

**Development of real-time NMR monitoring method and
elucidation of the deamination mechanism of APOBEC3G**

Keisuke Kamba

Table contents

Chapter 1

General Introduction

1-1. Enzymes, sophisticated biological catalyst.....	1
1-2. AIDS and HIV	2
1-3. Anti-HIV factor APOBEC3G	5
1-4. Vif, an accessory protein of HIV-1, destroys A3G.....	9
1-5. Time course study of A3G's deaminase reaction using NMR methods.....	9
1-6. The aim of this study.....	11
1-7. References.....	11

Chapter 2

Development of a deamination assay involving the real-time NMR monitoring method

2-1. Introduction.....	18
2-2. Results	
2-2-1. Deaminase assay involving the real-time-NMR monitoring method.....	18
2-3. Discussion	
2-3-1. Enzymatic properties of A3G_C as to CCC hotspots.....	20
2-4. Conclusion.....	21
2-5. Experimental Procedures.....	22
2-6. References.....	23

Chapter 3

Nucleic acid determinant for deamination by A3G_C

3-1. Introduction.....	25
3-2. Results	
3-2-1. The effect of ribonucleotide substitution at a single position of a substrate ssDNA on A3G_C's deaminase activity.....	26
3-2-2. The effect of 2'-OH substitution or abasic substitution in the flanking regions of a 5 nucleotide recognition sequence on deaminase activity.....	30
3-2-3. The effect of 2'-OH substitution or abasic substitution at the linker region connecting two CCC hotspots on the 3'→5' polarity.....	32
3-2-4. NaCl concentration dependence of the 3'→5' polarity.....	35
3-3. Discussion	
3-3-1. The recognition sequence of A3G_C for deamination.....	36
3-3-2. The nucleic acid determinant for A3G_C's sliding.....	36
3-3-3. Deamination mechanism of A3G_C.....	37

3-4. Conclusion.....	39
3-5. Experimental Procedures.....	40
3-6. References.....	41

Chapter 4

Deamination property of A3G_C coupled with sliding on the basis of the pH-dependence

4-1. Introduction.....	42
4-2. Results	
4-2-1. The effects of pH and denaturant on the structure of A3G_C.....	42
4-2-2. The effect of pH on the deaminase activity of A3G_C.....	45
4-2-3. The effect of pH on the 3'→5' polarity of A3G_C.....	47
4-3. Discussion	
4-3-1. The origin of ow deaminase activity in higher pH range.....	49
4-3-2. The difficulty of evaluating deaminase activity in lower pH range.....	49
4-3-3. The key interaction for 3'→5' polarity.....	50
4-4. Conclusion.....	50
4-5. Experimental Procedures.....	50
4-6. References.....	52

Chapter 5

Real-time monitoring of the deamination reaction of A3G_C in presence of the Vif protein complex

5-1. Introduction.....	54
5-2. Results	
5-2-1 Constructs of the Vif protein complex.....	54
5-2-2 Deamination assaying in the presence of the Vif protein complex.....	56
5-3. Discussion	
5-3-1 Vif-mediated deamination inhibition.....	56
5-4. Conclusion.....	58
5-5. Experimental Procedures.....	58
5-6. References.....	59

Chapter 6

Deamination activity of A3G_C toward 5-methylcytidine or 5-hydroxymethylcytidine

6-1. Introduction.....	61
6-2. Results	
6-2-1. Deaminase activity toward 5mC.....	62
6-2-2. Deaminase activity toward 5hmC.....	63

6-3. Discussion	
6-3-1. The effects on 5' modification for deaminase activity of A3G_C.....	63
6-4. Conclusion.....	63
6-5. Experimental Procedures.....	65
6-6. References.....	65

Chapter 7

General Conclusion.....	68
List of publications.....	70
List of presentations.....	71
List of Honors.....	75
Acknowledgment.....	76

Chapter 1

General Introduction

1-1. Enzymes, sophisticated biological catalysts

Living organisms possess an excellent mechanism that enables them to obtain and use energy to perform all life activities such as metabolism, growth and movement. Biological macromolecules such as carbohydrates, nucleic acids, lipids and proteins play critical roles in biological phenomena and act as components of organs. Enzymes are sophisticated protein catalysts, most of which are capable of accelerating only specific chemical reactions at neutral pH, mild pressure and physiological temperature (1). Radzicka and Wolfenden mentioned that enzymes increase the rate of a reaction about 5×10^5 – 1×10^{17} fold as compared to under non-enzymatic conditions (2). For example, yeast orotidine 5'-phosphate (OMP) decarboxylase completes OMP decarboxylation within 10 minutes, the half-time being 78,000,000 years without the enzyme. (2). These findings strongly suggest that a biological reaction would take place too slowly for living organisms to maintain their lives in the absence of enzymes. Most enzymes act not only within living organisms but also artificial environments. Therefore, knowledge regarding the mechanism of enzymes is attractive for developing efficient energy extraction systems and material synthesis methods with low environmental loading.

The specificity of enzymes arises from their three-dimensional structures, in which a small area termed active sites and their surroundings are particularly important (1). Substrate molecules bind to the active sites followed by the conversion into products. Therefore, the binding interface between an enzyme and a substrate should be complementary. On the other hand, most enzymes do not exhibit tight binding with their substrates, which indicates that binding strength would have a negative effect on catalytic efficiency because of decreased turn over. It is suggested that an active site without a substrate is somewhat more flexible than the rigid core of the protein to accurately recognize the substrate. Some enzymes require cofactors, non-protein chemical compounds such as metal ions, to recognize a substrate or complete a reaction. The crystal structure of a complex of staphylococcal nuclease (SNase), an endo-exonuclease that digests nucleic acids and preferentially cleaves single-stranded nucleic acids, with cofactor Ca^{2+} and substrate analog inhibitor deoxythymidine 3', 5'-bisphosphate (pdTp) is shown as an example of an enzyme-substrate (ES) complex in Figure 1-1 (3). The substrates in ES complexes are fixed in optimal orientations, which create ideal chemical environments for reactions to occur and result in lowering of activation energy barriers for the reactions (1). In this way, enzymes drastically enhance the specific reactions.

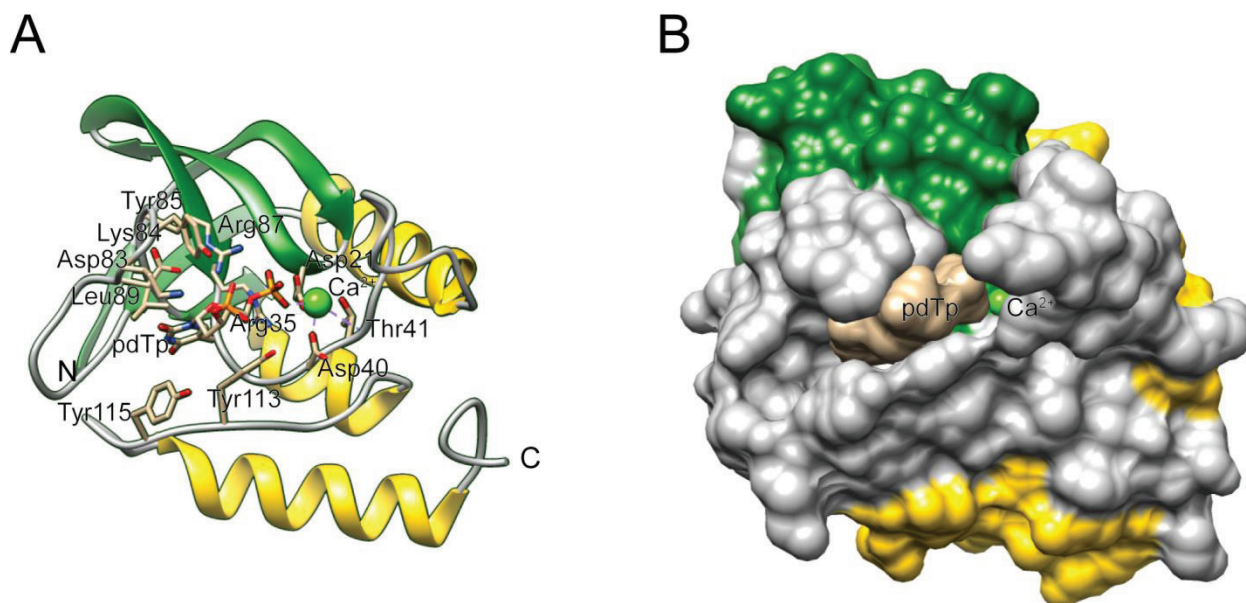


Figure 1-1.

A typical picture of an enzyme-substrate complex.

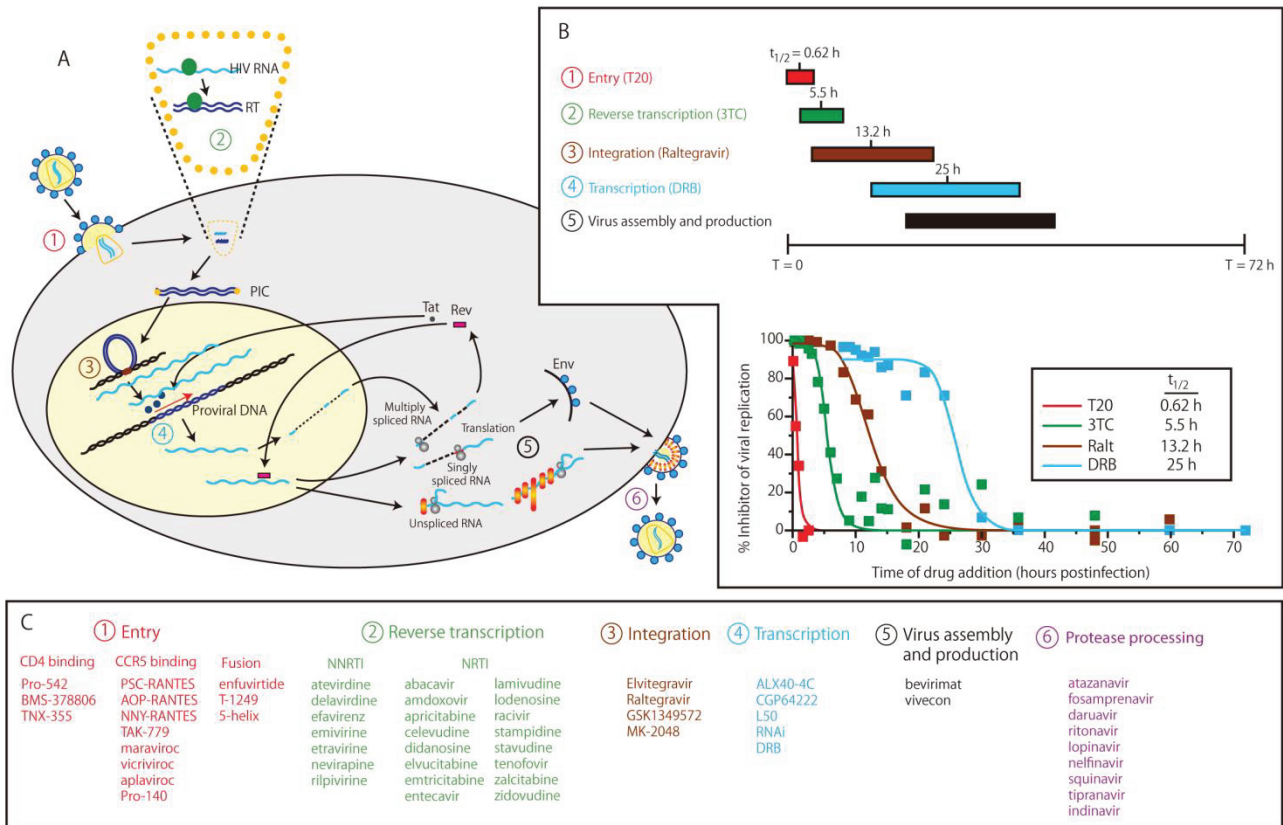
The crystal structure of the SNase complex (PDB ID: 1SNC) is shown as a ribbon diagram (A) and a molecular surface model (B) (3). The α -helices, β -strands and loops in SNase are colored yellow, green and gray, respectively. The amino acids comprising the active site are represented as a stick model in (A). The cofactor, Ca^{2+} , is presented as a light green sphere. The substrate analog inhibitor, pdTp, is shown as a stick model in (A), and its molecular surface is colored beige. Molecular graphics and analyses were performed with the UCSF Chimera package (4). Chimera was developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIGMS P41-GM103311).

1-2. AIDS and HIV

Acquired immune deficiency syndrome (AIDS) is a fatal condition caused by human immunodeficiency virus (HIV) (5, 6). Worldwide, an estimated 37 million people have been infected with the HIV and about 1.2 million people died from AIDS-related illnesses in 2014 (6). HIV breaks down the human immune system by destroying CD4^+ T cells (thymocytes cells that express a cluster of differentiation 4 glycoproteins on their surfaces) (5). There are two distinct types of HIV, of which the predominant one distributed throughout the world is HIV type 1 (HIV-1), the relatively uncommon HIV type 2 (HIV-2) being found mainly in West Africa (7). So far, no procedure or medication has been established to eliminate HIV from patients as well as kill their infected cells (5). This means that once a person is infected with HIV, they have to use antiviral drugs for the rest of their lives. Without treatment, almost all people infected with HIV inevitably will develop AIDS followed by death because a destroyed immune system cannot overcome infections by bacterial, viral,

fungal, or protozoan pathogens. To make the situation worse, mutations in the HIV genome occur so frequently that cessation of treatment may generate a drug-resistant virus and they cannot use the same drugs any longer (8). Therefore, novel HIV treatments as well as new type of anti-viral drug have been desired.

The six major steps of the HIV-1 life cycle as well as the current anti-HIV drugs are briefly reviewed in Figure 1-2 (8). The first step in the HIV-1 replication cycle is viral entry. The HIV-1 envelope protein gp120 has the ability to bind the CD4 receptor (as well as co-receptor CXCR4 or CCR5) on the surface of the CD4⁺ cell, which causes fusion of the viral and host membranes (5, 8). This phenomenon enables HIV-1 to insert its cell contents including the viral RNA genome into the host cell (5, 8). The second step is the reverse transcription required to convert the single-stranded HIV-1 RNA genome into double-stranded DNA (dsDNA) catalyzed by viral reverse transcriptase (RT), which is a multifunctional enzyme with RNA-dependent DNA polymerase, RNase-H, and DNA-dependent DNA polymerase activities (5, 8). The viral dsDNA synthesized by RT forms a large nucleoprotein complex, the preintegration complex (PIC) (5, 8). Thirdly, the viral dsDNA is transported to the nucleus and inserted into the host DNA (this inserted DNA is also called proviral DNA) by integrase (5, 8). Integration is necessary to maintain the viral genome in the infected cell and an imperative for propagating a new virus. In the transcription step, the fourth step of the viral life cycle, a new HIV RNA genome and viral mRNAs are synthesized by the host RNA synthesis system (5, 8). Note that the HIV Tat (trans-activator of transcription) protein drastically enhances the efficiency of viral transcription (8). The viral mRNAs are translated and viral proteins are produced, then the viral components are assembled and mature on the inner membrane in the fifth step. In the final stage, longer viral proteins are cleaved by proteases and a new HIV-1 assembly buds from the host cell (5, 8). The most effective treatments for HIV-1 are antiretroviral drug therapies (8).



1-3. Anti-HIV factor APOBEC3G

Apolipoprotein B messenger RNA-editing enzyme catalytic polypeptide-like 3G (APOBEC3G or A3G) inactivates HIV-1 by introducing a significant number of mutations into the viral genome (9–11), which is regarded as a potential drug target for HIV-1. A3G was originally identified as a human gene, named CEM15, encoded by human chromosome 22 that inhibits HIV-1 infection (12). Then CEM15 was renamed APOBEC3G because it possesses cytosine deaminase activity, which converts cytosines (Cs) into uracils (Us), and its amino acid sequence belongs to the APOBEC family (9, 13–15). This family, ‘APOBEC’, is named after apolipoprotein B mRNA-editing enzyme catalytic polypeptides 1 (APOBEC1) protein, which edits apoB-100 mRNA by deaminating Cs into Us. As a result of APOBEC1 deaminase activity, CAA is converted into UAA, and a stop codon is introduced into apoB-100 mRNA (16). A3G edits the newly reverse-transcribed minus strands of HIV-1 cDNA, single-stranded DNA (ssDNA), in the cytosol (9, 11, 13, 17). A3G’s deaminase activity, C to U conversion, leads to the introduction of a significant number of guanine (G) to adenine (A) mutations into the integrated viral genomic DNA, which causes inactivation of HIV-1 lacking virion infectivity factor (Vif) protein (9, 11–13, 17–19). A schematic representation of the anti-HIV activity of A3G is shown in Figure 1-3. A3G also exerts deamination-dependent anti-viral activity against many other retroviruses such as hepatitis B virus (20) and human T-cell leukemia virus type 1 (21).

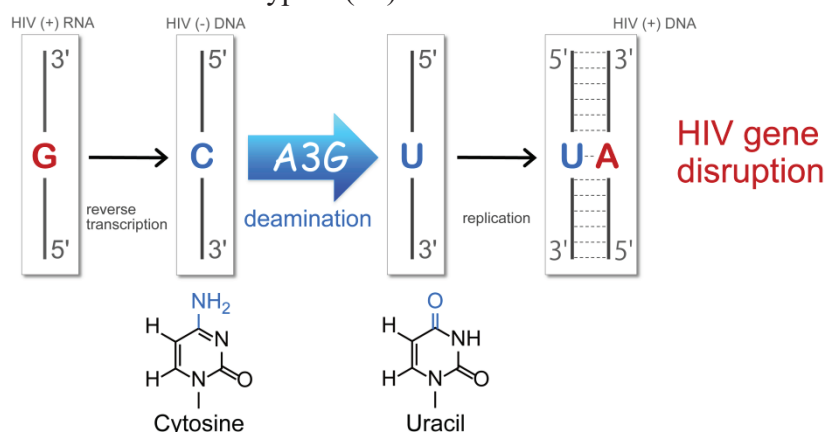


Figure 1-3.

Schematic representation of anti-HIV activity of A3G.

