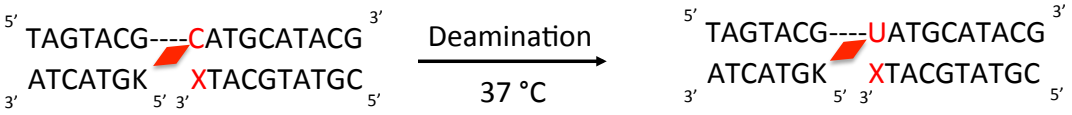


Figure 5.3: a) Scheme of deamination with sequence. b) UPLC chromatograms for deamination reaction with and without phosphate group. Sample: photo-cross-linked duplexes (2.5 µM) and ODN(G) or ODN(Gp) ((5 µM) in 50 mM sodium cacodylate buffer containing 100 mM NaCl were incubated at 37ºC and photo-split (312 nm, 2.0 mW/cm^2, 37ºC)). c) Extent of cytosine to uracil conversion after 24h incubation at 37 ºC.

Furthermore, deamination reaction was then carried out by changing counter base of target C. A mixture of ODN(tC) (ODN (C) and ODN (K) irradiated with 366 nm (9500 mW/cm^2, 4 ºC) and purified using HPLC) and ODN(X) (X is I, C, G, A, T, Ip, Cp, Gp, Ap, Tp) (5 µM each in 50 mM Na-cacodylate buffer (pH7.4) containing 100 mM NaCl) was incubated for 0 h～24 h at 37 ºC. To evaluate the effect of deamination with the counter base of target C, the sample was then irradiated at 312 nm (2.0 mW/cm^2, 37 ºC) and analyzed by UPLC. The results are shown in Figure 5.4. In the Unmodified bases (X=I, C, G, A, T), upon incubation for 24 h, scarce induction from Cytosine to Uracil was observed. However, when the terminal of the counter base was modified with phosphate group (X=Ip, Cp, Gp, Ap, Tp), the conversion of cytosine to uracil was easily induced by incubation at 24 h (Figure 5.5). This hinted that the presence of phosphate group at the terminal of the counter base affected the rate of deamination reaction to a great extent. Moreover, these results also indicate that the deamination reaction is faster when inosine is used as counter base of target cytosine, which corroborate with our previous studies.

a)



b)

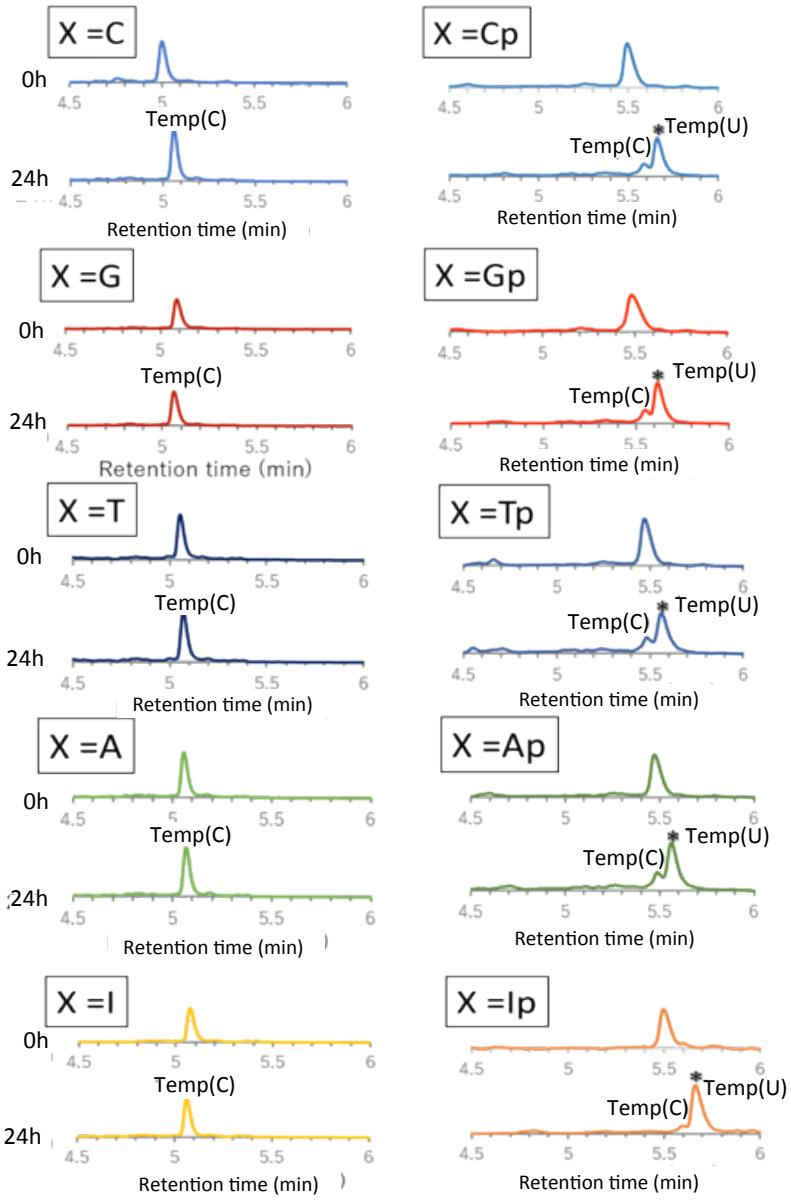


Figure 5.4: a) Scheme of deamination reaction with scheme. b) UPLC chromatograms for deamination reaction with and without phosphate group. Sample: photo-cross-linked duplexes (2.5 µM) and ODN(X) ((5 µM) in 50 mM sodium cacodylate buffer containing 100 mM NaCl were incubated at 37ºC (312 nm, 121.3 mW/cm^2, 37ºC)).

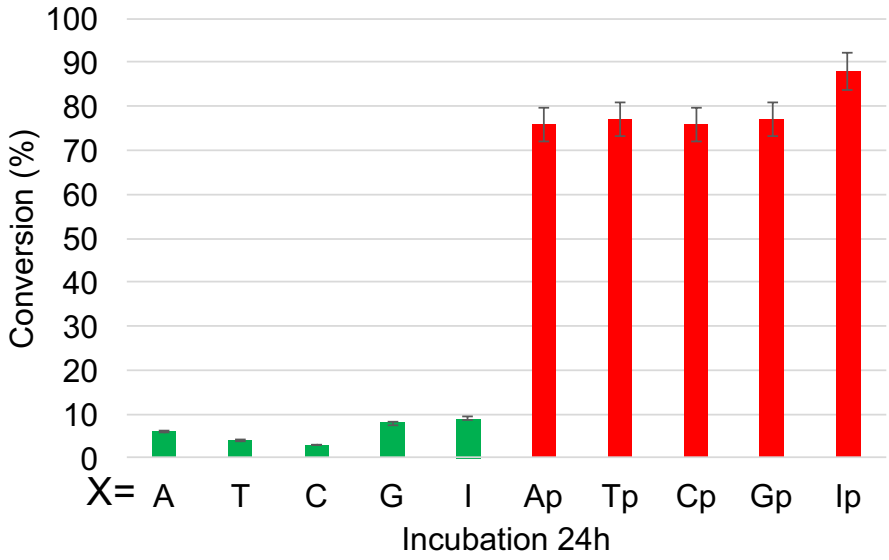


Figure 5.5: Extent of cytosine to uracil conversion after 24h incubation at 37 ºC.

Then, we tried to find whether the presence of phosphate group on cross-linker has any direct impact on the rate of deamination reaction. Therefore, we prepared ODN(K), ODN(pK), where ODN(K) doesn't possess phosphate group at 5' terminal whereas, ODN(pK) does.

Photo-crosslinked ODNs having ODN(C) and ODN(K or pK) were prepared by annealing equimolar mixture of the ODNs in 50 mM sodium cacodylate buffer having 100 mM sodium chloride and then irradiating using 366 nm UV radiation. The photo-crosslinked ODN(Y) (Y= K, pK) was then purified using HPLC and incubated at 37 ºC for 6h after addition of ODN(X), where X=G or Gp. Upon which, the ODNs were cleaved using 312 nm UV radiation and analysed by UPLC.

From the UPLC chromatograms (Figure 5.6), it has been observed that, upon using ODN (K) and ODN(G), no deamination has been observed in 3h, while a mere 2% conversion has been observed in 6h. When phosphate group was attached either on the photo-cross-linker or counter base, almost comparable deamination of cytosine to uracil has been observed, 52% with ODN(Kp) and 47% with ODN(Gp). On the other hand, when both counter base and photo-cross-linker are phosphate modified, the conversion of cytosine to uracil is ~90% in just 6h (figure 5.7). This indicates that the having more phosphate groups around the target cytosine accelerate the deamination reaction.

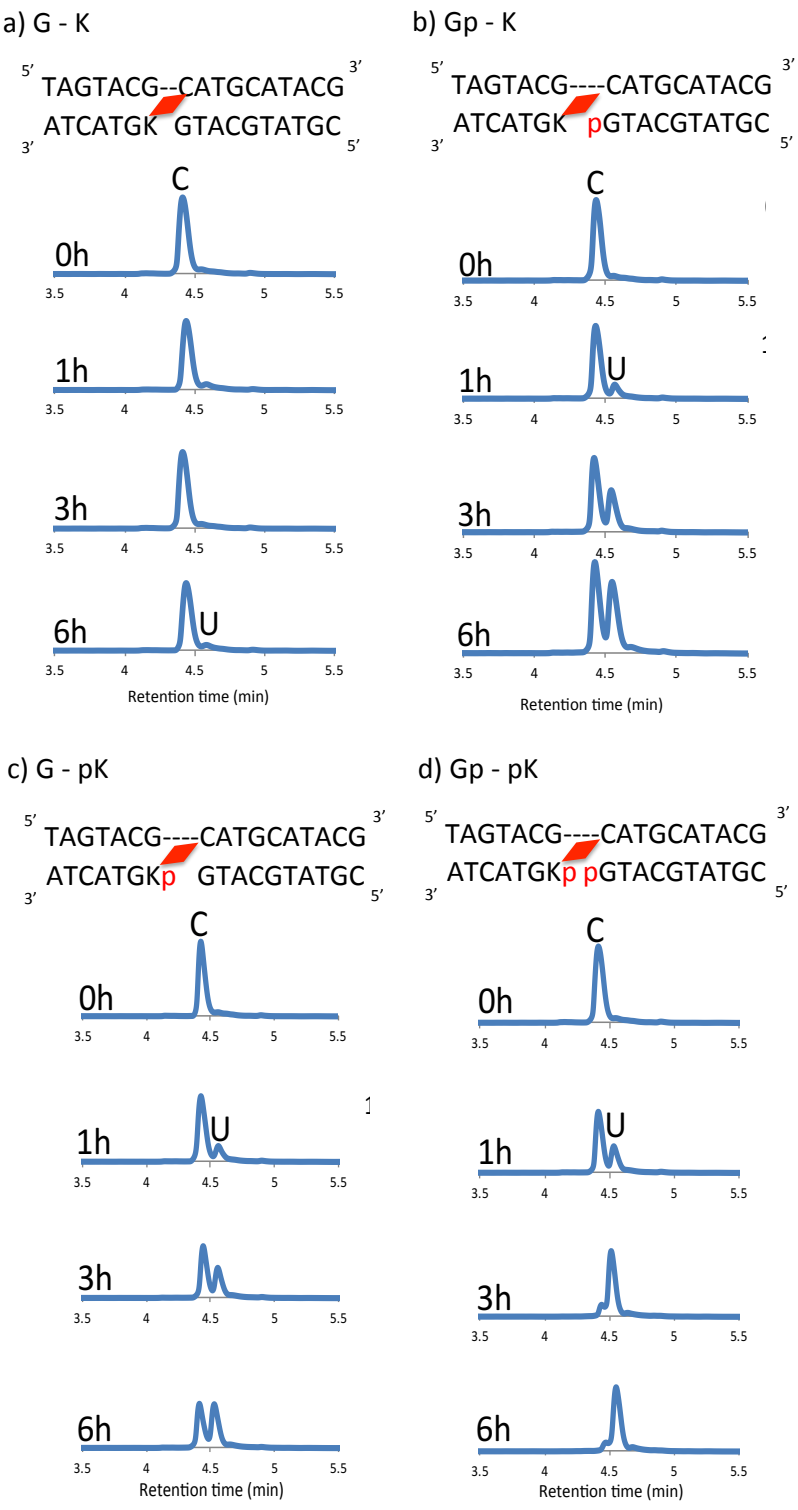


Figure 5.6: UPLC chromatograms for deamination reaction with a) no phosphate group, b) phosphate group only on counter base, c) phosphate group on only ^{CNV}K, d) phosphate group on both counter base and ^{CNV}K. Sample: photo-cross-linked duplexes (2.5 µM) and ODN(X) (5 µM) in 50 mM sodium cacodylate buffer containing 100 mM NaCl were incubated at 37ºC (312 nm, 121.3 mW/cm2, 37ºC)