A3G deaminates cytosines (C) into uracils (U) within the newly reverse-transcribed minus strands of HIV-1 DNA (9–11).

The deaminase activity of A3G is ssDNA-specific, it exhibiting no activity against RNA (22–24) or dsDNA (23–25). A3G also exerts sequence specific deaminase activity, which converts cytosine dinucleotide (CC) into CU (13, 24, 26). Interestingly, A3G deaminates triplet cytosine (CCC) hotspots, the most favorite target of A3G, converting them into CCU and then CUU (13, 23–28). A3G is prone to deaminate a CCC hotspot that is located close to the 5' end more effectively than one that is less close to the 5' end (25, 29–32). We call this deamination bias 3'→5' polarity,

which introduces a 5'→3' gradient of G to A mutations into the HIV gene (23, 33, 34). 3'→5' polarity cannot be explained by general kinetic models including the Michaelis–Menten theory (29). Furukawa and coworkers successfully explained 3'→5' polarity rationally by constructing a kinetic model in which A3G slides along ssDNA through direction-independent Brownian movement, and that deamination is more efficient when A3G approaches the target cytosine in the 3' to 5' direction rather than the opposite one (29). The essence of origin of the 3'→5' polarity proposed by Furukawa *et al.* is shown in Figure 1-4 (29). Senavirathne and coworkers performed a single-molecule fluorescence resonance energy transfer (sm-FRET) experiment, and analyzed the distance between A3G and ssDNA (31). They concluded that A3G scans ssDNA randomly and bidirectionally, favoring nonrandom 3' to 5' deamination (31), which agreed with the model proposed Furukawa *et al.* (29).

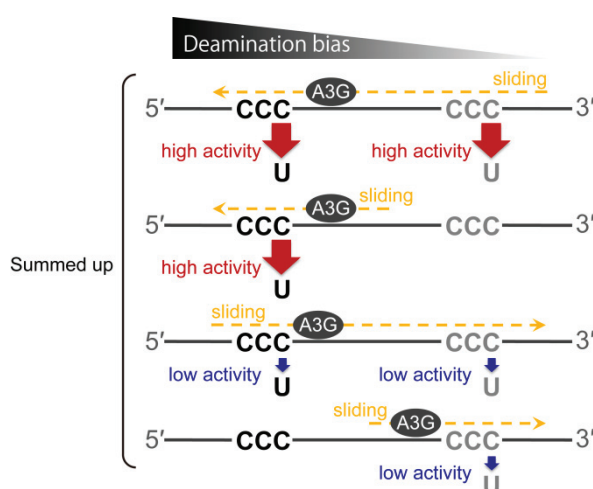


Figure 1-4.

The proposed origin of 3'→5' polarity.

A3G deaminates a CCC hotspot that is located close to the 5' end more effectively than ones that are less close to the 5' end. This phenomena is called 3'→5' polarity. A3G slides along ssDNA with no direction preference, and deamination rate depend on the direction of the sliding that confers the 3'→5' polarity (29).

The deamination mechanism of A3G has been studied by many scientists. A3G is composed of two zinc finger deaminase domains: an enzymatically inactive N-terminal domain (residues 1–192; A3G_N) and an active C-terminal domain (residues 193–384; A3G_C) (9–11, 35, 36). Each deaminase motif comprises a consensus amino-acid sequence, His or Cys-Xaa-Glu-Xaa_{23–28}-Pro-Cys-Xaa_{2–4}-Cys (Xaa denotes any amino acid) (9–11, 14). A3G_N is responsible for RNA binding and intrudes into the HIV virion during viral encapsidation (36–38). A3G_C retains deamination activity, on which the anti-viral effect of A3G depends (34, 36, 39). Each domain of A3G is reportedly involved in DNA binding (40) and deamination as well as 3'→5' polarity (25, 29–32). The deamination mechanism in the active site, shown in Figure 1-5, was previously proposed

based on the deamination mechanisms that were predicted in structural studies of a bacterial and yeast cytosine deaminases (9). Glutamic acid 259 activates the closely positioned water molecule and its nucleophilic attack drives the deamination reaction (9). Because of this, Glu259 is called the catalytic residue of A3G. The molecular structures of free A3G_C (27, 30, 41, 42) and truncated A3G_N like protein (43) were determined by X-ray crystallography and solution nuclear magnetic resonance (NMR) spectroscopy. Three different model structures of the A3G_C-ssDNA complex, the active site comprising the catalytic residue (Glu259) and a zinc finger (His257, Cys288, Cys291), have been proposed from three groups (11, 27, 30, 41, 44). Figure 1-6 shows the structure of A3G_C reported by Furukawa *et al.* (27), and the three proposed DNA binding paths (27, 30, 41). No A3G-ssDNA complex structure is available, although the A3G-ssDNA interaction has been investigated by means of various methods, such as electrophoretic mobility shift assaying (EMSA) (23–25, 27, 30, 35), steady-state rotational-anisotropy binding assays (25, 32), and atomic force microscopy (AFM) (32, 45–47), which suggested that A3G binds to ssDNA in a sequence-nonspecific manner. Furthermore, one-dimensional movement of full length A3G over ssDNA, called sliding, was directly observed on smFRET (31) and AFM (45), which indicates A3G undergoes multiple interactions to associate with ssDNA. It was also reported that full length A3G exhibits stronger activity (30–32) and 3'→5' polarity (30–32) than A3G_C even though A3G_C itself possesses deaminase activity (27–30) and 3'→5' polarity (29, 30). The findings make understanding of the deamination mechanism of A3G complicated.

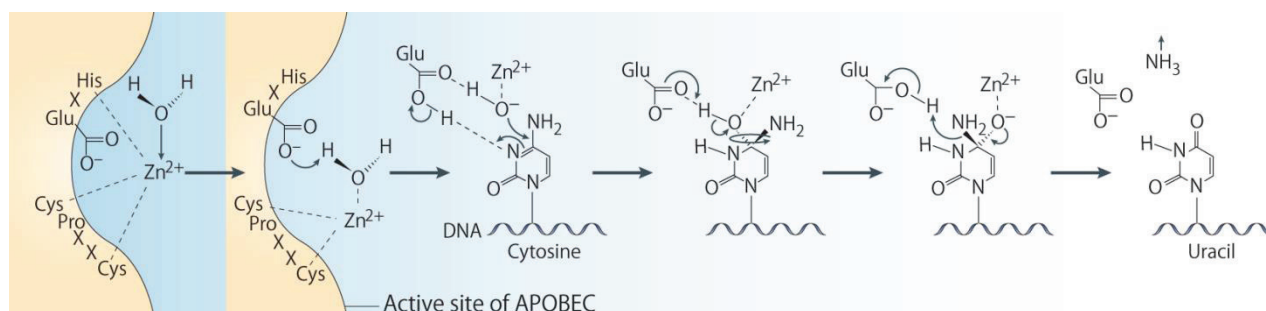


Figure 1-5.

A possible mechanism for A3G_C-dependent cytosine deamination.

First, the cysteine (Cys) and histidine (His) residues of the active site coordinate to a single zinc ion (Zn^{2+}); an incoming water molecule (H_2O) is shown. Second, a hydroxide ion (OH^-) is produced when the water reacts with glutamic acid and zinc. Only the zinc ion and the side chain of the catalytic glutamic acid (Glu) are illustrated beyond this step. Third, glutamic acid protonates N3 of the cytosine pyrimidine ring, which destabilizes the N3–C4 double bond and renders C4 susceptible to attack by the hydroxide ion. Fourth, the latter results in the formation of a tetrahedral (transition state) intermediate, as the proton (H^+) of the water molecule is sequestered by glutamic acid. Fifth, the amino group (NH_2) receives this proton in a reaction that causes cleavage of the C–N bond and formation of a C4–O double bond. Sixth, uracil and ammonia (NH_3) are released from the active site. This figure is Figure 2C in Harris and Liddament (9).

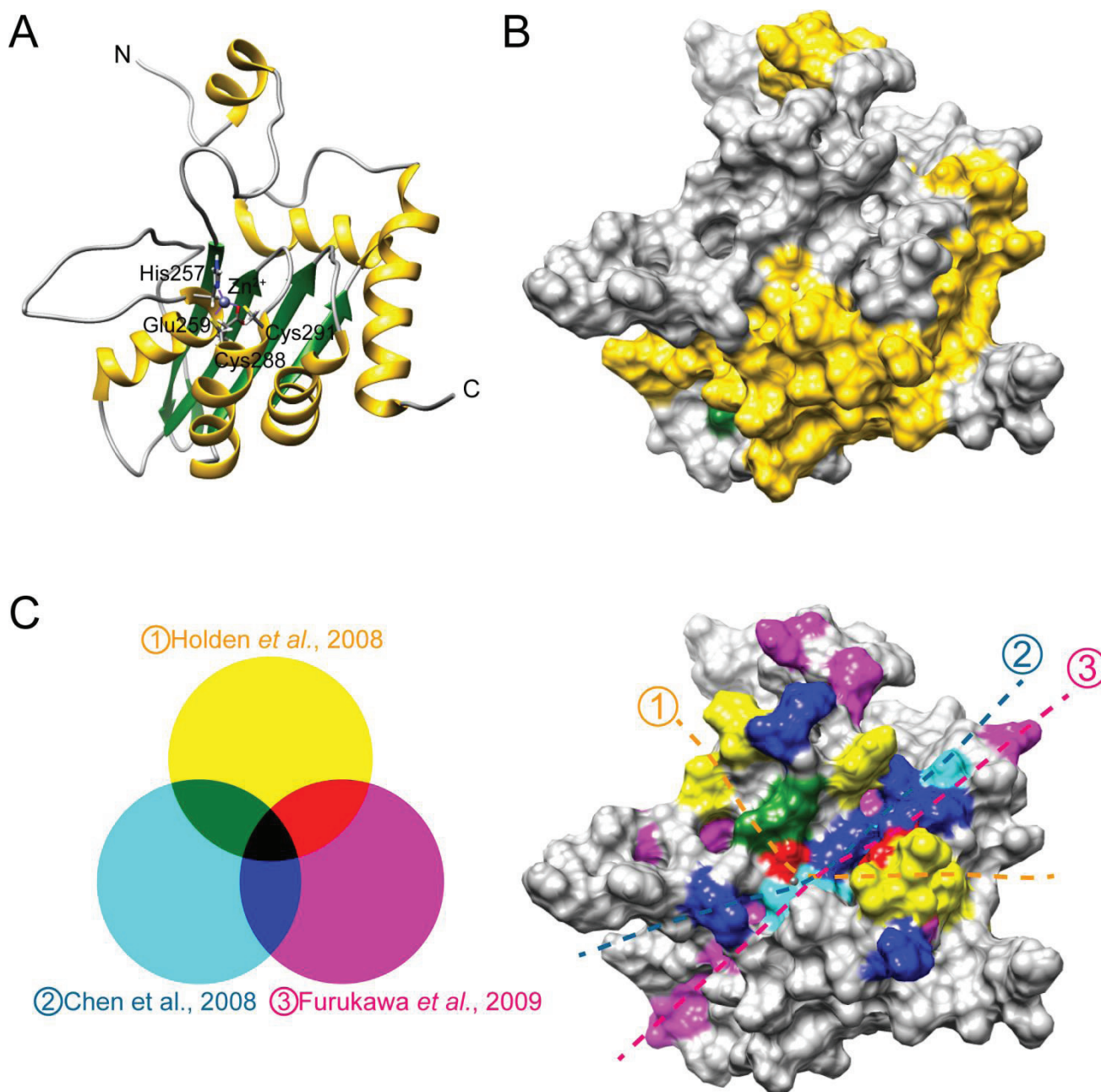


Figure 1-6.

The molecular structure of A3G_C.

The solution NMR structure of A3G_C (PDB ID: 2KBO) is shown as a ribbon diagram (A) and a molecular surface model (B) (27). The α -helices, β -strands and loops in A3G_C are colored yellow, green and gray, respectively. The amino acids comprising the active site (His257, Glu259, Cys288, Cys291) are represented as a stick model in (A). The cofactor, Zn^{2+} , is presented as a gray sphere. (C) The three proposed ssDNA binding paths. The colors used in this figure are shown on the left. Molecular graphics and analyses were performed with the UCSF Chimera package (4).

1-4. Vif, an accessory protein of HIV-1, destroys A3G

In spite of anti-viral factor A3G, infection occurs and develops into AIDS, which is caused by A3G inactivation owing to the viral defense system (5). Vif, a HIV-1 accessory protein that is not essential for viral replication, is responsible for annihilation of A3G (11, 48, 49). Figure 1-7 reviews three major mechanisms of liquidation of A3G by Vif (48). Vif mainly destroys A3G through the poly-ubiquitination on proteasomal degradation, which reduces the intracellular level of A3G (Figure 1-7, pathway 2) (50–54). Vif is able to bind A3G, and to recruit an E3 ubiquitin ligase complex composed of Elongin B (EloB), Elongin C (EloC), Cullin5 (Cul5), and RING box protein 2 (Rbx2), leading to the poly-ubiquitination following degradation of A3G (53–56). Vif grasps the A3G molecule by binding to A3G_N and recruits it into the E3 ubiquitin ligase complex (57–59). Then the four lysine residues in A3G_C are ubiquitinated (60). It has also been found that the transcription cofactor core-binding factor β subunit (CBF β), a part of the heterodimeric enhancer-binding protein that is associated with genetic regulation of hematopoiesis and osteogenesis, was additionally recruited to this ubiquitin ligase complex (61–63). Vif has also been shown to inhibit translation of A3G (Figure 1-7, pathway 3) (64–66). It was suggested that Vif binds to the 5'-untranslated region (UTR) of A3G mRNA, which results in inhibition of the translation of A3G (66). In the last pathway, Vif reportedly blocks packaging of A3G (Figure 1-7 pathway 1) (67). During virus assembly, A3G as well as Vif interact with the HIV-1 Gag precursor protein, Pr55^{Gag}, and the viral genomic RNA, which indicates that Vif acts as a competitive inhibitor against A3G packaging (37, 38, 68–70). Because of this A3G annihilation function, Vif has been paid great attention as a drug target for anti-HIV-1 treatment, and a HIV-1 inhibitor targeting Vif-dependent degradation of A3G is in the process of being developed (71).

1-5. Time course study of A3G's deaminase reaction using NMR methods

The time-resolved NMR method is a powerful tool for evaluating chemical properties and dynamics such as enzyme dynamics (72), chemical modifications in proteins (73), protein folding (74), metabolic dynamics (75) and for characterization of unstable intermediates in a chemical reaction (76). Katahira's group have utilized NMR method, monitoring the time course of one-dimensional proton (1D ¹H) NMR spectra of CCC within 10-mer ssDNA and performing its chemical shift assignment using two dimensional (2D) ¹H–¹³C heteronuclear single quantum correlation (HSQC) spectra, for evaluating enzymatic activity, and demonstrated that A3G converts CCC into CCU and then CUU sequentially for the first time (27). Then Furukawa and coworkers extended this method by applying 2D total correlated spectroscopy (TOCSY) spectra to improve the resolution of the spectra (29). After analyzing kinetic parameters obtained by deamination assaying, they proposed a novel kinetic model assuming A3G slides along ssDNA with no direction preference, and deamination rate depend on the direction of the sliding, and succeeded in explaining 3'→5' polarity. This 'real-time NMR monitoring method' has become a popular method for evaluating the deaminase activity of A3G (27–29, 43, 77) and other APOBEC family proteins (78, 79). This

‘real-time NMR monitoring method’ is expected to allow the evaluation of more complicated mechanisms such as the A3G-ssDNA interaction during the deamination and deamination reactions in the presence of Vif and the E3 ubiquitin ligase complex.

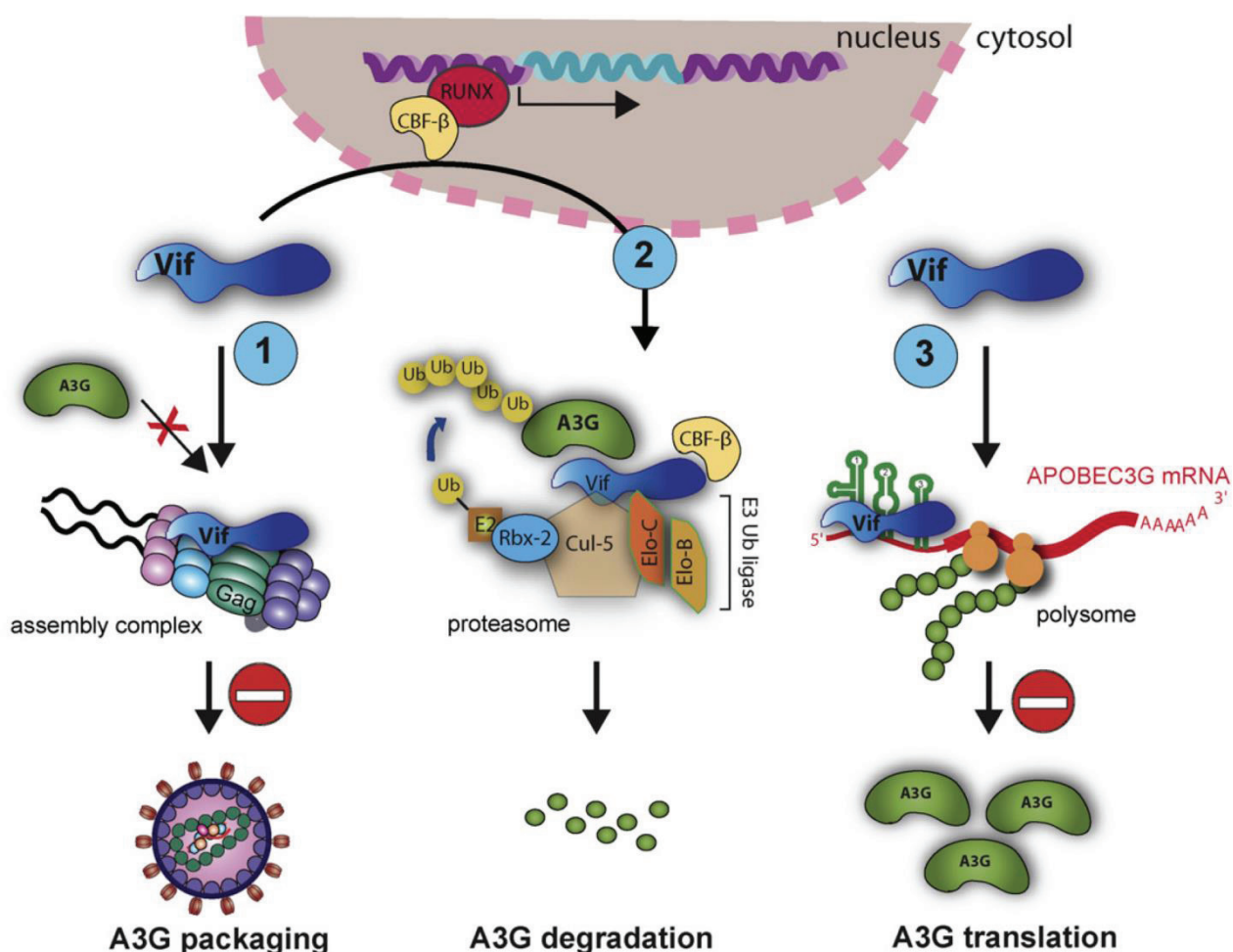


Figure 1-7.