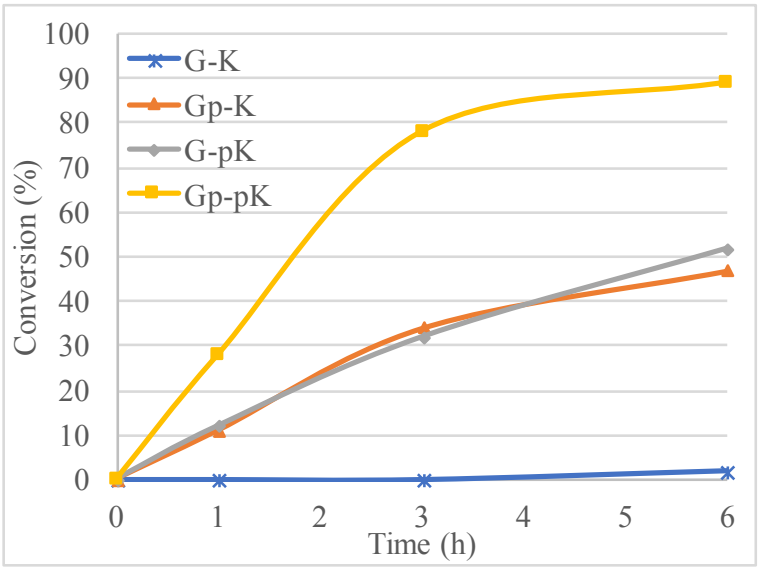
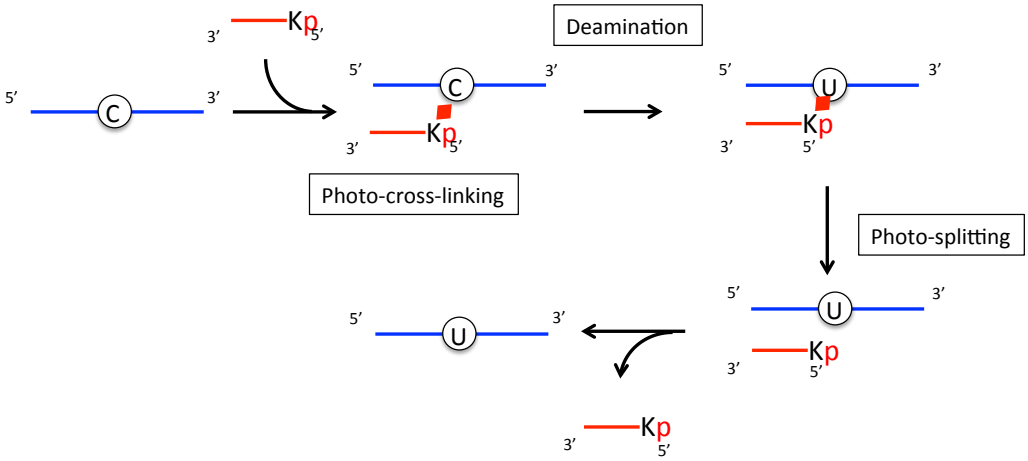


Figure 5.7: Time course of photochemical cytosine to uracil transition with free phosphate group on cross-linker or guanine.

Further, those results indicated that the free phosphate group accelerates the deamination of cytosine so we designed the simple system using template and ODN (pK). We demonstrated the deamination of cytosine using the system shown in scheme 5.3.



Scheme 5.3: Schematic illustration of photochemical cytosine to uracil transition using photo-active ODN modified with free phosphate group at 5'-end study the effect of free phosphate group without counter base.

The reaction compared deamination using previous method, having a full counter sequence with target having ^{CNV}K namely, ODN(K17), with the new system having a 7-mer ODN shortened till -1 position to that of cytosine, ODN(K) and 7-mer ODN with phosphate modification at 5 end, ODN(pK) (Fig 5.8). From the UPLC analysis, it was observed that while in 24 h, no deamination was observed using the ODN(K17) but only a little amount of cytosine was converted to uracil using ODN(K). Interestingly, close to 100% conversion was observed with OND(pK) which confirms that even without any counter base, separate or not, deamination can be achieved to completion using phosphate modification.

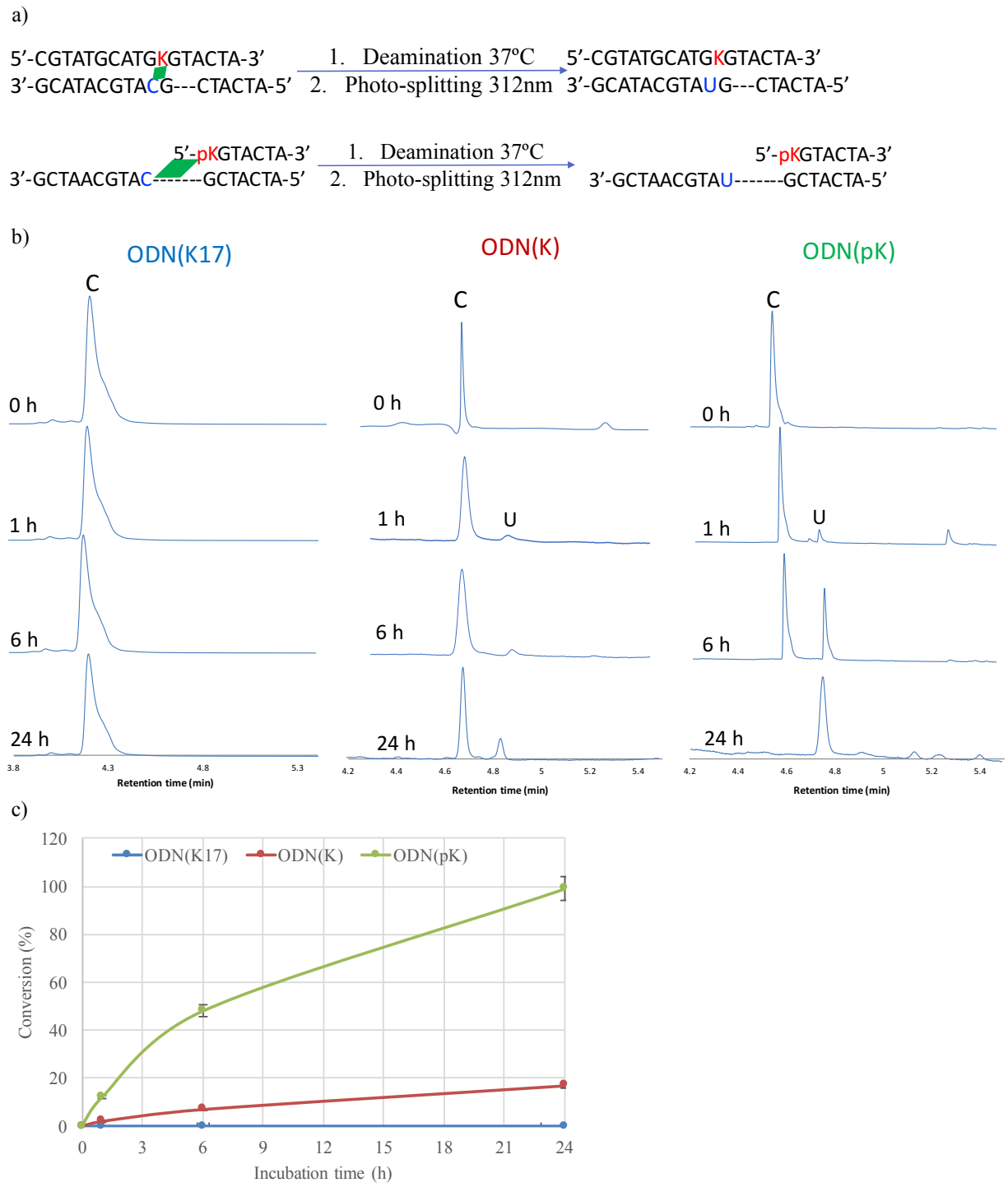


Figure 5.8: a) Scheme of deamination. b) UPLC chromatograms for deamination of ODN(tC) using ODN(K) and ODN(pK). Sample: photo-cross-linked duplexes (2.5 µM) in 50 mM sodium cacodylate buffer containing 100 mM NaCl were incubated at 37ºC (photo-splitting: 312 nm, 121.3 mW/cm2, 37ºC). c) Time course of deamination reaction.