HIV-1 Vif protein liquidate A3G.

Vif neutralizes A3G through different mechanisms. [1] Vif competes with A3G for binding to viral components like the nucleocapsid domain of the viral Gag protein and the genomic RNA. [2] Vif binding to A3G recruits an E3 ubiquitin ligase that mediates the polyubiquitination of A3G and its degradation. [3] Vif impairs the translation of A3G mRNA through an mRNA-binding mechanism. Taken together, these three different actions of Vif on translation, degradation and packaging not only deplete A3G from virus-producing cells but also prevent A3G from being incorporated into virions. This figure is Figure 1A in Batisse *et al.* (48).

1-6. The aim of this study

In this study, we have developed a NMR method for investigating the deamination mechanism of A3G_C. We found that A3G_C recognizes consecutive five nucleotides, the CCC hotspot being positioned at the center, and that base and sugar moieties are not recognized during sliding on NMR-based deamination assaying using nucleic acid analogs. Examination of the effects of different salt concentrations on the 3'→5' polarity indicated that the sliding is alleviated due to a decrease in the electrostatic interaction between A3G_C and the phosphate backbone. This finding was confirmed deamination assaying at different pHs. Our NMR-based enzymatic method is useful in presence of Vif and the E3 ubiquitin ligase complex. It was found that Vif and the E3 ubiquitin ligase complex may associate with A3G_C as well as ssDNA, which indicates Vif protein directly inhibits deamination as well as 3'→5' polarity. Finally, we applied this method to modified cytosines, 5-methylcytosine (5mC) or 5-hydroxymethylcytosine (5hmC), containing DNA, and we demonstrated for the first time that A3G_C retains deaminase activity against 5mC to produce thymine, although A3G_C cannot edit 5hmC into 5-hydroxymethyluracil.

1-7. References

- (1) Robinson, P. K. (2015) Enzymes: principles and biotechnological applications. *Essays Biochem.* **59**, 1–41.
- (2) Radzicka, A. and Wolfenden, R. (1995) A proficient enzyme. *Science* **267**, 90–93.
- (3) Loll, P. J. and Lattman, E. E. (1989) The crystal structure of the ternary complex of staphylococcal nuclease, Ca^{2+} and the inhibitor pdTp, refined at 1.65 Å. *Proteins* **5**, 183–201.
- (4) Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612.
- (5) Barré-Sinoussi, F., Ross, A. L. and Delfraissy, J. F. (2013) Past, present and future: 30 years of HIV research. *Nat. Rev. Microbiol.* **11**, 877–883.
- (6) *UNAIDS World AIDS Day Report* (2015).
- (7) Nyamweya, S., Hegedus, A., Jaye, A., Rowland-Jones, S., Flanagan, K. L., and Macallan, D. C. (2013) Comparing HIV-1 and HIV-2 infection: Lessons for viral immunopathogenesis. *Rev. Med. Virol.* **23** 221–240.
- (8) Arts, E. J. and Hazuda, D. J. (2012) HIV-1 antiretroviral drug therapy. *Cold Spring Harb. Perspect. Med.* **2**, doi: 10.1101/cshperspect.a007161.
- (9) Harris, R. S. and Liddament, M. T. (2004) Retroviral restriction by APOBEC proteins. *Nat. Rev. Immunol.* **4**, 868–867.
- (10) Kitamura, S., Ode, H. and Iwatani, Y. (2011) Structural features of antiviral APOBEC3 proteins are linked to their functional activities. *Front. Microbiol.* **2**, doi: 10.3389/fmicb.2011.00258.
- (11) Aydin, H., Taylor, M. W. and Lee, J. E. (2014) Structure-guided analysis of the human APOBEC3-HIV restrictome. *Structure* **22**, 668–684.

- (12) Sheehy, A. M., Gaddis, N. C., Choi, J. D. and Malim, M. H. (2002) Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* **418**, 646–650.
- (13) Harris, R. S., Bishop, K. N., Sheehy, A. M., Craig, H. M., Petersen-Mahrt, S. K., Watt, I. N., Neuberger, M. S. and Malim, M. H. (2003) DNA deamination mediates innate immunity to retroviral infection. *Cell* **113**, 803–809.
- (14) Harris, R. S., Petersen-Mahrt, S. K. and Neuberger, M. S. (2002) RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. *Mol. Cell* **10**, 1247–1253.
- (15) Jarmuz, A., Chester, A., Bayliss, J., Gisbourne, J., Dunham, I., Scott, J. and Navaratnam, N. (2002) An anthropoid-specific locus of orphan C to U RNA-editing enzymes on chromosome 22. *Genomics* **79**, 285–296.
- (16) Nakamuta, M., Oka, K., Krushka, J., Kobayashi, K., Yamamoto, M., Li, W. H. and Chan, L. (1995) Alternative mRNA splicing and differential promoter utilization determine tissue-specific expression of the apolipoprotein B mRNA-editing protein (ApoBec1) gene in mice. Structure and evolution of ApoBec1 and related nucleoside/nucleotide deaminases. *J. Biol. Chem.* **270**, 13042–13056.
- (17) Zhang, H., Yang, B., Pomerantz, R. J., Zhang, C., Arunachalam, S. C. and Gao, L. (2003) The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. *Nature* **424**, 94–98.
- (18) Lecossier, D., Bouchonnet, F., Clavel, F. and Hance, A.J. (2003) Hypermutation of HIV-1 DNA in the absence of the Vif protein. *Science* **300**, 1112.
- (19) Mangeat, B., Turelli, P., Caron, G., Friedli, M., Perrin, L. and Trono, D. (2003) Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature* **424**, 99–103.
- (20) Henry, M., Guétard, D., Suspène, R., Rusniok, C., Wain-Hobson, S. and Vartanian, J. P. (2009) Genetic editing of HBV DNA by monodomain human APOBEC3 cytidine deaminases and the recombinant nature of APOBEC3G. *PLoS One* **4**, doi: 10.1371/journal.pone.0004277.
- (21) Mahieux, R., Suspène, R., Delebecque, F., Henry, M., Schwartz, O., Wain-Hobson, S. and Vartanian, J. P. (2005) Extensive editing of a small fraction of human T-cell leukemia virus type 1 genomes by four APOBEC3 cytidine deaminases. *J. Gen. Virol.* **86**, 2489–2494.
- (22) Bishop, K. N., Holmes, R. K., Sheehy, A. M. and Malim, M. H. (2004) APOBEC-mediated editing of viral RNA. *Science* **305**, 645.
- (23) Yu, Q., König R., Pillai, S., Chiles, K., Kearney, M., Palmer, S., Richman, D., Coffin, J. M. and Landau, N. R. (2004) Single-strand specificity of APOBEC3G accounts for minus-strand deamination of the HIV genome. *Nat. Struct. Mol. Biol.* **11**, 435–442.
- (24) Iwatani, Y., Takeuchi, H., Strebel, K. and Levin, J. G. (2006) Biochemical activities of highly purified, catalytically active human APOBEC3G: correlation with antiviral effect. *J. Virol.* **80**, 5992–6002.
- (25) Chelico, L., Pham, P., Calabrese, P. and Goodman, M. F. (2006) APOBEC3G DNA deaminase

- acts processively 3' → 5' on single-stranded DNA. *Nat. Struct. Mol. Biol.* **13**, 392–399.
- (26) Langlois, M. A., Beale, R. C., Conticello, S. G. and Neuberger, M. S. (2005) Mutational comparison of the single-domained APOBEC3C and double-domained APOBEC3F/G anti-retroviral cytidine deaminases provides insight into their DNA target site specificities. *Nucleic Acids Res.* **33**, 1913–1923.
- (27) Furukawa, A., Nagata, T., Matsugami, A., Habu, Y., Sugiyama, R., Hayashi, F., Kobayashi, N., Yokoyama, S., Takaku, H. and Katahira, M. (2009) Structure, interaction and real-time monitoring of the enzymatic reaction of wild-type APOBEC3G. *EMBO J.* **28**, 440–451.
- (28) Harjes, S., Solomon, W. C., Li, M., Chen, K. M., Harjes, E., Harris, R. S. and Matsuo, H. (2013) Impact of H216 on the DNA binding and catalytic activities of the HIV restriction factor APOBEC3G. *J. Virol.* **87**, 7008–7014.
- (29) Furukawa, A., Sugase, K., Morishita, R., Nagata, T., Kodaki, T., Takaori-Kondo, A., Ryo, A., and Katahira, M. (2014) Quantitative analysis of location- and sequence-dependent deamination by APOBEC3G using real-time NMR spectroscopy. *Angew. Chem. Int. Ed. Engl.* **53**, 2349–2352.
- (30) Holden, L. G., Prochnow, C., Chang, Y. P., Bransteitter, R., Chelico, L., Sen, U., Stevens, R. C., Goodman, M. F. and Chen, X. S. (2008) Crystal structure of the anti-viral APOBEC3G catalytic domain and functional implications. *Nature* **456**, 121–124.
- (31) Senavirathne, G., Jaszczur, M., Auerbach, P. A., Upton, T. G., Chelico, L., Goodman, M. F. and Rueda, D. (2012) Single-stranded DNA scanning and deamination by APOBEC3G cytidine deaminase at single molecule resolution. *J. Biol. Chem.* **287**, 15826–15835.
- (32) Chelico, L., Prochnow, C., Erie, D. A., Chen, X. S. and Goodman, M. F. (2010) Structural model for deoxycytidine deamination mechanisms of the HIV-1 inactivation enzyme APOBEC3G. *J. Biol. Chem.* **285**, 16195–16205.
- (33) Suspène, R., Rusniok, C., Vartanian, J. P. and Wain-Hobson, S. (2006) Twin gradients in APOBEC3 edited HIV-1 DNA reflect the dynamics of lentiviral replication. *Nucleic Acids Res.* **34**, 4677–4684.
- (34) Kobayashi, T., Koizumi, Y., Takeuchi, J. S., Misawa, N., Kimura, Y., Morita, S., Aihara, K., Koyanagi, Y., Iwami, S. and Sato, K. (2014) Quantification of deaminase activity-dependent and -independent restriction of HIV-1 replication mediated by APOBEC3F and APOBEC3G through experimental-mathematical investigation. *J. Virol.* **88**, 5881–5887.
- (35) Haché, G., Liddament, M. T. and Harris, R. S. (2005) The retroviral hypermutation specificity of APOBEC3F and APOBEC3G is governed by the C-terminal DNA cytosine deaminase domain. *J. Biol. Chem.* **280**, 10920–10924.
- (36) Navarro, F., Bollman, B., Chen, H., König, R., Yu, Q., Chiles, K. and Landau, N. R. (2005) Complementary function of the two catalytic domains of APOBEC3G. *Virology* **333**, 374–386.
- (37) Khan, M. A., Kao, S., Miyagi, E., Takeuchi, H., Goila-Gaur, R., Opi, S., Gipson, C. L., Parslow, T. G., Ly, H. and Strebel, K. (2005) Viral RNA is required for the association of APOBEC3G with human immunodeficiency virus type 1 nucleoprotein complexes. *J. Virol.* **79**, 5870–5874.

- (38)Khan, M. A., Goila-Gaur, R., Opi, S., Miyagi, E., Takeuchi, H., Kao, S. and Strebel, K. (2007) Analysis of the contribution of cellular and viral RNA to the packaging of APOBEC3G into HIV-1 virions. *Retrovirology* **4**, doi:10.1186/1742-4690-4-48.
- (39)Browne, E. P., Allers, C. and Landau, N. R. (2009). Restriction of HIV-1 by APOBEC3G is cytidine deaminase-dependent. *Virology* **387**, 313–321.
- (40)Shindo, K., Li, M., Gross, P. J., Brown, W. L., Harjes, E., Lu, Y., Matsuo, H. and Harris, R. S. (2012) A comparison of two single-stranded DNA binding models by mutational analysis of APOBEC3G. *Biology (Basel)* **1**, 260–276.
- (41)Chen, K. M., Harjes, E., Gross, P. J., Fahmy, A., Lu, Y., Shindo, K., Harris, R. S. and Matsuo, H. (2008) Structure of the DNA deaminase domain of the HIV-1 restriction factor APOBEC3G. *Nature* **452**, 116–119.
- (42)Harjes, E., Gross, P. J., Chen, K. M., Lu, Y., Shindo, K., Nowarski, R., Gross, J. D., Kotler, M., Harris, R. S. and Matsuo, H. (2009) An extended structure of the APOBEC3G catalytic domain suggests a unique holoenzyme model. *J. Mol. Biol.* **389**, 819–832.
- (43)Kouno, T., Luengas, E. M., Shigematsu, M., Shandilya, S. M., Zhang, J., Chen, L., Hara, M., Schiffer, C. A., Harris, R. S. and Matsuo, H. (2015) Structure of the Vif-binding domain of the antiviral enzyme APOBEC3G. *Nat. Struct. Mol. Biol.* **22**, 485–491.
- (44)Bransteitter, R., Prochnow, C. and Chen, X. S. (2009) The current structural and functional understanding of APOBEC deaminases. *Cell. Mol. Life Sci.* **66**, 3137–3147.
- (45)Shlyakhtenko, L. S., Lushnikov, A. Y., Miyagi, A., Li, M., Harris, R. S. and Lyubchenko, Y. L. (2012) Nanoscale structure and dynamics of ABOBEC3G complexes with single-stranded DNA. *Biochemistry* **51**, 6432–6440.
- (46)Shlyakhtenko, L. S., Lushnikov, A. Y., Miyagi, A., Li, M., Harris, R. S. and Lyubchenko, Y. L. (2013) Atomic force microscopy studies of APOBEC3G oligomerization and dynamics. *J. Struct. Biol.* **184**, 217–225.
- (47)Shlyakhtenko, L. S., Dutta, S., Banga, J., Li, M., Harris, R. S. and Lyubchenko, Y. L. (2015) APOBEC3G interacts with ssDNA by two modes: AFM studies. *Sci. Rep.* **5**, doi: 10.1038/srep15648.
- (48)Batisse, J., Guerrero, S., Bernacchi, S., Sleiman, D., Gabus, C., Darlix, J. L., Marquet, R., Tisné, C. and Paillart, J. C. (2012) The role of Vif oligomerization and RNA chaperone activity in HIV-1 replication. *Virus Res.* **169**, 361–376.
- (49)Desimmie, B. A., Delviks-Frankenberry, K. A., Burdick, R. C., Qi, D., Izumi, T. and Pathak, V. K. (2014) Multiple APOBEC3 restriction factors for HIV-1 and one Vif to rule them all. *J. Mol. Biol.* **426**, 1220–1245.
- (50)Marin, M., Rose, K. M., Kozak, S. L. and Kabat, D. (2003) HIV-1 Vif protein binds the editing enzyme APOBEC3G and induces its degradation. *Nat. Med.* **9**, 1398–1403.
- (51)Conticello, S. G., Harris, R. S. and Neuberger, M. S. (2003) The Vif protein of HIV triggers degradation of the human antiretroviral DNA deaminase APOBEC3G.

Curr. Biol. **13**, 2009-2013.

- (52) Sheehy, A. M., Gaddis, N. C. and Malim, M. H. (2003) The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif. *Nat. Med.* **9**, 1404–1407.
- (53) Yu, X., Yu, Y., Liu, B., Luo, K., Kong, W., Mao, P. and Yu, X. F. (2003) Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. *Science* **302**, 1056–1060.
- (54) Mehle, A., Strack, B., Ancuta, P., Zhang, C., McPike, M. and Gabuzda D. (2004) Vif overcomes the innate antiviral activity of APOBEC3G by promoting its degradation in the ubiquitin-proteasome pathway. *J. Biol. Chem.* **279**, 7792–7798.
- (55) Luo, K., Xiao, Z., Ehrlich, E., Yu, Y., Liu, B., Zheng, S. and Yu, X. F. (2005) Primate lentiviral virion infectivity factors are substrate receptors that assemble with cullin 5-E3 ligase through a HCCH motif to suppress APOBEC3G. *Proc. Natl. Acad. Sci. USA* **102**, 11444–11449.
- (56) Mehle, A., Thomas, E. R., Rajendran, K. S. and Gabuzda, D. (2006) A zinc-binding region in Vif binds Cul5 and determines cullin selection. *J. Biol. Chem.* **281**, 17259–17265.
- (57) Mehle, A., Wilson, H., Zhang, C., Brazier, A. J., McPike, M., Pery, E. and Gabuzda, D. (2007) Identification of an APOBEC3G binding site in human immunodeficiency virus type 1 Vif and inhibitors of Vif-APOBEC3G binding. *J. Virol.* **81**, 13235–13241.
- (58) Uyttendaele, I., Lavens, D., Catteuw, D., Lemmens, I., Bovijn, C., Tavernier, J. and Peelman, F. (2011) Random mutagenesis MAPPIT analysis identifies binding sites for Vif and Gag in both cytidine deaminase domains of Apobec3G. *PLoS One* **7**, doi: 10.1371/journal.pone.0044143.
- (59) Letko, M., Booiman, T., Kootstra, N., Simon, V. and Ooms, M. (2015) Identification of the HIV-1 Vif and human APOBEC3G protein interface. *Cell Rep.* **13**, 1789–1799.
- (60) Iwatani, Y., Chan, D. S., Liu, L., Yoshii, H., Shibata, J., Yamamoto, N., Levin, J. G., Gronenborn, A. M. and Sugiura, W. (2009) HIV-1 Vif-mediated ubiquitination/degradation of APOBEC3G involves four critical lysine residues in its C-terminal domain. *Proc. Natl. Acad. Sci. USA* **106**, 19539–19544.
- (61) de Bruijn, M. F. and Speck, N. A. (2004) Core-binding factors in hematopoiesis and immune function. *Oncogene* **23**, 4238–4248.
- (62) Jäger, S., Kim, D. Y., Hultquist, J. F., Shindo, K., LaRue, R. S., Kwon, E., Li, M., Anderson, B. D., Yen, L., Stanley, D., Mahon, C., Kane, J., Franks-Skiba, K., Cimermanic, P., Burlingame, A., Sali, A., Craik, C. S., Harris, R. S., Gross, J. D. and Krogan, N. J. (2011) Vif hijacks CBF- β to degrade APOBEC3G and promote HIV-1 infection. *Nature* **481**, 371–375.
- (63) Zhang, W., Du, J., Evans, S. L., Yu, Y. and Yu, X. F. (2011) T-cell differentiation factor CBF- β regulates HIV-1 Vif-mediated evasion of host restriction. *Nature* **481**, 376–379.
- (64) Kao, S., Khan, M. A., Miyagi, E., Plishka, R., Buckler-White, A. and Strebel, K. (2003) The human immunodeficiency virus type 1 Vif protein reduces intracellular expression and inhibits packaging of APOBEC3G (CEM15), a cellular inhibitor of virus infectivity. *J. Virol.* **77**, 11398–11407.
- (65) Stopak, K., de Noronha, C., Yonemoto, W. and Greene, W. C. (2003) HIV-1 Vif blocks the

- antiviral activity of APOBEC3G by impairing both its translation and intracellular stability. *Mol. Cell* **12**, 591–601.
- (66) Mercenne, G., Bernacchi, S., Richer, D., Bec, G., Henriët, S., Paillart, J. C. and Marquet, R. (2010) HIV-1 Vif binds to APOBEC3G mRNA and inhibits its translation. *Nucleic Acids Res.* **38**, 633–646.
- (67) Opi, S., Kao, S., Goila-Gaur, R., Khan, M. A., Miyagi, E., Takeuchi, H. and Strebel, K. (2007) Human immunodeficiency virus type 1 Vif inhibits packaging and antiviral activity of a degradation-resistant APOBEC3G variant. *J. Virol.* **81**, 8236–8246.
- (68) Schäfer, A., Bogerd, H. P. and Cullen, B. R. (2004) Specific packaging of APOBEC3G into HIV-1 virions is mediated by the nucleocapsid domain of the gag polyprotein precursor. *Virology* **328**, 163–168.
- (69) Burnett, A. and Spearman, P. (2007) APOBEC3G multimers are recruited to the plasma membrane for packaging into human immunodeficiency virus type 1 virus-like particles in an RNA-dependent process requiring the NC basic linker. *J. Virol.* **81**, 5000–5013.
- (70) Luo, K., Liu, B., Xiao, Z., Yu, Y., Yu, X., Gorelick, R. and Yu, X. F. (2004) Amino-terminal region of the human immunodeficiency virus type 1 nucleocapsid is required for human APOBEC3G packaging. *J. Virol.* **78**, 11841–11852.
- (71) Pery, E., Sheehy, A., Nebane, N. M., Brazier, A. J., Misra, V., Rajendran, K. S., Buhrlage, S. J., Mankowski, M. K., Rasmussen, L., White, E. L., Ptak, R. G. and Gabuzda, D. (2015) Identification of a novel HIV-1 inhibitor targeting Vif-dependent degradation of human APOBEC3G protein. *J. Biol. Chem.* **290**, 10504–10517.
- (72) Boehr, D. D., Dyson, H. J. and Wright, P. E. (2006) An NMR perspective on enzyme dynamics. *Chem. Rev.* **106**, 3055–3079.
- (73) Okuda, M. and Nishimura, Y. (2015) Real-time and simultaneous monitoring of the phosphorylation and enhanced interaction of p53 and XPC acidic domains with the TFIIH p62 subunit. *Oncogenesis* **4**, doi: 10.1038/oncsis.2015.13.
- (74) Kuwata, K., Shastry, R., Cheng, H., Hoshino, M., Batt, C. A., Goto, Y. and Roder, H. (2001) Structural and kinetic characterization of early folding events in beta-lactoglobulin. *Nat. Struct. Biol.* **8**, 151–155.
- (75) Fukuda, S., Nakanishi, Y., Chikayama, E., Ohno, H., Hino, T. and Kikuchi, J. (2009) Evaluation and characterization of bacterial metabolic dynamics with a novel profiling technique, real-time metabolotyping. *PLoS One* **4**, doi: 10.1371/journal.pone.0004893.
- (76) Nakakoshi, M., Ueda, M., Sakurai, S., Asakura, K., Utsumi, H., Miyata, O., Naito, T. and Takahashi, Y. (2007) Direct observation of the unstable intermediates in radical addition reaction by using an interfacing microchip combined with an NMR. *Magn. Reson. Chem.* **45**, 989–992.
- (77) Lu, X., Zhang, T., Xu, Z., Liu, S., Zhao, B., Lan, W., Wang, C., Ding, J. and Cao, C. (2015) Crystal structure of DNA cytidine deaminase ABOBEC3G catalytic deamination domain suggests a binding mode of full-length enzyme to single-stranded DNA.

J. Biol. Chem. **290**, 4010–4021.

- (78) Byeon, I. J., Ahn, J., Mitra, M., Byeon, C. H., Hercík, K., Hritz, J., Charlton, L. M., Levin, J. G. and Gronenborn, A. M. (2013) NMR structure of human restriction factor APOBEC3A reveals substrate binding and enzyme specificity. *Nat. Commun.* **4**, doi: 10.1038/ncomms2883.
- (79) Mitra, M., Hercík, K., Byeon, I. J., Ahn, J., Hill, S., Hinchee-Rodriguez, K., Singer, D., Byeon, C. H., Charlton, L. M., Nam, G., Heidecker, G., Gronenborn, A. M. and Levin, J. G. (2014) Structural determinants of human APOBEC3A enzymatic and nucleic acid binding properties. *Nucleic Acids Res.* **42**, 1095–1110.

Chapter 2

Development of a deamination assay involving the real-time NMR monitoring method

2-1. Introduction

The deamination activity of A3G has been evaluated in vitro by means of various methods such as measurement of the G to A mutation frequency in HIV-1-infected cells (1–5) and in the *Escherichia coli* Rif^r gene (6–8). The most commonly used deaminase assay is a gel-based biochemical deamination assay, i.e. the DNA glycosylase assay (1, 8–13). For this assay, substrate ssDNAs are mixed with A3G, and after the reaction, the products are collected and treated with uracil-DNA glycosylase (UDG), by which the resultant uracil bases are cleaved off. Subsequent alkaline treatment results in cleavage of the backbone through hydrolysis at the abasic (base lacking) sites. The final products can then be visualized by denaturing polyacrylamide gel electrophoresis (PAGE), the band intensities being quantified. This UDG-based assay requires multiple steps to finally detect deamination, however, it is useful in the sense that it is high-throughput.

Katahira's group have been utilizing the real-time NMR monitoring method to analyze the deamination reaction of A3G (14, 15). The advantage of our method over other biochemical methods is the ability to analyze the time course of a reaction at atomic resolution and the requirement of only a single step, owing to its non-destructive manner. In this chapter, we demonstrate and improve the methodology to estimate the deamination reaction by A3G_C using single CCC-containing ssDNA.

2-2. Results

2-2-1. Deaminase assay involving the real-time-NMR monitoring method

The experimental procedure for the method is illustrated in Figure 2-1A. We add A3G_C to the solution containing ssDNA and mix them thoroughly, which is defined as reaction time point zero. The final concentrations of A3G_C and substrate DNA are 2 μ M and 200 μ M, respectively. The mixed solution is immediately transferred to an NMR tube, which is loaded into the magnet. Subsequently, a series of 2D TOCSY spectra are obtained at different reaction time points. The conversion of ATTCCCGATT (s2) to ATTCCUGATT and then ATTCUUATT is monitored as an example in Figure 2-1. Each panel in Figure 2-1B shows the regions of 2D TOCSY spectra for H5–H6 correlation signals of cytosines and uracils. The three signals in the top left panel in the left column correspond to the C4, C5, and C6 residues of s2. The intensity of the C6 signal gradually decreases and it eventually disappears, which is coupled with the emergence of the U6 signal and a gradual increase in its intensity (Figure 2-1B, panels 1.8 h, 3.0 h, and 12 h). Then the C5 signal decreases and the U5 signal appears and increases (Figure 2-1B, panels 12 h, 72 h, and 187 h). Note that the signals not originating from deaminated residues were possibly perturbed by deamination.

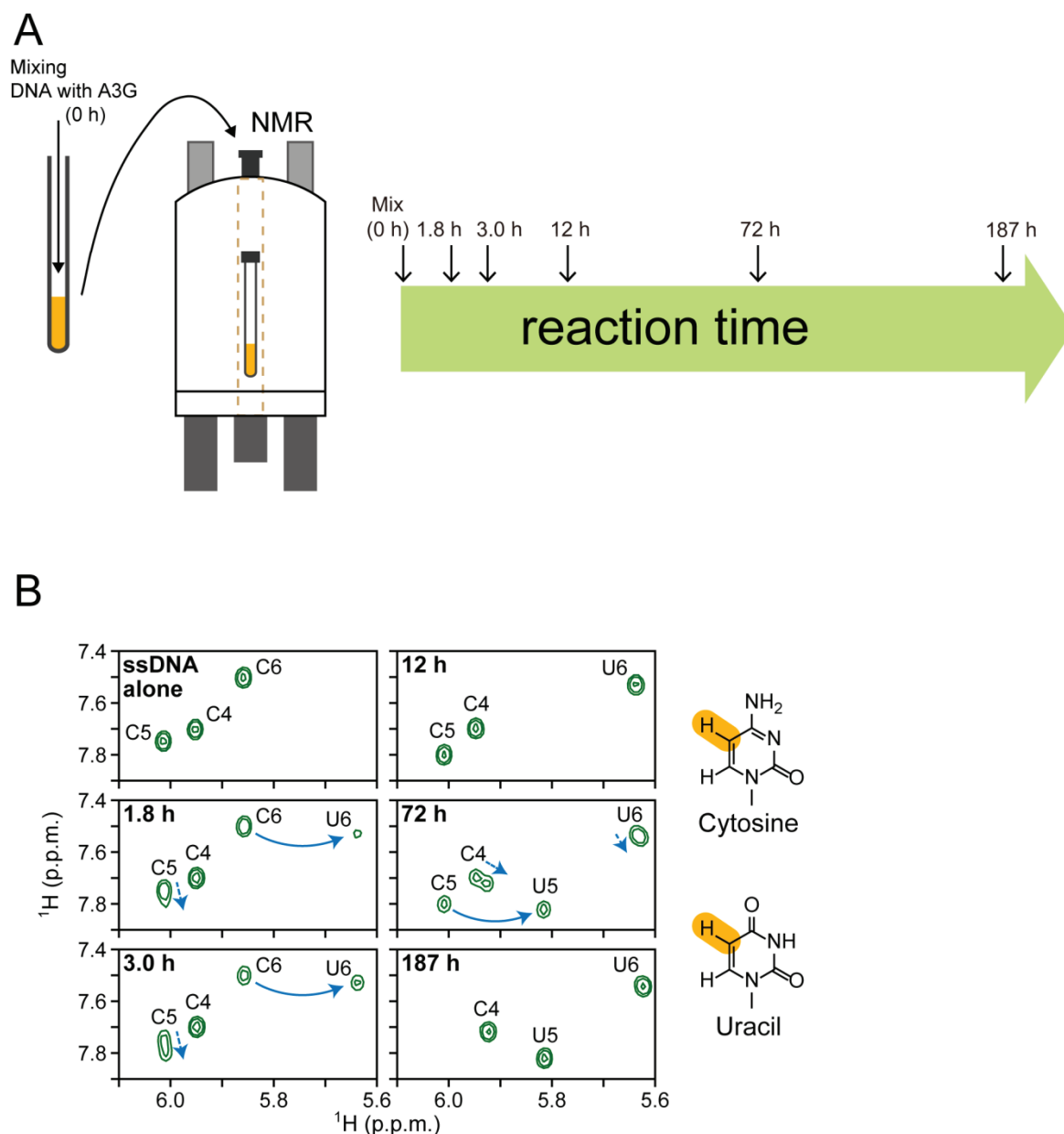


Figure 2-1.

Overview of real-time NMR monitoring.

(A) Illustration of the experimental procedure. The solutions containing A3G_C and ssDNA substrate were mixed (reaction time point 0 h), and then the mixture was transferred to an NMR tube.

(B) The NMR tube was then loaded into the magnet. A series of 2D TOCSY spectra was recorded at different reaction time points. Here, the reaction for ATTCCCGATT (s2) is monitored as an example. C to U conversion occurred at C6 then C5. The panels show the regions of 2D TOCSY spectra for H5–H6 correlation signals of cytosines and uracils at different reaction time points (1.8, 3.0, 12 h, 72 h and 187 h). The spectrum of "ssDNA (s2) alone" is shown for comparison. The structural formulae of cytosine and uracil are shown (protons H5 and H6 are highlighted in yellow). The signals for C6 and C5 are connected with those for U6 or U5 by solid arrows, respectively. The signals for C4, C5 and U6 are perturbed (broken arrows) due to the C to U conversion.

We then determine the intensities of these signals and follow the time courses of their changes. In the current case, the signal intensity at each time point is plotted in Figure 2-2A, B. The plots for C6 and U6 were each fitted to a single exponential function, equations (2.2) and (2.1) in the Experimental Procedure section, by which the rate constant, k , was found to be $0.28 \pm 0.01 \text{ h}^{-1}$ and $0.27 \pm 0.02 \text{ h}^{-1}$, respectively. This k value can be used as an index of activity. The time course of the C5 signal after conversion of C6 into U6 is plotted in Figure 2-2B, which shows the maximum. The increase in the initial phase in Figure 2-2B is due to deamination of C6 to U6, and the decrease in that in the later phase are indicated as C5 to U6 conversion. Then, the curvature was fitted by the double exponential function shown in equation (2-3), k values (k_1 , $0.29 \pm 0.01 \text{ h}^{-1}$, and k_2 , $0.0102 \pm 0.0010 \text{ h}^{-1}$) being calculated. This k_1 value corresponds to k calculated from the plot in Figure 2-2A. Additionally, the deamination rate obtained from the U5 increase fitted by equation (2-2) was $0.00095 \pm 0.0010 \text{ h}^{-1}$, which corresponds to k_2 . These findings indicate that the deaminase activity of A3G_C against CCC is 30 times higher than that against CC. Enzymatic activity, depending on specific activity, may be different for each purification lot. In order to ignore the lot difference, activity was normalized as a ratio, taking the rate of deamination of C6 as 1.0. Figure 2-2C is a plot of relative activity. In this study, relative activity is used as an index of deamination activity.

2-3. Discussion

2-3-1. Enzymatic properties of A3G_C as to CCC hotspots

In the Results section, we demonstrated that our real-time NMR monitoring method is a powerful tool for evaluating the deamination reaction catalyzed by A3G_C of atomic resolution. A3G exhibits sequential deamination activity that converts CCC into CCU and then CUU, which may be the basis of the difference in the enzymatic preferences for the sequences (CCC and CC). It was interesting that the time course of the deamination reaction can be used to evaluate a single exponential function as a chemical reaction even if the substrate concentration is much higher compared to that of A3G_C. These findings mean that substrates are saturated by A3G_C or that the active site of A3G_C is unsaturated even under substrate-excess conditions. Harjes and coworkers reported that they were not able to saturate the enzyme although they used very high substrate concentrations of up to 10 mM, about 1,000 times that of A3G_C, the initial speed of the reaction increasing linearly with substrate concentration (16). This finding suggested that either the reaction cannot be described by a Michaelis-Menten-type mechanism or saturation of the enzyme occurs at a substrate concentration much higher than 1,000-fold A3G_C (16). The rate constant for ssDNA at pH 7.4 without the enzyme is about $1 \times 10^{-10} \text{ s}^{-1}$, (ca. $4 \times 10^{-7} \text{ h}^{-1}$) (17). This finding means A3G_C accelerate deamination reaction about 7×10^5 times, being not so effective as the most effective group of enzymes (18). Furthermore, compared to other APOBEC family protein, A3G_C in Furukawa *et al.* (14), the same construct as used in this study, was about 1,000-fold less active than the same amount of APOBEC3A (19). These findings indicate that the active site of A3G_C may have the ability to bind to much more substrate CCC. Anyway, we demonstrated that the deaminase

activity of A3G_C can be evaluated as a single or double exponential function.