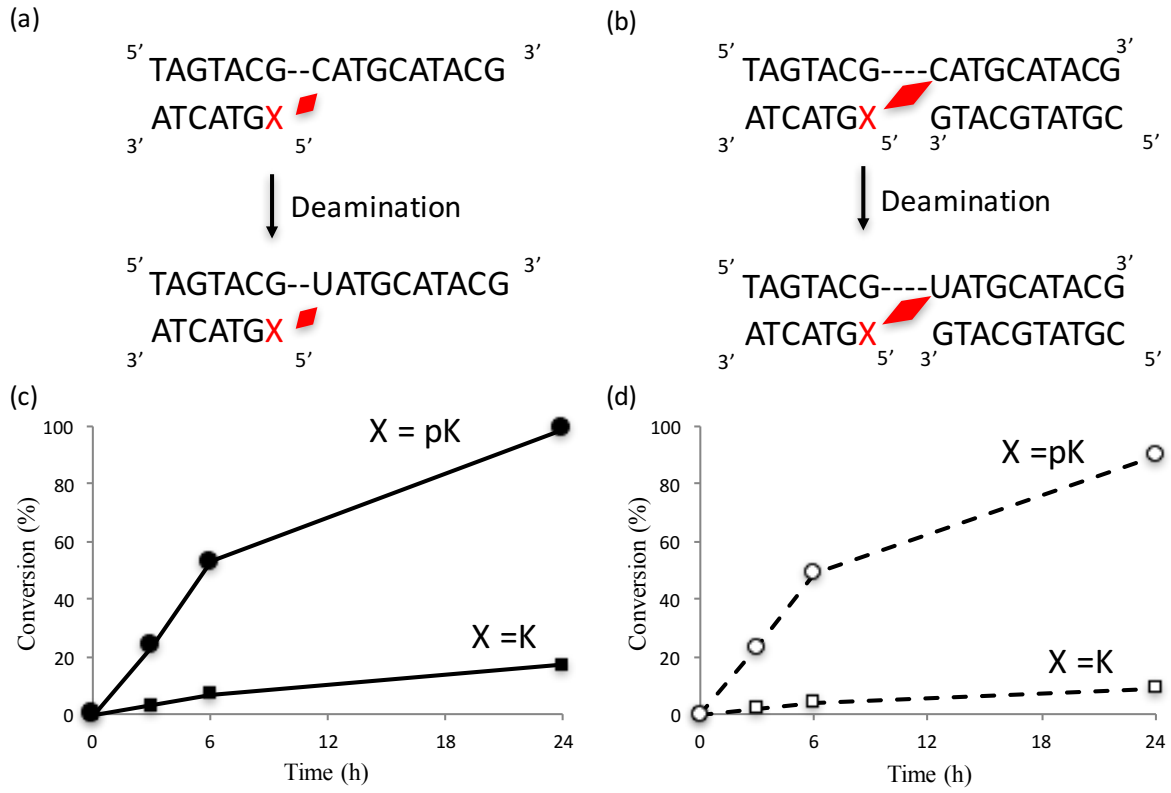


Figure 5.9: Time course of cytosine to uracil conversion in 24 at 37°C. (a) Novel system using temp(C) and ODN(pK) and (b) time curse of deamination. (c) Previous system using temp(C), ODN(pK) and ODN(G) and (d) time course of deamination.

Further comparing 2-ODN system with 3-ODN system(Figure 5.9), it has been observed that the X = pK has same conversion of cytosine to uracil regardless of presence of counter base containing ODN. This indicated that this novel system using ODN(tempC) and ODN(pK) accelerate deamination of cytosine even without counter ODN.

Furthermore, we tried to change the polarity on the phosphate group by using another photo-cross-linker, phosphorothioate cyanovinylcarbazole (psK or ps^{CNV}K) and carried out the deamination reaction for 24 h.

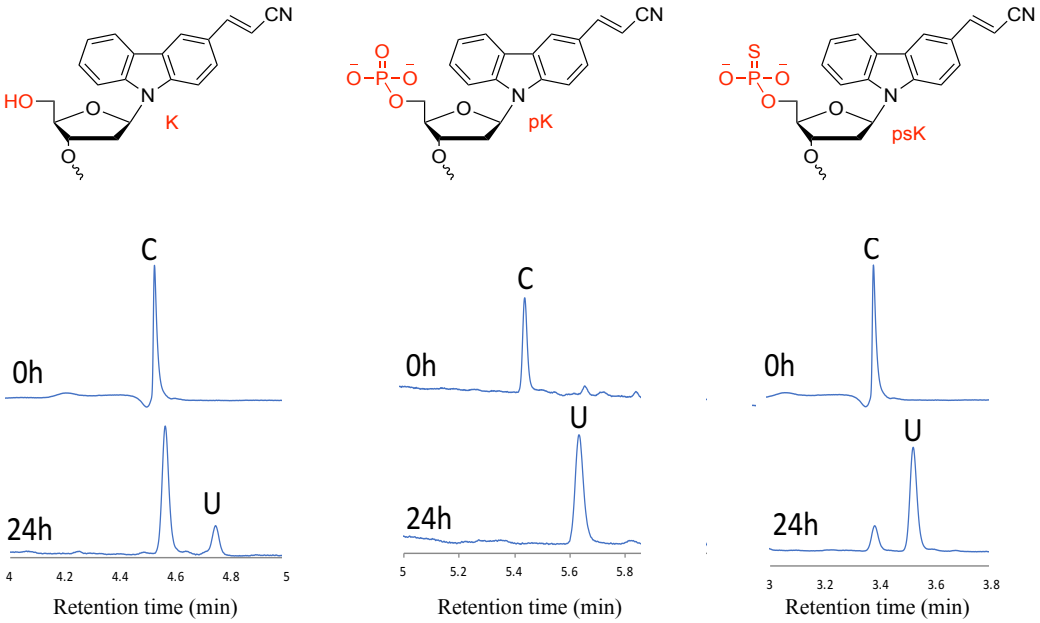


Figure 5.10: UPLC chromatograms for deamination reaction with different modifications of photo-cross-linker. a) ODN(K<>C), b) ODN(pK<>C), c) ODN(psK<>C) Sample: photo-cross-linked duplexes (2.5 µM) in 50 mM sodium cacodylate buffer containing 100 mM NaCl were incubated at 37ºC (312 nm, 121.3 mW/cm2, 37ºC).

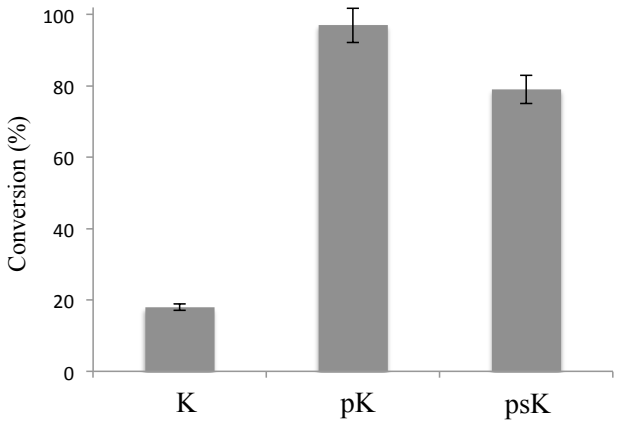


Figure 5.11: Extent of cytosine to uracil conversion with ODN(K), ODN(pK), and ODN(psK)

Using various modification of photo-cross-linkers, like ^{CNV}K, p^{CNV}K, and ps^{CNV}K, deamination of cytosine was carried out (figure 5.10). It was observed that using pK, almost 100% conversion has been observed in 24h when no complementary strand has been used while with phosphorothioate modified ^{CNV}K, the conversion is around 80% (figure 5.11). Whereas, when unmodified ^{CNV}K has been used as photo-cross-linker, <20% conversion has been observed in 24 h.

To explain the difference in the rate of conversion, logP of all the photo-adducts was calculated. Table 5.3 shows the logP values and conversion ratio of using each photo-active ODN. It Is interesting here to note that the LogP values of the photo-cross-linking ODNs are in direct relation with the conversion ratio. Therefore, higher hydrophilicity of the ODN (pK) related to the higher rate of deamination reaction and likewise, the lower hydrophilicity in the ODN (K) pertains to the least amount of uracil formation. Therefore, presence of phosphate (or phosphorothioate) leading to higher polarity and hydrophilicity affects the incoming nucleophile by activating them for the nucleophilic attack on cytosine.

Table 5.3: LogP value of ODNs.

ODN(X)	Sequence	logP (Difference*)	Conversion (%)	pKa[19]
ODN(pK)	5'-p^{CNV}KGTACTA-3'	0.98 ±0.12 (0.24)	95 ±4.96	6.5
ODN(psK)	5'-ps^{CNV}KGTACTA-3'	1.02 ±0.10 (0.2)	85 ±4.42	5.0
ODN(K)	5'-^{CNV}KGTACTA-3'	1.22 ±0.09 (0)	18 ±1.32	-

*logP – logP of ODN(K)

Further, we tried to find the effect of different metal ions on the deamination of cytosine to check our hypothesis of metal assisted nucleophilic attack on the cytosine. For this purpose, we used sodium, magnesium, zinc, and potassium (figure 5.12).

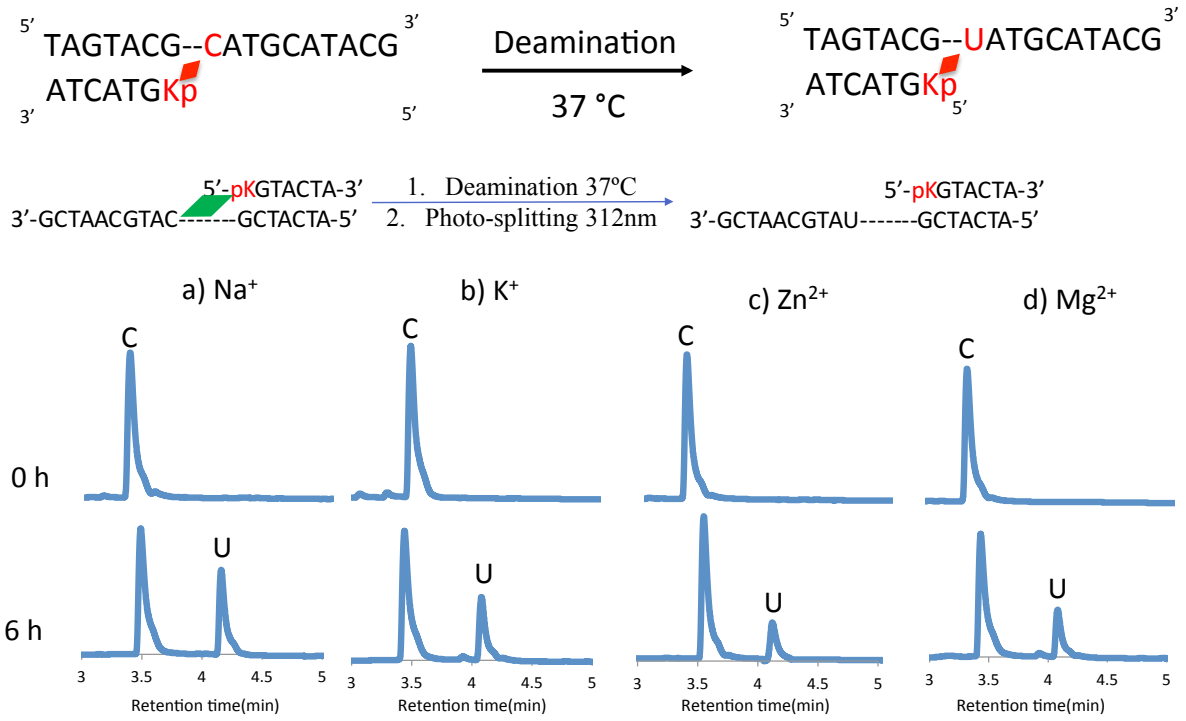


Figure 5.12: UPLC chromatograms for deamination reaction with different metals salts in Tris-HCl buffer. Sample: photo-cross-linked duplexes (2.5 µM) in 50 mM Tris-HCl buffer containing salts of different metals were incubated at 37ºC (312 nm, 121.3 mW/cm2, 37ºC). NaCl: 100 mM, KCl: 100 mM, $ZnCl_2$: 20 mM, $MgCl_2$: 20 mM.

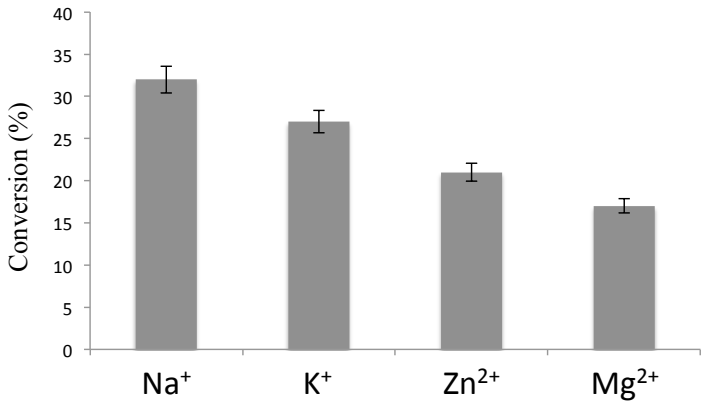


Figure 5.13: Extent of cytosine to uracil conversion in 6h.

It was found that the maximum conversion has been observed when sodium was used as the buffering salt and could give highest rate of deamination reaction (figure 5.13). It is interesting here to observe that the mono-basic salts are better for deamination reaction (K and Na) than the di-basic salts (Zn and Mg), which might be due to formation of ionic bonds between the metal and oxygen ion of the phosphate group. In case of monobasic salt, one of the charged oxygens of phosphate forms the ionic bond with metal cation which leaves one oxygen free to

form hydrogen bond with the incoming nucleophile and the cation can further stabilize and activate the nucleophile by interacting with the oxygen of nucleophilic water. This helps in increasing electron density on the nucleophile which further facilitates the nucleophilic attack on cytosine. It is also interesting to note here that among the mono basic cations; sodium shows more affinity towards the acceleration of nucleophilic attack as it is more electropositive in nature than potassium thus further strengthening the nucleophlicity of the water molecule. Whereas, in case of dibasic cations, both the oxygens of phosphate forms ionic bond with the metal ion and forms a structure similar to cyclic monophosphate which deprive the hydrogen bond to form with the incoming nucleophile and thus slows down the reaction.