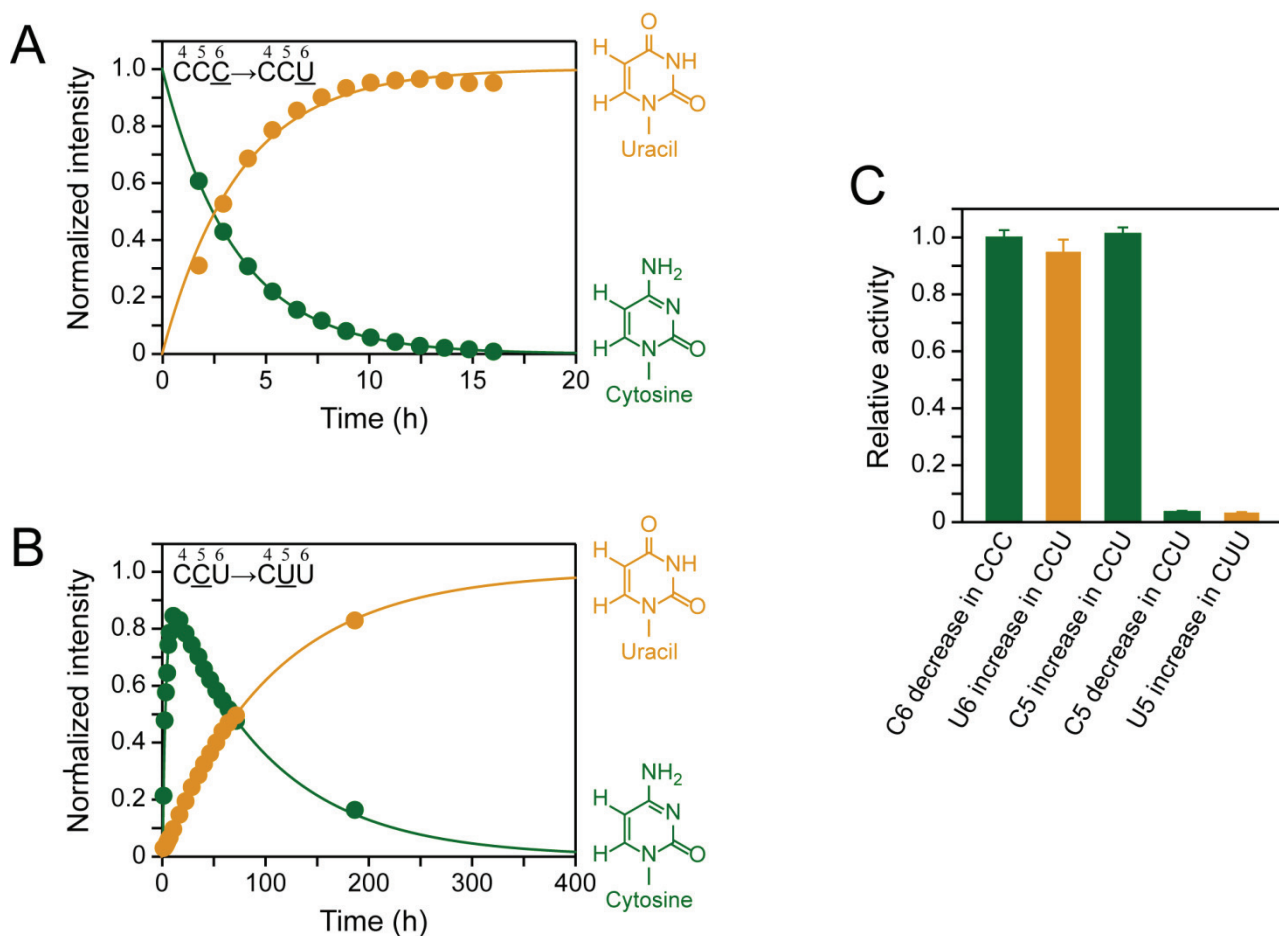


Figure 2-2.

Deamination activity of A3G against 10-mer ssDNA with a single hotspot.

(A) The intensities of H5–H6 correlation signals for C6 and U6 at different reaction time points are plotted. (B) The intensities of H5–H6 correlation signals for C5 and U5 at different reaction time points are plotted as in (A). (C) The data were fitted to a single or double exponential function, by which the rate constant, k , was obtained. The obtained k s were normalized as a ratio, taking the rate of deamination of C6 as 1.0.

2-4. Conclusion

In this chapter, we presented methodology for real-time NMR spectroscopy in detail, and described the ease of its usage and its applicability to systems exhibiting different temporal and spatial resolution.

2-5. Experimental Procedures

Preparation of wild-type APOBEC3G

Recombinant A3G_C (residues 193–384) was expressed and purified as described in Furukawa *et al.* with minor modification (14). The pET15b-A3G_C expression plasmid, kindly provided by Dr. Nagata (Kyoto University), was used to express N-terminal hexahistidine (His6)-affinity tagged A3G_C (14). *Escherichia coli* BL21(DE3/RIL) cells (Stratagene) were transformed with the vector and grown in LB or TB medium at 37°C until OD 600<0.6, and expression was induced by adding 1 mM IPTG for 4 hours. Cells were harvested by centrifugation (3,000 g) for 20 min at 4°C. The cells containing A3G_C were lysed by sonication in a buffer comprising 20 mM Tris-HCl (pH 7.5) and 500 mM NaCl. The supernatant was collected by centrifugation (12,000 g) for 20 min at 4°C. The soluble fraction was loaded onto a Ni Sepharose resin column (GE Healthcare), equilibrated with a solution comprising 20 mM Tris-HCl (pH 7.5), 500 mM NaCl and 30 mM imidazole. A3G_C was eluted with a linear imidazole gradient, 30–500 mM. The protein solution was dialyzed three times against a buffer comprising 20 mM Tris-HCl (pH 7.5), 500 mM NaCl, 10 μ M ZnCl₂ and 5 mM DTT at 4°C. The His6 tag was cleaved with thrombin protease (GE Healthcare) in 20 mM Tris-HCl (pH 7.5), 500 mM NaCl, 5 mM MgCl₂, 10 μ M ZnCl₂ and 5 mM DTT at room temperature overnight. After removal of thrombin with Benzamidine Sepharose resin (GE Healthcare), the solution was loaded onto a Ni Sepharose resin column equilibrated with a solution comprising 20 mM Tris-HCl (pH 7.5), 500 mM NaCl, 10 μ M ZnCl₂, 5 mM DTT and 30 mM imidazole. The flow-through fraction was dialyzed more than three times against the storage buffer comprising 20 mM Tris-HCl (pH 7.5), 30 mM NaCl, 10 μ M ZnCl₂ and 5 mM DTT. Finally, the obtained A3G_C was passed through a syringe filter with a 0.45 μ m pore size PES membrane, and then stored until use at 4°C.

Preparation of substrate DNA

The substrate ssDNA, synthesized with a DNA synthesizer and purified by high performance liquid chromatography (HPLC), was purchased from Fasmac Co., Ltd.

NMR spectroscopy

All NMR spectra were recorded at 25°C on a Bruker DRX600 spectrometer equipped with a cryogenic probe and a Z-gradient (Bruker Biospin). Samples were prepared by mixing A3G_C in storage buffer with measurement buffer. The measurement buffer was prepared as storage buffer in the Preparation of wild-type APOBEC3G section, except that 5% deuterium oxide was added. The concentration of A3G_C was set at 2.0 μ M, while that of the ssDNA was adjusted to 200 μ M. After mixing A3G_C and ssDNA, a series of two-dimensional 2D TOCSY spectra with a mixing time of 20 ms were recorded at different reaction time points (ca. 75 min to obtain a single 2D TOCSY spectrum), by which the deamination of cytosines of ssDNA could be monitored in real-time. Resonance assignments of the substrate ssDNAs were made using 2D TOCSY with the aid of 2D

^1H - ^{13}C HSQC spectroscopy and also by comparing the chemical shifts with the assignments for ATTCCAATT previously reported (14). Water signal suppression was achieved with 3-9-19 watergate, and echo-antiecho pulse schemes in each 2D TOCSY and 2D ^1H - ^{13}C HSQC measurement (14, 15). Spectra were processed with NMRPipe (20), and analyzed using SPARKY (21).

The intensities of the peaks of either cytosine or uracil were plotted against time analyzed as follows. The data were fitted to a single or double exponential function, using the following equations, by which apparent deamination rate constants were derived:

$$I(t)_{\text{C6}} = I_s \exp(-kt) + I_b \quad (2.1)$$

$$I(t)_{\text{U6 or U5}} = I_s(1 - \exp(-kt)) + I_b \quad (2.2)$$

$$I(t)_{\text{C5}} = -I_{s1} \exp(-k_1 t) + I_{s2} \exp(-k_2 t) + I_b \quad (2.3)$$

, where I_b corresponds to the baseline intensity of the spectrum, I_s the maximum signal intensity of each phase, and k the apparent deamination rate constant, which was obtained by least-squares analysis. The k value was used as an index of deaminase activity. The error was estimated as follows. Firstly, the error of the NMR signal intensity was obtained from the base level noise of each NMR spectrum. Then the error of the rate constant was calculated from data sets constructed by Monte Carlo simulation using the error of the NMR signal intensity. Finally, the errors of the relative activity were obtained by means of error propagation calculation.

2-6. References

- (1) Harris, R. S., Bishop, K. N., Sheehy, A. M., Craig, H. M., Petersen-Mahrt, S. K., Watt, I. N., Neuberger, M. S. and Malim, M. H. (2003) DNA deamination mediates innate immunity to retroviral infection. *Cell* **113**, 803–809.
- (2) Zhang, H., Yang, B., Pomerantz, R. J., Zhang, C., Arunachalam, S. C. and Gao, L. (2003) The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. *Nature* **424**, 94–98.
- (3) Lecossier, D., Bouchonnet, F., Clavel, F. and Hance, A.J. (2003). Hypermutation of HIV-1 DNA in the absence of the Vif protein. *Science* **300**, 1112.
- (4) Browne, E. P., Allers, C. and Landau, N. R. (2009). Restriction of HIV-1 by APOBEC3G is cytidine deaminase-dependent. *Virology* **387**, 313–321.
- (5) Kobayashi, T., Koizumi, Y., Takeuchi, J. S., Misawa, N., Kimura, Y., Morita, S., Aihara, K., Koyanagi, Y., Iwami, S. and Sato, K. (2014) Quantification of deaminase activity-dependent and -independent restriction of HIV-1 replication mediated by APOBEC3F and APOBEC3G through experimental-mathematical investigation. *J. Virol.* **88**, 5881–5887.
- (6) Haché, G., Liddament, M. T. and Harris, R. S. (2005) The retroviral hypermutation specificity of APOBEC3F and APOBEC3G is governed by the C-terminal DNA cytosine deaminase domain. *J. Biol. Chem.* **280**, 10920–10924.
- (7) Chen, K. M., Harjes, E., Gross, P. J., Fahmy, A., Lu, Y., Shindo, K., Harris, R. S. and Matsuo, H. (2008) Structure of the DNA deaminase domain of the HIV-1 restriction factor APOBEC3G.

Nature **452**, 116–119.

- (8) Shindo, K., Li, M., Gross, P. J., Brown, W. L., Harjes, E., Lu, Y., Matsuo, H. and Harris, R. S. (2012) A comparison of two single-stranded DNA binding models by mutational analysis of APOBEC3G. *Biology (Basel)* **1**, 260–276.
- (9) Yu, Q., König R., Pillai, S., Chiles, K., Kearney, M., Palmer, S., Richman, D., Coffin, J. M. and Landau, N. R. (2004) Single-strand specificity of APOBEC3G accounts for minus-strand deamination of the HIV genome. *Nat. Struct. Mol. Biol.* **11**, 435–442.
- (10) Navarro, F., Bollman, B., Chen, H., König, R., Yu, Q., Chiles, K. and Landau, N. R. (2005) Complementary function of the two catalytic domains of APOBEC3G. *Virology* **333**, 374–386.
- (11) Chelico, L., Pham, P., Calabrese, P. and Goodman, M. F. (2006) APOBEC3G DNA deaminase acts processively 3' → 5' on single-stranded DNA. *Nat. Struct. Mol. Biol.* **13**, 392–399.
- (12) Iwatani, Y., Takeuchi, H., Strebel, K. and Levin, J. G. (2006) Biochemical activities of highly purified, catalytically active human APOBEC3G: correlation with antiviral effect. *J. Virol.* **80**, 5992–6002.
- (13) Holden, L. G., Prochnow, C., Chang, Y. P., Bransteitter, R., Chelico, L., Sen, U., Stevens, R. C., Goodman, M. F. and Chen, X. S. (2008) Crystal structure of the anti-viral APOBEC3G catalytic domain and functional implications. *Nature* **456**, 121–124.
- (14) Furukawa, A., Nagata, T., Matsugami, A., Habu, Y., Sugiyama, R., Hayashi, F., Kobayashi, N., Yokoyama, S., Takaku, H. and Katahira, M. (2009) Structure, interaction and real-time monitoring of the enzymatic reaction of wild-type APOBEC3G. *EMBO J.* **28**, 440–451.
- (15) Furukawa, A., Sugase, K., Morishita, R., Nagata, T., Kodaki, T., Takaori-Kondo, A., Ryo, A., and Katahira, M. (2014) Quantitative analysis of location- and sequence-dependent deamination by APOBEC3G using real-time NMR spectroscopy. *Angew. Chem. Int. Ed. Engl.* **53**, 2349–2352.
- (16) Harjes, S., Solomon, W. C., Li, M., Chen, K. M., Harjes, E., Harris, R. S. and Matsuo, H. (2013) Impact of H216 on the DNA binding and catalytic activities of the HIV restriction factor APOBEC3G. *J. Virol.* **87**, 7008–7014.
- (17) Frederico, L. A., Kunkel, T. A. and Shaw, B. R. (1990) A sensitive genetic assay for the detection of cytosine deamination: determination of rate constants and the activation energy. *Biochemistry* **29**, 2532–2537.
- (18) Radzicka, A. and Wolfenden, R. (1995) A proficient enzyme. *Science* **267**, 90–93.
- (19) Byeon, I. J., Ahn, J., Mitra, M., Byeon, C. H., Hercík, K., Hritz, J., Charlton, L. M., Levin, J. G. and Gronenborn, A. M. (2013) NMR structure of human restriction factor APOBEC3A reveals substrate binding and enzyme specificity. *Nat. Commun.* **4**, doi: 10.1038/ncomms2883.
- (20) Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293.
- (21) Goddard, T. D. and Kneller, D. G. (2006) SPARKY 3. *San Francisco: University of California.*

Chapter 3

Nucleic acid determinant for deamination by A3G_C

3-1. Introduction

Previously, it was reported the sequence of ssDNA outside of CCC hot spot affects deaminase activity of A3G. Several groups reported CCCA/G are much favored than CCCT or CCCC by A3G (1, 2). Chelico and coworkers found sequence context within nine nucleotides CCC hot spot is positioned in center affects deaminase activity. The most favored target in 13 represented sequences, AAACCCAAA or ATACCCAAA, are approximately 40 folds stronger than the last one, TTTCCCTTT (2). Substrate-binding features and the deamination specificity of A3G were analyzed by means of so-called nucleotide analog interference mapping (3). With this method, ssDNA substrates containing single or multiple nucleoside analogs, that have a modification in the base, sugar or phosphate moiety, are used for the UDG-based assay (3). Introduction of 2' -O-methyl modification, methyl phosphotriester modification, or an abasic site at a single position in the ssDNA revealed that the sugar and phosphate moieties of the nucleotides within the CCC hotspot and its vicinity are important for substrate recognition by A3G (3). Moreover, introduction of pyrimidine base analogs at positions adjacent to the deamination target site showed that A3G dictates the exocyclic groups in pyrimidines 1–2 nt 5' to the target cytosine, therefore the authors concluded that the base moieties of these nucleotides are critical for A3G to recognize the target cytosine (3).

Additionally, we wondered if this substrate recognition mechanism corresponds to sliding or not. Furukawa and coworker previously utilized real-time NMR spectroscopy to monitor the deamination reaction, with which one can detect the reaction directly in a site-specific manner with high temporal and spatial resolution (4, 5). The advantage of this method is that it is sensitive as to weak interactions because highly concentrated ssDNA (100 μ M order) can be used. The real-time NMR method first revealed that A3G deaminates the third cytosine of a CCC hotspot much faster than the second one (CCC) (4). In the subsequent study, the deamination reactions of ssDNA containing two CCC hotspots were monitored and it was revealed that full length A3G as well as A3G_C can deaminate the two hotspots with 3'→5' polarity (5). Furthermore, by construction of a kinetic model, in which nonspecific protein-ssDNA binding and sliding processes are incorporated, the authors quantitatively analyzed the 3'→5' polarity of deamination by A3G_C (5). It was revealed that the 3'→5' polarity of A3G_C can be rationally explained by introducing the sliding direction-dependent catalytic rate (5). The analysis provided the values for various kinetic parameters, and importantly the catalytic rate of A3G_C was shown to be higher (ca. five-fold) when A3G_C approaches the target cytosine in the 3' to 5' direction rather than in the opposite direction (5).

Here, we applied the real-time NMR method to analyze the catalytic activity of A3G_C toward ssDNA oligonucleotides containing a nucleotide analog at a single or multiple positions to identify nucleic acid determinants for deamination. Then, long ssDNA oligonucleotides, each

containing two CCC hotspots connected by a long linker sequence of either deoxyribonucleotides, ribonucleotides or abasic deoxyribonucleotides were used as substrates to identify the chemical moieties of nucleotides that play a key role in the sliding of A3G_C. To confirm the importance of the identified chemical moieties, the dependency of the 3'→5' polarity on the NaCl concentration was examined.

3-2. Results

3-2-1. The effect of ribonucleotide substitution at a single position of a substrate ssDNA on A3G_C's deaminase activity

The real-time NMR method was applied to monitor the deaminase activity of A3G_C toward a series of DNA nucleotides, each containing ribonucleotide substitution (more specifically, 2'-OH substitution) at a single position. The standard substrate was a 10-mer DNA nucleotide, AAACCCGAAA (s3-1_DNA in Figure 3-1A), the corresponding ribonucleotides at positions 2 to 9 in the other substrates each being substituted. Figure 3-1B shows examples of signal assignment using ^1H - ^{13}C HSQC (H5-C5 correlation signals) spectra for three substrates, s3-1_DNA, s3-1_R4 and s3-1_R6. The three green signals in each spectrum in Figure 3-1B, are those of C4, C5, and C6 of the CCC hotspots of the substrates before the addition of A3G_C. Several hours after the addition of A3G_C (see legend to Figure 3-1B), the spectra of the different substrates showed different outcomes. In the cases of s3-1_DNA and s3-1_R4, along with conversion of the target cytosine to uracil (C6 to U6), the signal of C6 disappeared and decreased, respectively, and the U6 signal appeared. On the other hand, for s3-1_R6, in which the target cytosine is substituted by a ribonucleotide, no signal change was evident, indicating that deamination failed. Thus, it was clearly shown that the real-time NMR method can distinguish the reactivity for the different substrates. For detailed analysis with the real-time NMR method, we chased the intensity change of the H5-H6 TOCSY correlation signal with the targeted cytosine, C6, in each substrate. Figure 3-1C demonstrates the TOCSY spectra of about 2 h (green) and 36 h (yellow) after mixing obtained for s3-1_DNA and its analog (s3-1_R3 to s3-1_R7 in Figure 3-1A). The time course of the C6 intensity of Figure 3-1C are plotted in Figure 3-2. It is clear that the decrease in the C6 intensity of four chimeric substrates shown in Figure 3-2 are slower than that of DNA. These results indicates that ribonucleotide substitution (2'-OH substitution) at the third to seventh position, that corresponds to five sequential nucleotides CCC hotspot positioned in center, negatively affects A3G_C's activity. We performed this analysis for the other substrates listed in Figure 3-3A, (s3-1_R2, s3-1_R8, and s3-1_R9, and a deamination rate constant, which is used as an index of activity, was calculated. Similar to Figure 3-1, the TOCSY spectra of three DNA analog shown in Figure 3-3A are also demonstrated in Figure 3-3B and time course of C6 of those chimeric DNA are also plotted and fitted against one exponential function to calculate the rate constant of the reaction (Figure 3-3 (C)).

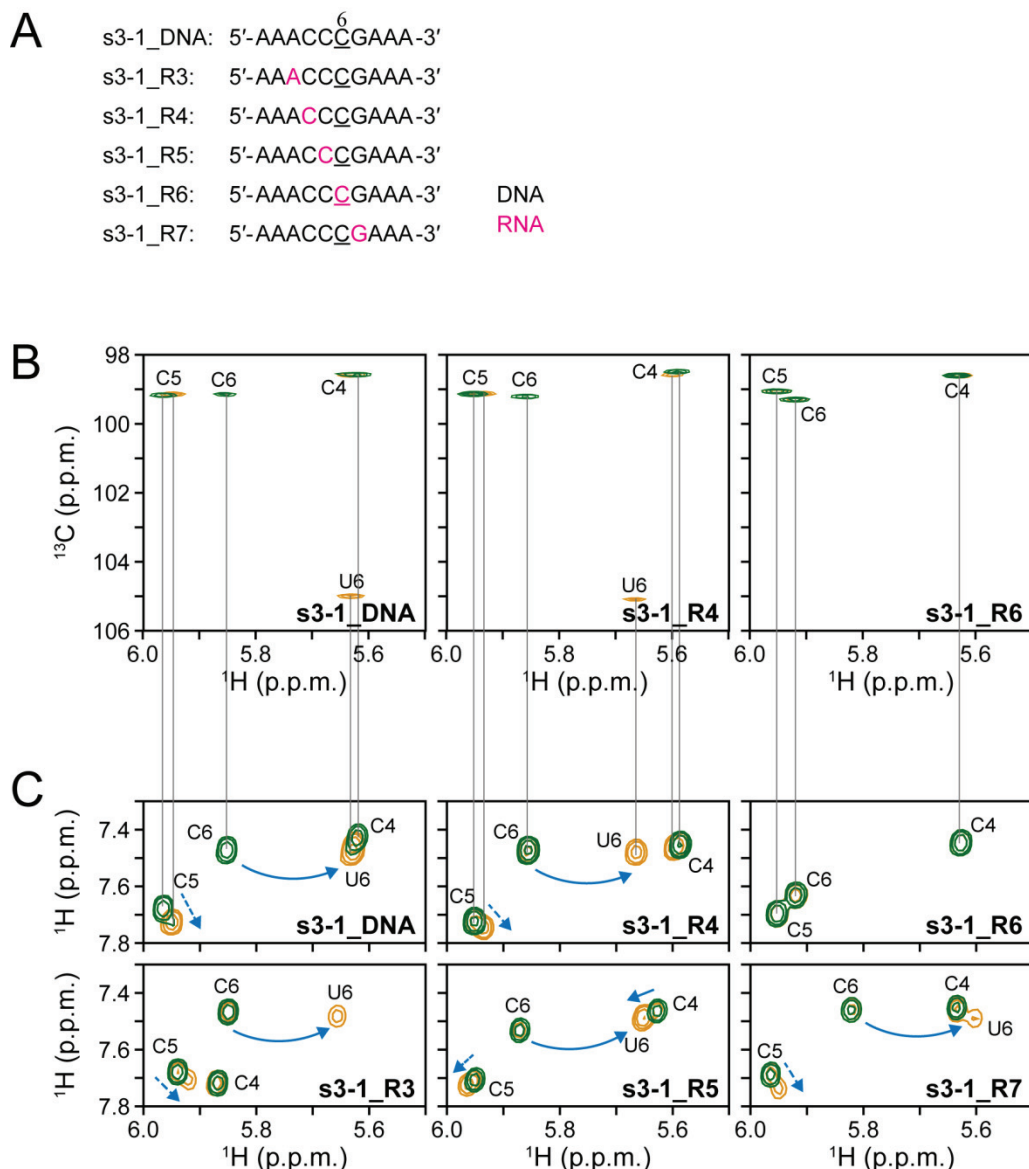


Figure 3-1.

Examples of deamination assay for chimeric DNA containing a single DNA→RNA substitution.

(A) The sequences of chimeric nucleotides. Deoxyribonucleotide and ribonucleotide are colored black and magenta, respectively. (B, C) Chemical shift assignment of chimeric DNAs shown in (A). ^1H - ^{13}C HSQC spectra (B) and TOCSY spectra (C, up) of each ssDNA substrate, AAACCCGAAA (s3-1_DNA), AA^ACCCGAAA (s3-1_R4), and AAACCC^GGAAA (s3-1_R6), in which C to U conversion occurs at underscored C6. The signals originated from the same molecule are connected by gray line. Green signals in ^1H - ^{13}C HSQC spectra present before addition of A3G_C, while yellow signals present 107 hours (s3-1_DNA), 74 hours (s3-1_R4), and 90 hours (s3-1_R6) after addition of A3G_C. Green and yellow signals in TOCSY spectra each present about 2 and 36 hours after addition of A3G_C. Signals of C6 and U6 are connected by a solid arrow, while a broken arrow indicates the signal perturbation of C5 due to the C6 to U6 conversion.

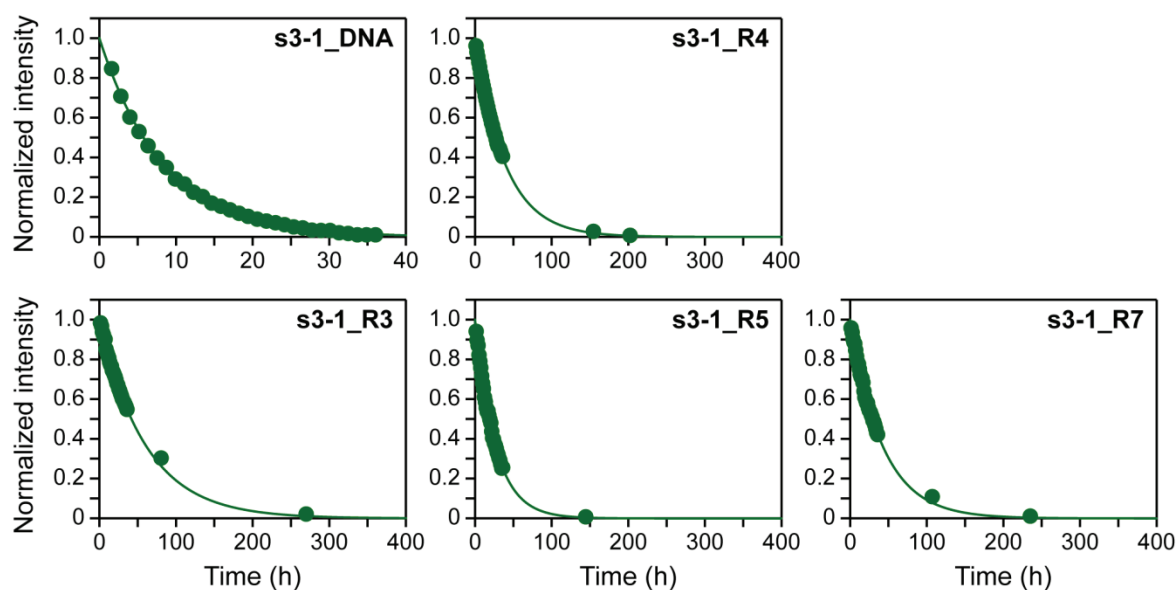


Figure 3-2.

Deamination activity against chimeric DNA containing a single DNA→RNA substitution within the central five sequential nucleotides.

The time course of the intensity change of C6 in TOCSY spectra shown in Figure 3-1(C) were monitored.

To ignore the difference of specific activity of two A3G lots, relative activity, relative k value, was also calculated for further quantitative analysis as shown in Chapter 2. Figure 3-4 compares the relative activities of A3G_C toward ssDNA substrates, each having a single ribonucleotide substitution at different positions. The relative activity of A3G_C toward RNA substrate (s3-1_RNA) is also included as a negative control. A3G_C failed to deaminate this substrate as well as s3-1_R6, the latter being mentioned above, and in both cases a 2'-OH substitution was introduced at the targeted cytosine. Figure 3-4 indicates that the negative effect of a 2'-OH substitution on reactivity is great when the substitution is introduced within a CCC hotspot, and also at the positions immediately flanking the 5' and 3' of the CCC hotspot, i.e. at the 5 nucleotide recognition sequence positioning the CCC hotspot at the center.

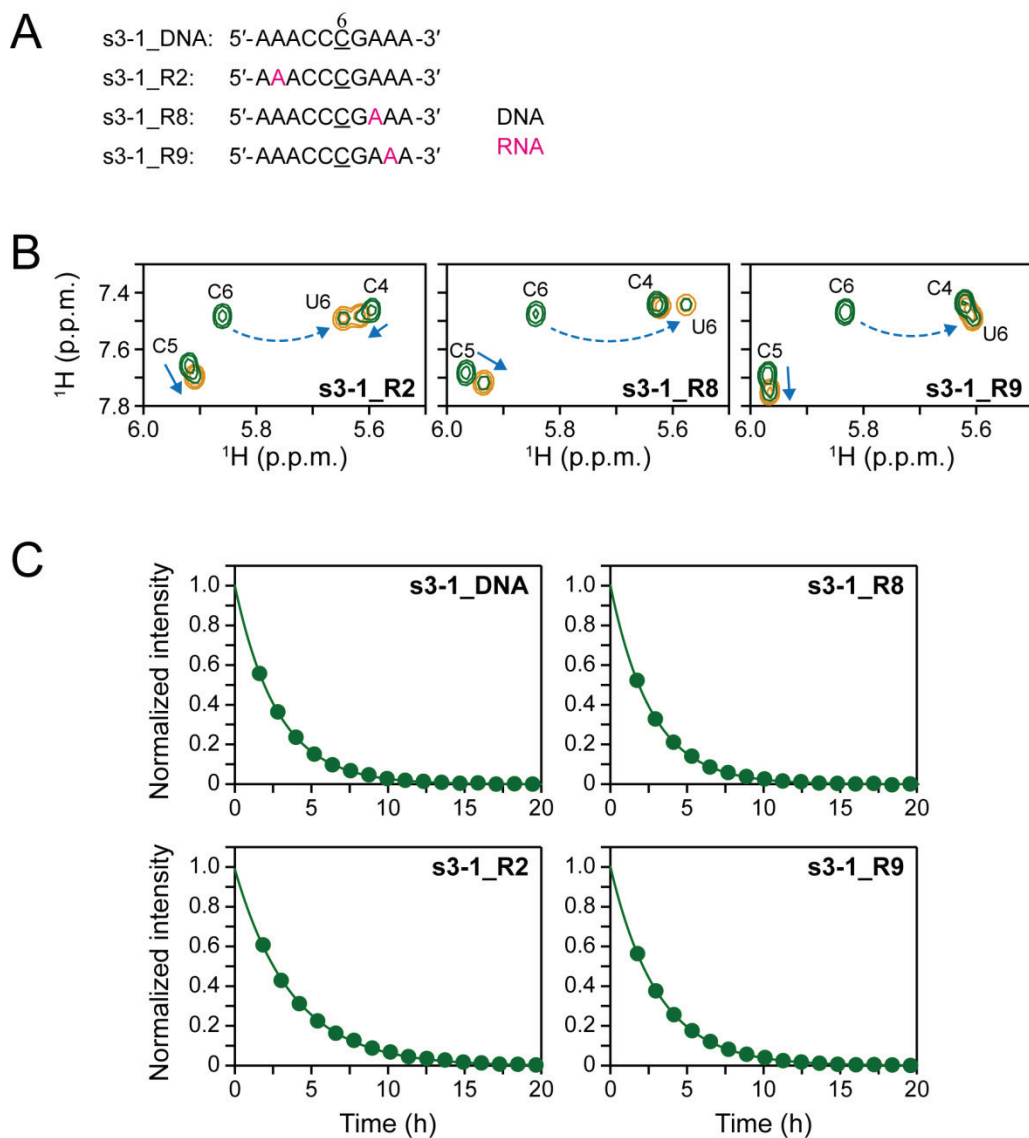


Figure 3-3.

Deamination activity against chimeric DNA containing a single DNA→RNA substitution outside of the central five sequential nucleotides.

(A) The sequences of chimeric nucleotides. Deoxyribonucleotide and ribonucleotide are colored black and magenta, respectively. (B) Chemical shift assignment of TOCSY spectra of each ssDNA substrate in (A). Green and yellow signals in TOCSY spectra each present about 2 and 36 hours after addition of A3G_C. Signals of C6 and U6 are connected by a solid arrow, while a broken arrow indicates the signal perturbation of C5 due to the C6 to U6 conversion. (C) The time course of the intensity change of C6 in TOCSY spectra shown in (B) was monitored.

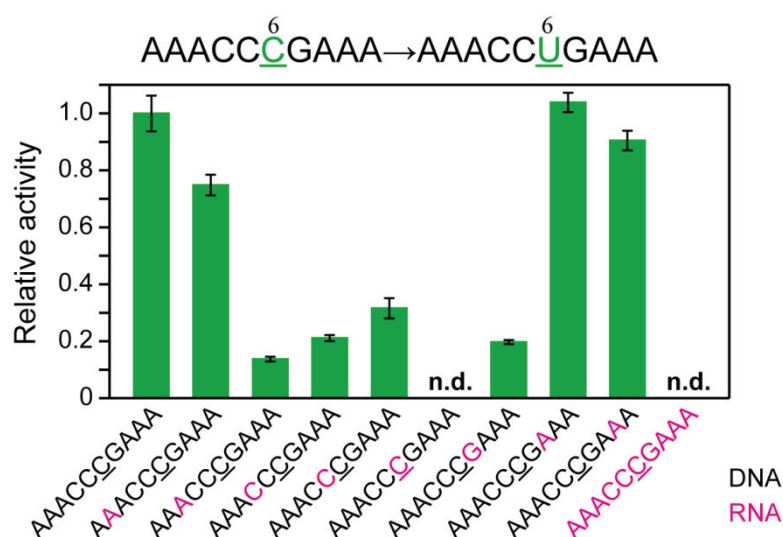


Figure 3-4.

Deamination activity against chimeric DNA containing a single DNA→RNA substitution.

The activity against 10-mer DNA (black) and RNA (magenta) are shown as control. Relative activity was defined as the ratio of the apparent deamination rate constant (k) obtained from the time course of C6 signal intensity of each chimeric nucleotide to that obtained for DNA substrate.

3-2-2. The effect of 2'-OH substitution or abasic substitution in the flanking regions of a 5 nucleotide recognition sequence on deaminase activity

In the previous section we showed that a 5 nucleotide recognition sequence, containing a CCC hotspot at its center, is important for A3G_C to exert its full deaminase activity. Here, we attempted to determine whether or not the regions flanking the 5' and 3' to this 5 nucleotide recognition sequence play any role in the reactivity of the substrates. A series of 10-mer DNA nucleotides, each carrying either a ribonucleotide or abasic deoxyribonucleotide in the 5'- and 3'-flanking regions, was subjected to real-time monitoring.

As shown in Figure 3-5, the deaminase activities of A3G_C toward ssDNA substrates with multiple ribonucleotide substitutions were quantified. s3-1_D3-6/R, residues 3–6 being deoxyribonucleotides and the other residues being ribonucleotides, has a ribonucleotide substitution at residue G7 that is one of the 5 nucleotide recognition sequence and was indicated to be critical for deamination in the previous section. s3-1_D4-7/R (Figure 3-5A) also has a substitution in the 5 nucleotide recognition sequence at residue A3. The activity toward these substrates was extremely low (Figure 3-5B, C), which is consistent with the results in Figure 3-4. Surprisingly, although the deoxyribonucleotides are retained in the 5 nucleotide recognition sequence and ribonucleotide substitutions are introduced only outside of the 5 nucleotide recognition sequence, the reactivity of the substrate s3-1_D3-7/R turned out to be rather low. This indicates that the sugar moiety of the nucleotides in the regions outside the 5 nucleotide recognition sequence also affects the deamination

by A3G_C, which was not apparent in the single substitution experiments shown in Figure 3-4. It is supposed that the effect of a substitution outside the 5 nucleotide recognition sequence on the activity is small but that the effect becomes evident when such a substitution is introduced at multiple positions. The substrates s3-1_D2-7/R and s3-1_D3-8/R each retain deoxyribonucleotides in the 5 nucleotide recognition sequence, and additionally at the 5'- and 3'- flanking positions of the 5 nucleotide recognition sequence, respectively. The reactivity of s3-1_D2-7/R was higher than that of s3-1_D3-8/R. The further interpretation of these findings are discussed in the Discussion section.