From all the deamination results and LogP analysis, it can be concluded that the phosphate group on either the cross-linker or the counter base can accelerate the deamination of photo-cross-linked cytosine and the more the number of the phosphate groups, more the acceleration. Moreover, sodium is better for the deamination reaction than other metal ions due to its efficiency towards co-ordination to oxygen. Further, the protonation of phosphate plays a role as the metal can only bind to the oxygen of phosphate when there is no hydrogen to activate the attack of nucleophile. Also, phosphorothioates are equally effective in accelerating the deamination reaction which gives this technique hope in the *in vivo* applications as the phosphorothioate modified ODNs are less prone to degradation in the cytosol. And above all, the accelerating effect of phosphate group (either 1 or 2) can be attributed to the hydrophilicity of free phosphate based on the LogP values.

Thus, based on the role of phosphate group at 5’ end of the photo-active ODN, we hypothesize a mechanism based on the mechanism of phosphate based pseudo hairpin loop formation and transposition carried out by DNA transposome. In which case, a free phosphate group at 5’ end forms a pseudo hairpin loop with the complementary strand’s 3’-OH group through hydrogen bonding, which, in turn activated the OH for attack on the circular duplex DNA leading to transposition of the DNA strand. Analogous to this mechanism, in our study, the phosphate group at 5’ end of the photo-cross-linker is in close proximity to the target amino group of cross-linked cytosine (figure 5.14). This phosphate group accommodates water molecules due to high hydrophilicity and increases the electron density on the oxygen of water, thereby making it better nucleophile and due to fixed spatial position due to hydrogen bonding, the oxygen is in perfect position to attack the C4 carbon of cytosine leading to formation of a

pseudo S_N2 reaction intermediate which accelerates the elimination of ammonia molecule from cytosine followed by rearrangements of electron thus forming uracil.

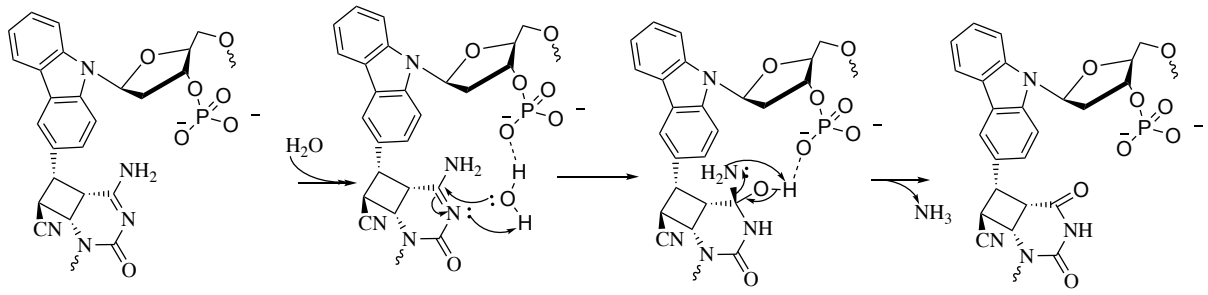


Figure 5.14: Proposed mechanism of deamination of photo-cross-linked cytosine through the phosphate based activation.

Furthermore, to assess the capability of this scheme in the RNA, we designed ORN(tC), the target RNA having the similar sequence to the target ODNs used in the previous experiments and the deamination reaction was carried out using the said RNA photo-cross-linked with ODN(K) or ODN(pK), (5 µM cross-linked ODN in 50 mM sodium cacodylate buffer with 100 mM NaCl) was incubated at 37ºC for 24 h followed by photo-splitting using 312 nm UV irradiation at 37ºC. HPLC analysis was then carried out to ascertain the conversion of cytosine to uracil. Figure 5.15 shows the sequence and scheme of the deamination reaction along with HPLC chromatograms and the extent of conversion from cytosine to uracil in case of both ODN(K) and ODN(pK). It has been observed that the RNA can be easily targeted using the photo-chemical method for site-directed mutagenesis and the conversion of cytosine to uracil is similar to that when DNA was used as target, i.e 10% with ODN(K) and >90% using ODN(pK). RNA as a target is easier to handle for the photo-chemical site-directed cytosine to uracil conversion as RNA is single stranded and photo-active ODN can act in the cytosol itself.

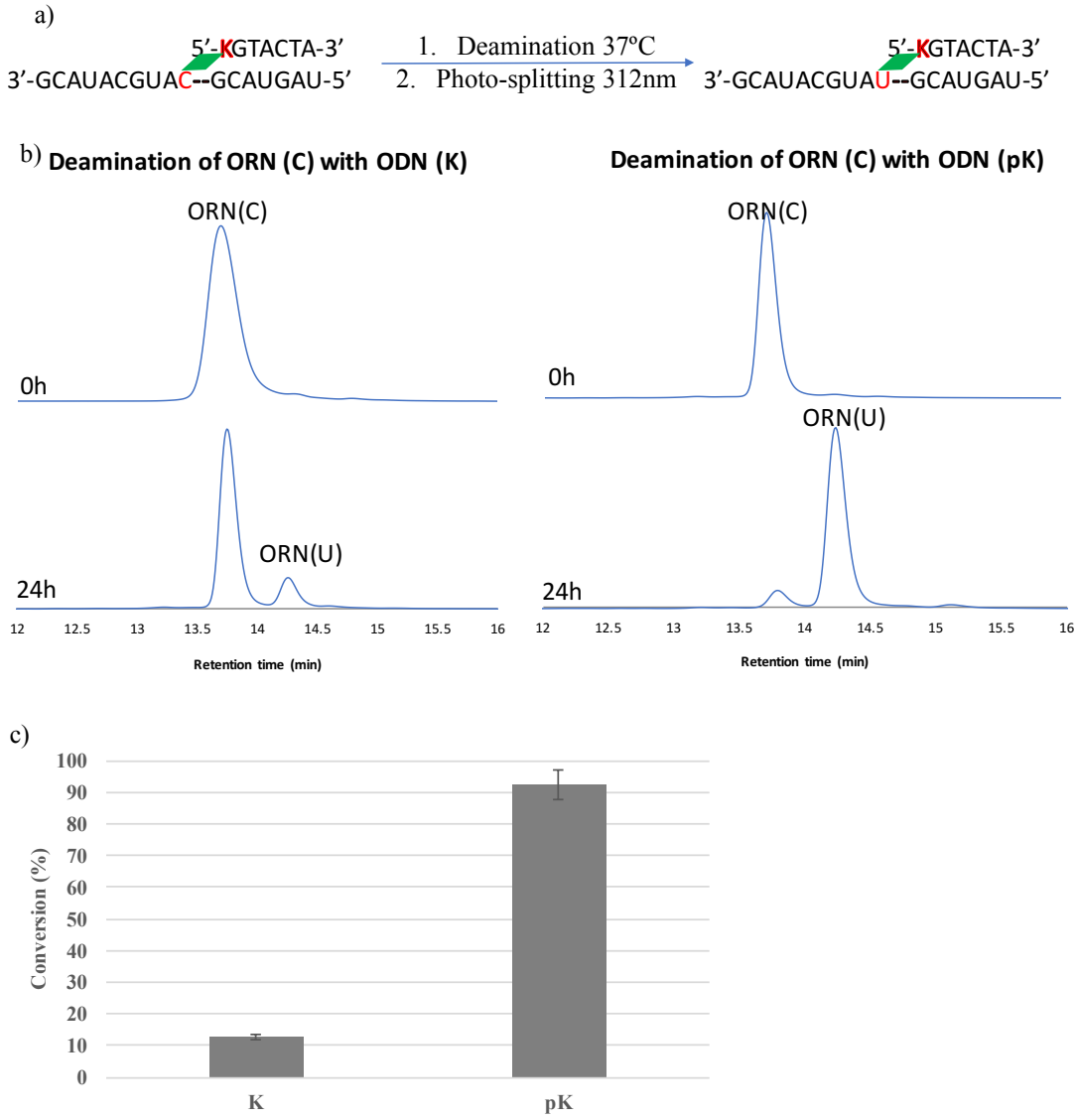


Figure 5.15: a) Scheme of deamination reaction using RNA. b) UPLC chromatograms for deamination of ORN(tC) using ODN(K), ODN(pK), and ODN(psK). Sample: photo-cross-linked duplexes (2.5 µM) in 50 mM sodium cacodylate buffer containing 100 mM NaCl were incubated at 37ºC (312 nm, 121.3 mW/cm^2, 37ºC). c) Extent of cytosine to uracil conversion.

Conclusions

In the conversion of cytosine to uracil in DNA using reversible DNA photo-cross-linking reaction, we found that the deamination reaction rate was influenced by the counter base of target cytosine, where inosine has maximum effect on acceleration of the deamination reaction. Site-specific nucleoside editing from Cytosine to Uracil in DNA by introducing the phosphate group can be accelerated to a very high rate. This technique, after being accelerated under physiological conditions, has broad potential for *in vivo* photochemical nucleic acid editing.

References

1. Dusheck J, **(2002)** *Natural History*, 111, 52-59
2. Crick F, **(1970)** *Nature,* 1970, 227(5258), 561-563
3. Ikebe S, Tanaka M, Ozawa T, **(1995)** *Mol. Brain Res.,* 28(2), 281-295
4. Sakuraba H, Oshima A, Fukuhara Y et al., **(1990)** *Am. J. Hum. Genet.,* 47(5), 784-789
5. Li J, Uversky V N, Fink A L, **(2001)** *Biochemistry*, 40(38), 11604-11613
6. deVries D D, vanEngelen B G M, Gabreels F J M, Ruitenbeek W, Oost B A V, **(1993)** *Ann. Neurol.,*, 34, 410-412
7. Lake N J, Compton A G, Rahman S, Thorburn D R, **(2016)** *Ann. Neurol.,* 79, 190-203
8. Baertling F, Rodenburg R J, Schaper J et al., **(2014)** *J. Neurol. Neurosurg. Psychiatry,* 85, 257-265
9. Lamont P J, Surtees R, Woodward C E, Leonard J V, Wood N V, Harding A E, **(1998)** *Arch. Dis. Child,* 79, 22.
10. Watanabe M, Nobuta A, Tanaka J, Asaka M, **(1996)** *Int. J. Cancer* 67, 264.
11. Sanger J D, Joung J K, **(2014)** *Nature Biotchnology,* 32, 347-355
12. Baker M, **(2012)** *Nat Meth,* 9(1), 23-26
13. Fu Y, Foden J A, Khayter C, Maeder M L, Reyon D, Joung J K, Sander J D, **(2013)** *Nat. Biotechnol.*, 31, 822-826
14. Tanaka K, Okamoto A, **(2007)** *Bioorg. Med. Chem. Lett.,* 17, 1912-1915
15. (a) Fujimoto K, Hiratsuka-konishi K, Sakamoto T, Yoshimura Y, **(2010)** *ChemBioChem,* 11, 1661-1664; (b) Fujimoto K, Hiratsuka-konishi K, Sakamoto T, Yoshimura Y, **(2010)** *Chem. Commun.*, 2010, 46, 7545-7547
16. Sethi S, Ooe M, Sakamoto T, Fujimoto K, **(2017)** *Molecular BioSys.* 13(6), 1152-1156
17. Sethi S, Takashima Y, Nakamura S, Fujimoto K, **(2017)** *Bioorg Med Chem Lett.*, 27(16), 3905-3908
18. Sethi S, Nakamura S, Fujimoto K, **(2018)** *Molecules,* 23(4), 828
19. Klenchin V A, Czyz A, Goryshin I Y, Gradman R, Lovell S, Rayment I, Reznikoff W S, **(2008)** *Nucleic Acids Research*, 36 (18) 5855–5862.

Chapter 6: General Conclusion

Chapter 2: The focus was to find the best counter-base of cytosine, based on hydrogen bonding, for photo-chemical site-directed mutagenesis using 3-cyanovinylcarbazole as the photo-active nucleoside which can crosslink with the target cytosine to afford cytosine to uracil transformation at physiological conditions. Different counter bases like guanine (G), inosine (I), 2-aminopurine (P), nebularine (R), and 5-nitroindole (N) were used to find the best counter base. Among all the bases, it was found that P, R, and N are not suitable counter base for photo-chemical cytosine to uracil transformation as when using these bases, no deamination reaction takes place. While in case of G, the deamination reaction is very slow and only 5% conversion is observed in 72h reaction time. Thus, the best counter base among the bases was found to be inosine which gives 35-40% in 72h reaction time at 37ºC having the optimal hydrogen bonding pattern before and after photo-cross-linking.

Chapter 3: In this chapter, the role of hydrophilicity and polarity of photo-cross-linker was discussed. Various derivatives of vinyl carbazole, like 3-cyanovinyl carbazole (^{CNV}K), 3-amidovinylcarbazole (^{NH2V}K), 3-methoxyvinylcarbazole (^{OMeV}K), and 3-carboxylvinylcarbazole (^{OHV}K), were used for studying the micro-environment around the target cytosine crosslinked to photo-cross-linker during the deamination of cytosine. It was discovered that the hydrophilicity and polarity of the photo-cross-linker plays a crucial role in the deamination of cytosine to uracil via photo-cross-linking. ^{OMeV}K having the least hydrophilicity gave the least rate of reaction for the deamination reaction at varying temperature (90, 70, 50, and 37 ºC) while the highest reaction rate was observed with ^{OHV}K, which is most polar among the cross-linkers based on the polarity index (Log P). Thus it was concluded that the hydrophilicity and polarity around target cytosine are deciding factors in case of deamination reaction of cytosine via photo-cross-linking.

Chapter 4: Based on the findings of chapter 2 and 3, the overall micro-environment around the target cytosine for the mutation of cytosine to uracil via vinylcarbazole based photo-cross-linking was studied. A combination of counter bases (guanine (G), inosine (I), and cytosine (C)) and photo-cross-linkers (^{CNV}K, ^{NH2V}K, and ^{OHV}K) were used in the ODN to study the best match for acceleration of deamination of cytosine to uracil at physiological conditions. It turned out that the best combination of counter base and photo-cross-linker is inosine and ^{OHV}K which could give ~70% conversion of cytosine to uracil in 7 days at physiological conditions, which could be extended to ~90% in 20 days. Thus, the micro-environment around cytosine, including

hydrogen bonding, hydrophilicity, and polarity of counter base and photo-cross-linker are key players for the photo-cross-link assisted deamination of cytosine to uracil.