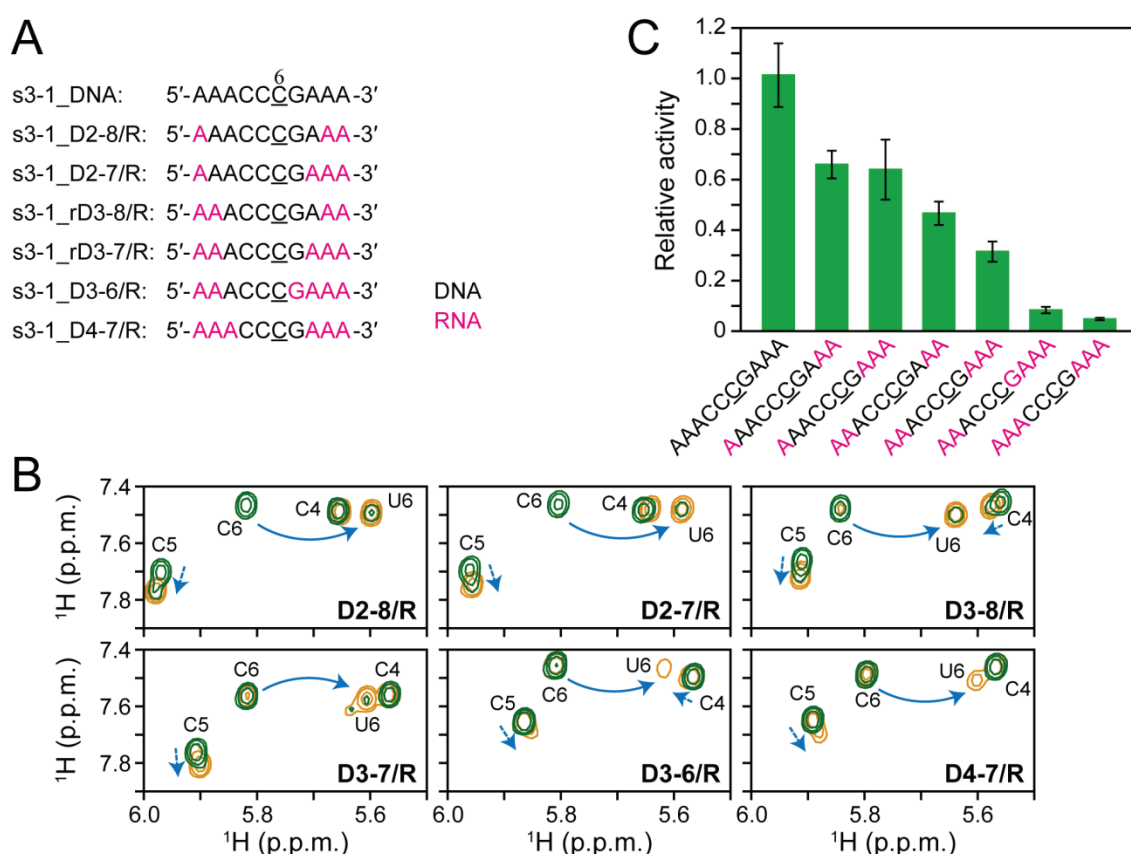


Figure 3-5.

Real-time monitoring of deamination reactions for substrates containing multiple ribonucleotides.

(A) The sequences of chimeric nucleotides. (B) Chemical shift assignment of TOCSY spectra of each ssDNA substrate in (A). Green and yellow signals in TOCSY spectra each present about 2 and 36 hours after addition of A3G_C. Signals of C6 and U6 are connected by a solid arrow, while a broken arrow indicates the signal perturbation of C5 due to the C6 to U6 conversion. (C) Relative activity was calculated by dividing the deamination rate constant for a substrates containing ribonucleotides substitution.

The deaminase activities of A3G_C toward ssDNA substrates that contain multiple abasic deoxyribonucleotides in the 5'- and 3'-flanking regions were quantified, as shown in Figure 3-6. The substrates s3-1_D3-6/X and s3-1_D4-7/X exhibited extremely low reactivity (Figure 3-6B, C). This indicates that not only the sugar moiety but also the base moiety of the 5 nucleotide recognition sequence is critical for the reactivity. The substrate s3-1_D3-7/X retains deoxyribonucleotides in the 5 nucleotide recognition sequence and abasic substitutions are introduced outside of the 5 nucleotide recognition sequence. This substrate also showed a great reduction of reactivity, which indicates that the base moiety of the nucleotides in the regions outside the 5 nucleotide recognition sequence also affects the deamination reaction of A3G_C. Each of the substrates s3-1_D2-7/X and s3-1_D3-8/X retains bases in the 5 nucleotide recognition sequence, and additionally at the 5'- and 3'- flanking positions of the 5 nucleotide recognition sequence, respectively. The reactivity of s3-1_D2-7/X was lower than that of s3-1_D3-8/X and was as low as that of s3-1_D3-7/X. The further interpretation of these findings are discussed in the Discussion section.

3-2-3. The effect of 2'-OH substitution or abasic substitution at the linker region connecting two CCC hotspots on the 3'→5' polarity

In the former study, we used an ssDNA substrate containing two CCC hotspots, connected by a long linker region, to monitor the deamination reaction, and revealed that A3G_C can deaminate the two hotspots with 3'→5' polarity (5). It was also revealed that the 3'→5' polarity is caused by sliding, because the polarity disappeared when sliding was inhibited through the formation of a DNA duplex at a linker region (5). Here, we substituted the linker region of an ssDNA substrate with a poly-ribonucleotide or poly-abasic deoxyribonucleotide and then subjected it to real-time monitoring. Thus, we attempted to reveal the effects of these substitutions on the 3'→5' polarity of A3G_C's deamination activity.

Figure 3-7A shows the design of chimeric DNAs containing two hotspots connected by linker sequence with poly- adenine deoxynucleotide tail, one is control DNA comprise poly- adenine deoxynucleotide and linker portion of the other is replaced to poly- adenine ribonucleotide or abasic deoxyribonucleotide, respectively. Figure 3-7B shows an example of real-time NMR monitoring using TOCSY spectra. Presented in green are the H5–H6 TOCSY correlation signals of the residues within the 5'-CCC hotspot (C4, C5, and C6) and 3'-hotspot (C32, C33, and C34) A3G_C at the ca. 2 hours (initial) time point. Upon C to U conversion at positions C6 and C34, all the signals were perturbed. The signals at the 36 hours time point are shown as an example of the perturbation (yellow). It can be seen clearly that the signals of C6 and C34 are separate from other signals, and therefore changes in intensity can be followed unambiguously. Thus, we plotted the normalized intensity against time for these signals (Figure 3-7C). The rate constant was calculated for each substrate and the relative activity was obtained (Figure 3-7D). The values for the relative activity of A3G_C toward C6 (in 5'-CCC hotspot) are almost the same for s3-2_poly-dA linker, s3-2_poly-rA

linker and s3-2_abasic linker. Similarly, the values for the relative activity of A3G_C toward C34 (in 3'-CCC hotspot) are also almost the same for s3-2_poly-dA linker, s3-2_poly-rA linker and s3-2_abasic linker. Using these values, the 3'→5' polarity of A3G_C deamination activity toward each substrate was quantified as *P*, the ratio of the rate constant obtained for C6 to that obtained for C34 (Figure 3-7E). *P* for s3-2_poly-rA linker and s3-2_poly-rA linker were almost the same as that

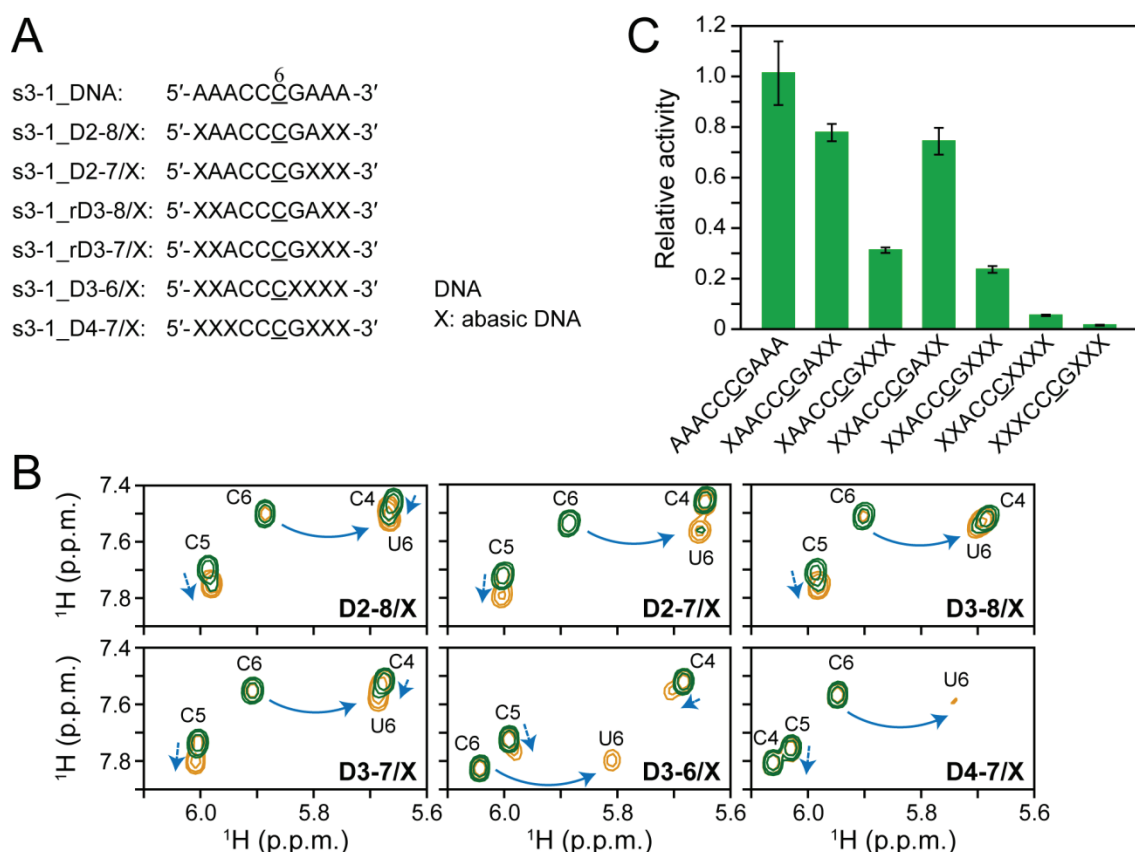


Figure 3-6.

Real-time monitoring of deamination reactions for substrates containing multiple abasic deoxyribonucleotides.

(A) The sequences of chimeric nucleotides. X corresponds to abasic deoxyribonucleotide. (B) Chemical shift assignment of TOCSY spectra of each ssDNA substrate in (A). Green and yellow signals in TOCSY spectra each present about 2 and 36 hours after addition of A3G_C. Signals of C6 and U6 are connected by a solid arrow, while a broken arrow indicates the signal perturbation of C5 due to the C6 to U6 conversion. (C) Relative activity was calculated by dividing the deamination rate constant for a substrates containing X substitutions.

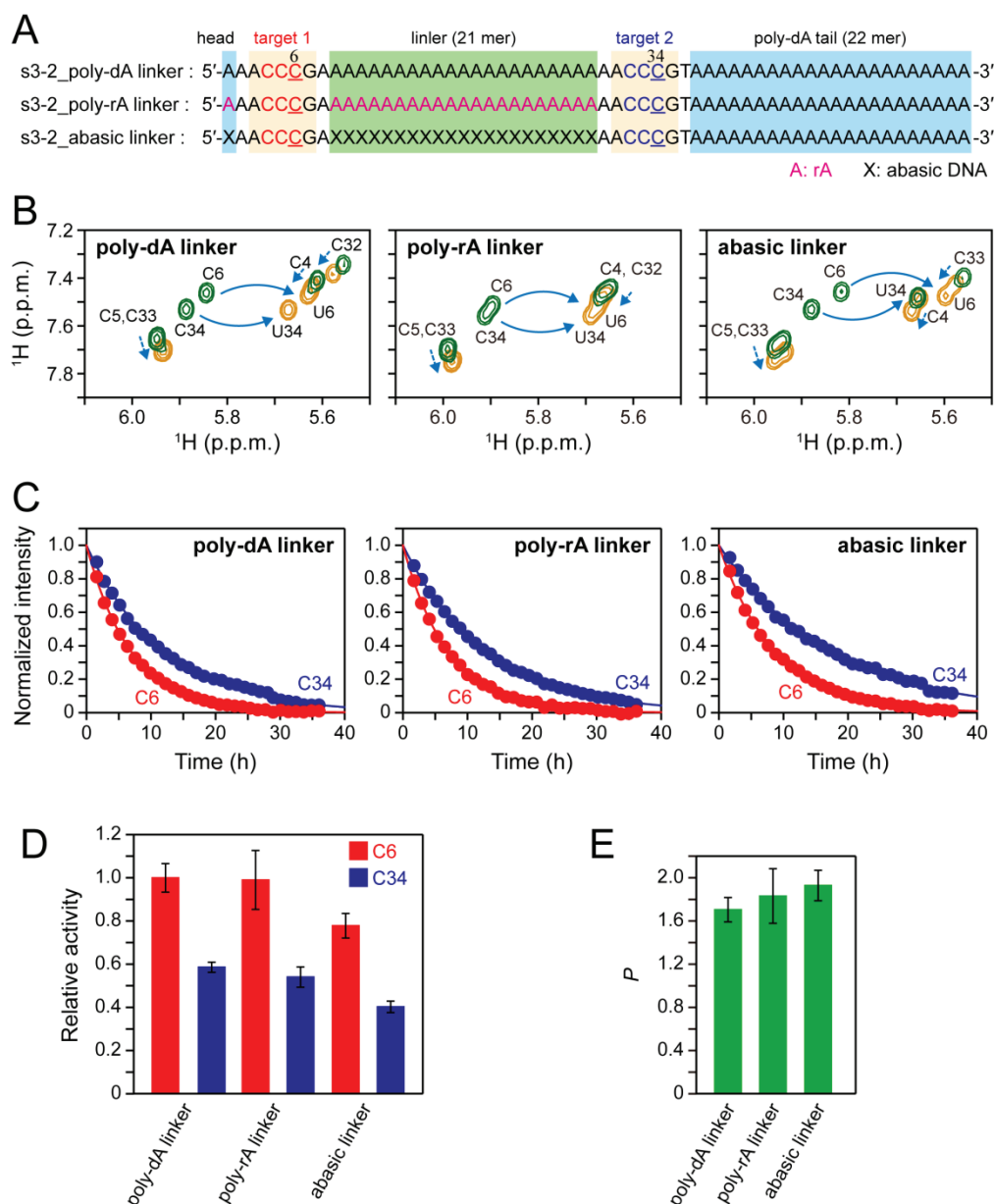


Figure 3-7.

The effect of the DNA to RNA or baselacking replacement at long DNA linker region on the 3'→5' polarity.

(A) Design of the chimeric nucleotides. Two CCC hot spots are connected by the linker, one is control DNA and the others are contains poly- adenine ribonucleotide (rA or A) or abasic coxyribonucleotide (X), respectively. The cytosines deaminated by A3G_C are underlined. (B-C) Deamination reaction against each nucleotide shown in A was monitored. TOCSY spectra of the each nucleotide were shown 2 hours (green) and 36 hours (yellow) after addition of A3G-C are superimposed (B). Signals of C6 and C34 are connected with those of U6 and U34 by solid arrows, respectively. The signal perturbations of the cytosines which is not involved the C to U conversion are also detected (broken arrows). The time course of the signal intensity for C6 (red) and C34 (blue) of the each nucleotide was displayed in (C) (D) Deamination activities toward C6 and C34 of each nucleotide shown in (A). Relative activity was defined as the ratio of the apparent deamination rate constant k obtained from the time course of C6 or C34 signal intensity of each nucleotide to that obtained for C6 of DNA substrate. (E) The 3'→5' polarity against each nucleotide shown in (A) is displayed. P is defined as the ratio of the apparent rate constant k obtained for C6 to that obtained for C34 for each nucleotide.

for s3-2_poly-dA linker. These findings indicate that either a 2'-OH substitution or abasic substitution in the linker region connecting two CCC hotspots does not have a negative effect on the 3'→5' polarity of A3G_C's deamination activity. Since the chemical moiety that is conserved in these three substrates is the phosphate backbone, we hypothesized that the sugar and base moieties are not as important as the phosphate backbone for A3G_C to slide along ssDNA. In other words, we assume that the phosphate backbone tethers A3G_C to ssDNA via an electrostatic interaction when A3G_C slides on an ssDNA.

3-3-4. NaCl concentration dependence of the 3'→5' polarity

Our data suggested that the phosphate backbone of ssDNA contributes to its sliding. To confirm this, we analyzed the effect of the NaCl concentration on the 3'→5' polarity of A3G_C's deamination activity. In Figure 3-8A, the values for the relative activity of A3G_C as to positions C6 and C34 (underlined) of the s3-3 substrate (AAACCCGAA₂₁AACCCGTAA) are plotted against NaCl concentrations. The relative activity was calculated as the ratio of the rate constant obtained for C6 or C34 of s3_DNA at each NaCl concentration to that obtained for C6 of s3-3 at 30 mM NaCl. The plot of relative activity is bell-shaped (Figure 3-8), while the values for the 3'→5' polarity (P) is changed monotonically against NaCl concentrations (Figure 3-8), which are rationally explained in the Discussion section.

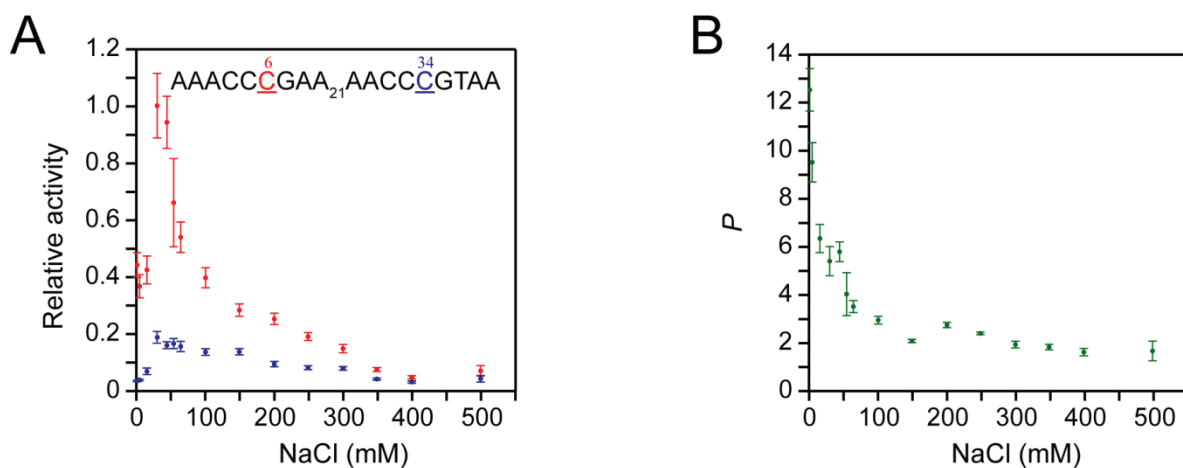


Figure 3-8.