Chapter 5: Based on the previous chapters we realized that the micro-environment around the target cytosine is deciding factor for rate of cytosine to uracil conversion via photo-cross-linking. Although, the reaction rate is very rather slow at physiological conditions even when inosine is counter base and ^{OHV}K is photo-cross-linker. Thus, a different approach to accelerate the rate of deamination reaction was used in which the ODN containing photo-cross-linker was divided into two parts between the counter base and photo-cross-linker. The adjoining part was modified with phosphate group at the terminal of counter base to increase the hydrophilicity near the cytosine. It was observed that upon the phosphate group modification near cytosine, ~100% conversion of cytosine to uracil was observed in just 24 h. Furthermore, we removed the ODN with counter base and modified the photo-cross-linker end with phosphate group to study the rate of reaction without hydrogen bonding and high hydrophilicity. It was found that the rate of reaction increased multifold with the modification giving ~100% conversion from cytosine to uracil in 3h at physiological conditions.

These results indicate that the deamination of cytosine to uracil is feasible at physiological conditions and heating to very high temperature is no more necessary to achieve the site-directed mutagenesis via photo-cross-linked cytosine. This has opened vast opportunities to use this enzyme free system in the biological samples at reduced cost and complexity to afford specific and site-directed cytosine to uracil conversions for the treatment of various genetic disorders like Leigh's syndrome.

Achievements

Journal Articles

1. Siddhant Sethi, Minako Ooe, Takashi Sakamoto, Kenzo Fujimoto

Effect of nucleobase change on cytosine deamination through DNA photo-cross-linking reaction via 3-cyanovinylcarbazole nucleoside. **Molecular Biosystems, 2017,** 13(6), 1152-1156

2. Siddhant Sethi, Yasuharu Takashima, Shigetaka Nakamura, Kenzo Fujimoto

Effect of substitution of photo-cross-linker in photochemical cytosine to uracil transition in DNA. **Bioorganic and Medicinal Chemistry Letters, 2017,** 27(16), 3905-3908

3. Siddhant Sethi, Shigetaka Nakamura, Kenzo Fujimoto

Study of Photochemical Cytosine to Uracil Transition via Ultrafast Photo-Cross-Linking Using Vinylcarbazole Derivatives in Duplex DNA. **Molecules, 2018**, 23(4), 828

4. Siddhant Sethi, Nozomi Honda, Shigetaka Nakamura, Kenzo Fujimoto

Super-fast deamination of cytosine in dsDNA using phosphate modified 3-cyanovilylcarbazole assisted photo-crosslinking. **Submitted**

International conferences

1. Siddhant Sethi, Minako Ooe, Takashi Sakamoto, Kenzo Fujimoto.

Photochemical site-directed mutagenesis using 3-cyanovinylcarbazole modified oligdeoxyonucleotide.

42nd International Symposium on Nucleic Acids Chemistry, September, 2015, in Himeji. (Poster)

2. Siddhant Sethi, Takashi Sakamoto, Kenzo Fujimoto.

3-Cyanovinylcarbazole Modified Oligodeoxyrionucleotide Based Photo-chemical DNA Editing as a Tool for Site-directed Mutagenesis

1st Nucleic Acid and Drug Society of Japan, Kyoto, 2015, in Kumamoto. (Poster)

3. Siddhant Sethi, Takashi Sakamoto, Kenzo Fujimoto.

Effects of hydrogen bonding on the cytosine deamination in photo-cross-linked DNA duplex.

43rd International Symposium on Nucleic Acids Chemistry, September, 2016, in Kumamoto. (Poster)

4. Siddhant Sethi, Yasuharu Takashima, Shigetaka Nakamura, Kenzo Fujimoto

Effects of Hydrogen Bonding and Vinylcarbazole Derivatives on 3-Cyanovinylcarbazole Mediated Photo-Cross-Linking Induced Cytosine Deamination

International Conference on Nucleic Acids, June 2017 in Venice. (e-poster)

Domestic conferences

1. Siddhant Sethi, Takashi Sakamoto, Kenzo Fujimoto.

Site-directed mutagenesis using 3-cyanovinylcarbazole modified oligodeoxyribonucleotide based photochemical DNA editing.

Hokuriku Chemical Society of Japan, November 2015 in Kanazawa. (Poster)

2. Siddhant Sethi, Takashi Sakamoto, Kenzo Fujimoto.

Photochemical DNA Editing Using 3-Cyanovinylcarbazole Modified Oligodeoxyribonucleotide As a Method for Site-directed Mutagenesis

96th Annual Meeting of Chemical Society of Japan, March 2016 in Kyoto. (Poster)

3. Siddhant Sethi, Minako Ooe, Takashi Sakamoto, Kenzo Fujimoto.

Evaluation of role of hydrogen bonding in deamination of 3-cyanovinylcarbazole mediated photo-cross-linked cytosine in DNA duplex

Hokuriku Chemical Society of Japan, November 2016 in Fukui. (Poster)

4. Siddhant Sethi, Takashi Sakamoto, Kenzo Fujimoto.

Evaluation of hydrogen bonding for deamination of 3-cyanovinylcarbazole mediated photo-cross-linked cytosine in DNA duplex

10th Bio-related Chemistry Symposium, September 2016 in Kanazawa. (Poster)

5. Siddhant Sethi, Minako Ooe, Takashi Sakamoto, Shigetaka Nakamura, Kenzo Fujimoto.

Effect of counter base of cytosine on deamination of 3-cyanovinyl-carbazole mediated photo-cross-linked cytosine in DNA duplex

137th Annual Meeting of Pharmaceutical Society of Japan, March 2017 in Sendai. (Poster)

6. Siddhant Sethi, Yasuharu Takashima, Shigetaka Nakamura, Kenzo Fujimoto.

Non-enzymatic approach towards cytosine deamination using 3-cyanovinylcarbazole assisted photo-cross-linking

19th Annual Meeting of The RNA Society of Japan, July 2017 in Toyama. (Poster)

Acknowledgement

Firstly, I would like to thank School of Material Science, Japan Advanced Institute of Science and Technology, Japan for providing me the opportunity to work under the brilliant minds of JAIST and giving me the exposure to the rich research experience.

I would like to express my sincere regards and gratitude to **Prof. Kenzo Fujimoto** for his all-embracing support and guidance for quality work and for providing a well-equipped research environment that helped me complete my research well on time.

I would also like to extend my heartiest thanks to Asst. Prof. Takashi Sakamoto and Asst. Prof. Shigetaka Nakamura, who assisted me throughout my research; guided me in all aspects of the work and gave me hope when I needed it the most. Their constant support and caring nature helped me gather my mind in times of need. Thank you for keeping my spirit up!

I give my gratitude to all the members of the Fujimoto Laboratory.

I extend a special thanks to Almighty, my parents (Mr. Devinder Kumar Sethi and Mrs. Rashmi Sethi) and family for their support during my study period and boosting my morale when I needed it the most. I would also like to acknowledge my friends, especially Ms. Yogita, Ms. Olena, Mr. Shubhamoy, Ms. Eena, Mr. Madhav, and Mr. Shubhayan who helped and supported me in every possible way.