The effect of the NaCl concentration on the 3'→5' polarity of A3G_C's deamination activity.

(A) Relative activity was defined as the ratio of the rate constant obtained for C6 or C34 of AAACCCGAA₂₁AACCCGTAA (s3-3) at each NaCl concentration to that obtained for C6 of s3-3 at 30 mM NaCl. (B) 3'→5' polarity values (P) are plotted against NaCl concentration. The P at each NaCl concentration is defined as the ratio of k obtained for C6 to that obtained for C34.

3-3. Discussion

3-3-1. The recognition sequence of A3G_C for deamination

In the Results section we demonstrated that the deamination activity of A3G_C toward chimeric DNA dramatically decreased with a nucleotide substitution in the central five sequential nucleotide region (Figure 3-1–3-6). These results indicate that A3G_C requires the 5 nucleotide recognition sequence for deamination. Similar results were obtained on nucleotide analog interference mapping analysis, which involving a UDG-based assay (3). We assume that introduction of 2'-OH caused a direct steric clash between the hydroxyl group and the surface of the catalytic pocket of A3G_C and/or an indirect one caused by conversion of the preferred sugar ring puckering from the C2'-endo to the C3'-endo conformation. A direct and/or indirect clash is supposed to prevent C6 from properly fitting into the catalytic pocket. A similar idea was proposed for human activation-induced deaminase (AID), a member of the APOBEC protein family, to interpret the reduction of activity caused by a 2'-F substitution in substrate DNA (6).

In Figure 3-5, the reactivity of s3-1_D2-7/R is higher than that of s3-1_D3-8/R, which indicates that the introduction of a 2'-OH substitution at the 5'-flanking position of the 5 nucleotide recognition sequence has a higher negative effect on the deamination activity than one at the 3'-flanking position. Thus, the requirement of being a deoxyribonucleotide is higher at the 5'-flanking position than at the 3'-flanking position of the 5 nucleotide recognition sequence. In Figure 3-6, the reactivity of s3-1_D2-7/X is lower than that of s3-1_D3-8/X, and is as low as that of s3-1_D3-7/X, which indicates that an abasic substitution at the 3'-flanking position of the 5 nucleotide recognition sequence has a higher negative effect on the deamination activity of A3G_C than one at the 5'-flanking position. Thus, the base moiety is more highly required at the 3'-flanking position than at the 5'-flanking position of the 5 nucleotide recognition sequence. Altogether, the substitution of the chemical moieties of nucleotides revealed that the regions relevant to the activity of A3G_C extend outside the 5 nucleotide recognition sequence defined in the previous section, and that a deoxyribose and a base are required at the 5'- and 3'-flanking positions of the 5 nucleotide recognition sequence, respectively.

3-3-2. The nucleic acid determinant for A3G_C's sliding.

In Figure 3-7, we demonstrates that the sugar and base moieties do not affect the deaminase activity of A3G_C or 3'→5' polarity. Thus, we assume that the electrostatic interaction via the phosphate backbone may be important when A3G_C slides on an ssDNA. Figure 3-8, we evaluate the effect of NaCl concentration on deaminase activity as well as 3'→5' polarity. The relative activity of A3G_C at both C6 and C34 was the highest at 30 mM NaCl. As the concentration of NaCl increased from 30 mM, the relative activity of A3G_C decreased. This is due to destabilization of the complex of A3G_C with the hotspot through a decrease of the electrostatic interaction. The decreased electrostatic interaction also causes the abortion of sliding, which results in a decrease of

activity. The relative activity of A3G_C also decreased as the NaCl concentration decreased from 30 mM. In this case, the complex of A3G_C with the hotspot is stabilized excessively under low NaCl conditions, which could result in interference with turn over. Additionally, sliding slows down on tight interaction of A3G_C with the DNA strand, which results in a reduction of activity. In Figure 3-8B, the values for the 3'→5' polarity (P) of A3G_C are plotted against NaCl concentrations. Interestingly, P was highest at the lowest NaCl concentration investigated in the present study (1 mM). P decreased rapidly as the concentration of NaCl increased, and eventually reached the value of approximately one around the concentration of 400 mM. This indicates that the efficiencies of deamination at the C6 and C34 sites became the same, namely, the 3'→5' polarity was lost, which was probably due to the alleviation of sliding through facilitated dissociation of A3G_C from the substrate ssDNA. P increased rapidly as the concentration of NaCl decreased. It is assumed that the sliding speed itself decreased at low NaCl concentrations due to tight electrostatic interaction of A3G_C with the DNA strand. However, it is also assumed that A3G_C hardly dissociated from the DNA strand and continued to slide along it, which resulted in the high 3'→5' polarity value.

3-3-3. Deamination mechanism of A3G_C

In the previous section, we mentioned two kinds of interface of nucleic acids, a 5 nucleotide recognition sequence and an electrostatic interaction via the phosphate backbone. Here we discuss the interface in A3G_C. The charge and basic amino acids (arginine, histidine and lysine) distribution on molecular surface are shown in Figure 3-9A, B, respectively. A positive charge (blue) as well as basic amino acids are spread out over the molecular surface of A3G_C, which does not correspond to the proposed DNA binding groove (4, 7, 8). On the other hand, the distribution of a negative charge is similar to the proposed DNA binding groove shown by Chen *et al.* (7) or Furukawa *et al.* (4). It is also possible that many residues take part in A3G_C's sliding and we cannot focus on only a few amino acids of this time. These findings indicate A3G_C uses individual binding surfaces when it slides along DNA and when it binds to DNA for deamination. In other words, A3G_C possesses two different binding mechanisms. Previously, Furukawa and coworkers reported that A3G_C slides along DNA without directional preference and that the catalytic rate depends on the direction of sliding (5). Adding a new finding, we propose a deamination mechanism for A3G_C (Figure 3-10). A3G_C requires the 5 nucleotide recognition sequence to exert deaminase activity (Figure 3-19A). Before binding this area, A3G_C slides along the phosphate backbone via an electrostatic interaction with no directional preference to search this area. 3'→5' polarity may cause the difference of entry, and switch the interface of sliding and binding with the recognition sequence (Figure 3-10C).

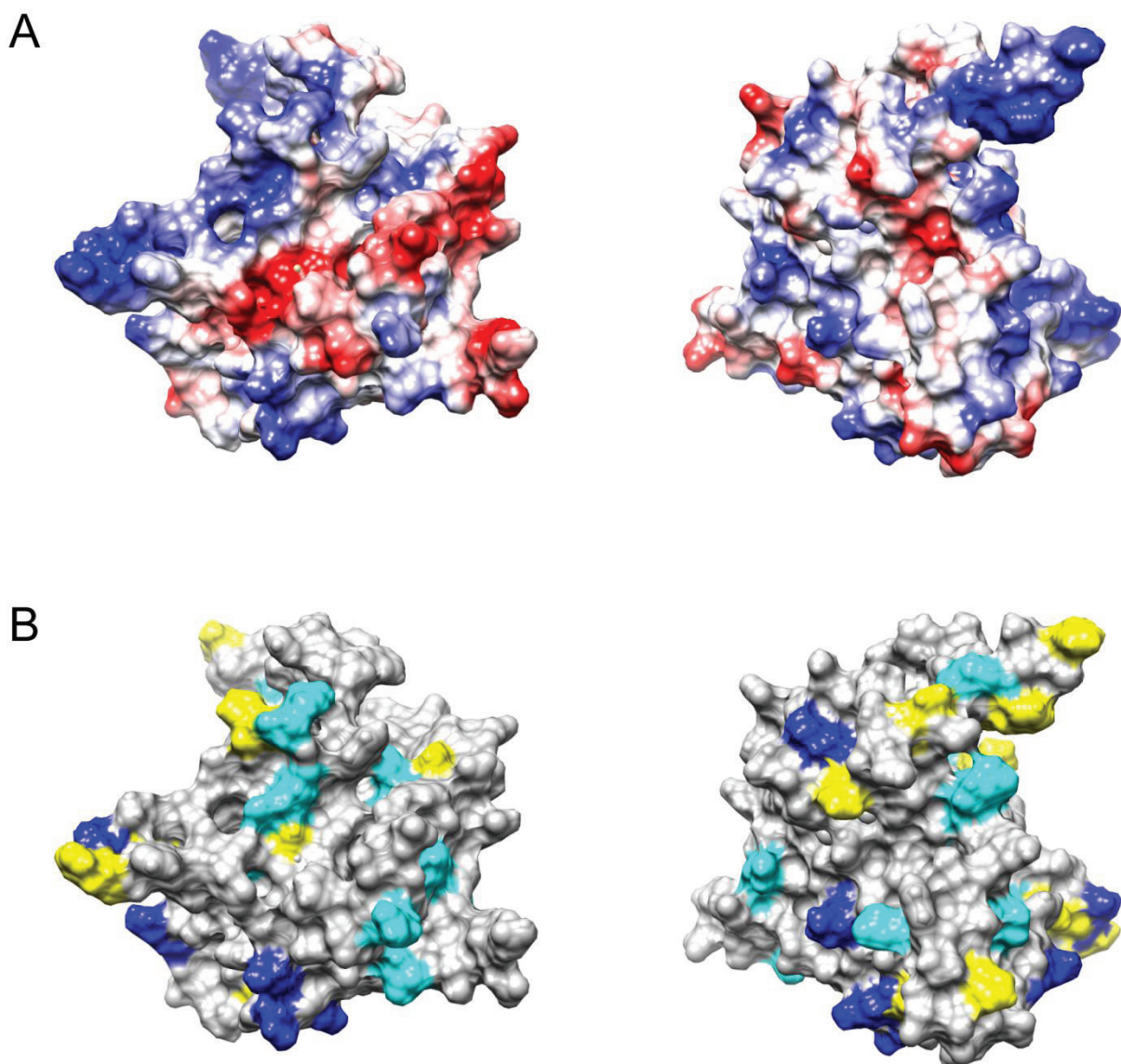


Figure 3-9.

The distribution of a positive charge on A3G_C.

(A) The charge distribution on the molecular surface is shown. Positive and negative charged regions are shown in blue and red, respectively. (B) The basic amino acids distribution on the molecular surface are highlighted. The arginine, histidine and lysine residues are colored cyan, yellow and blue, respectively. Molecular graphics and analyses were performed with the UCSF Chimera package (9).

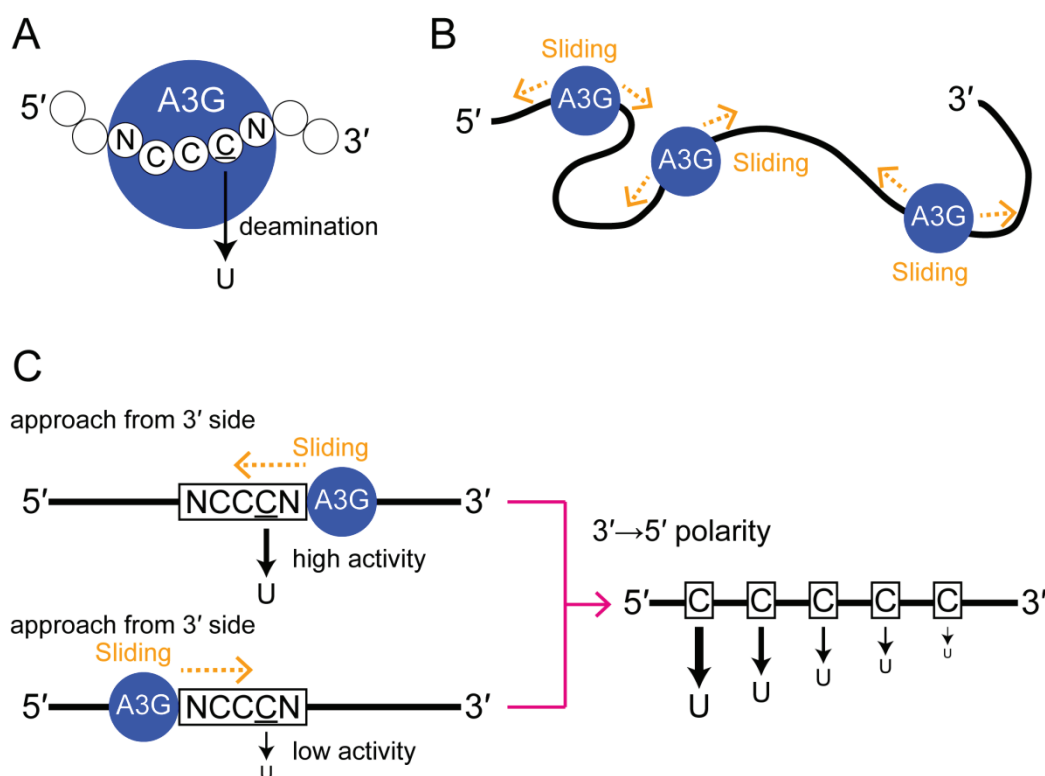


Figure 3-10.

Deamination mechanism for A3G_C proposed in this study.

(A) A3G_C requires the 5 nucleotide recognition sequence to exert deaminase activity. (B) Before binding this area, A3G_C slides along the phosphate backbone via an electrostatic interaction with no directional preference to search this area. (C) The proposed origin of 3'→5' polarity is demonstrated. 3'→5' polarity may cause the difference of entry, and switch the interface of sliding and binding with the recognition sequence. These properties generate the deamination bias.

3-4. Conclusion

In this chapter, we investigated the nucleic acid determinants for deamination by A3G_C using a real-time NMR method in combination with a series of ssDNA substrates each carrying a nucleotide analog at a single or multiple positions. Firstly, we showed that the sugar and base moieties of the 5 consecutive nucleotides, positioning the CCC hotspot at the center, as well as the 5'- and 3'-flanking regions of the 5 consecutive nucleotides, are important for A3G_C to exert its deamination activity. Secondly, we showed that A3G_C can tolerate the introduction of either a 2'-OH substitution in the sugar moiety or the removal of the base in the consecutive poly-nucleotide to slide along an ssDNA and exert the 3'→5' polarity. This suggests that the electrostatic interaction between A3G_C and the phosphate backbone of an ssDNA is the key for sliding. This idea was confirmed by the fact that the 3'→5' polarity is dependent on the NaCl concentration.

3-5 Experimental Procedures

Preparation of the protein and DNA substrates

A3G_C was expressed purified as described previously in Chapter 2. The obtained A3G_C was dialyzed against the storage buffer comprising 20 mM Tris-HCl, pH 7.5, 30 mM NaCl, 5 mM DTT, and 10 μ M ZnCl₂, and then stored until use at 4°C. All substrate ssDNAs were purchased from Fasmac Co., Ltd.

NMR spectroscopy

All NMR spectra were recorded at 25°C with a DRX600 spectrometer equipped with a cryogenic probe and a Z-gradient (Bruker Biospin). A final concentration of 200 μ M each DNA substrate was dissolved in 20 mM Tris-HCl (pH 7.5), 30 mM NaCl, 10 μ M ZnCl₂, 5 mM DTT, and 5% deuterium oxide. A final concentration of 0.8 or 2 μ M A3G_C was used. To examine the NaCl concentration-dependence of A3G_C's deaminase activity, the NaCl concentration was set to 1, 5, 15, 30, 45, 55, 65, 100, 150, 200, 250, 300, 350, 400, or 500 mM. After the addition of A3G_C, 2D TOCSY (mixing time, 20 ms), 2D ¹H-¹³C HSQC spectra of each DNA substrate were recorded at different time points to monitor the deamination reaction in real-time. Water signal suppression was achieved with 3-9-19 watergate, and echo-antiecho pulse schemes in each 2D TOCSY and 2D ¹H-¹³C HSQC measurement (4, 5). Resonance assignments of the oligonucleotides were carried out using 2D TOCSY and 2D ¹H-¹³C HSQC spectra (4, 5). Spectra were processed with NMRPipe (10), and analyzed using SPARKY (11). The intensities of the peaks of either cytosine or uracil were plotted against time and analyzed as in the previous section. The data were fitted to a single exponential decaying function, using the following equation, by which apparent deamination rate constants were derived:

$$I(t)_C = I_s \exp(-kt) + I_b \quad (3 - 1)$$

, where I_b corresponds to the baseline intensity of the spectrum, $I_s + I_b$ the signal intensity at time point zero for cytosine, and k the apparent deamination rate constant, which was always obtained through linear least-squares analysis. The value for the apparent deamination rate constant was used as an index of activity. We defined P with the following equation:

$$P = \frac{k_{(C6)}}{k_{(C34)}} \quad (3 - 2)$$

, where $k_{(C6)}$ and $k_{(C34)}$ are the apparent deamination rate constants at positions C6 and C34 of a 38-mer ssDNA (s3_DNA). P can be regarded as a measure of the magnitude of 3'→5' polarity. The error was estimated as follows. Firstly, the error of the NMR signal intensity was obtained from the base level noise of each NMR spectrum. Then the error of the rate constant was calculated from data sets constructed by Monte Carlo simulation using the error of the NMR signal intensity. Finally, the errors of the relative activity and the P were obtained by means of error propagation calculation.

3-6. References

- (1) Yu, Q., König R., Pillai, S., Chiles, K., Kearney, M., Palmer, S., Richman, D., Coffin, J. M. and Landau, N. R. (2004) Single-strand specificity of APOBEC3G accounts for minus-strand deamination of the HIV genome. *Nat. Struct. Mol. Biol.* **11**, 435–442.
- (2) Chelico, L., Pham, P., Calabrese, P. and Goodman, M. F. (2006) APOBEC3G DNA deaminase acts processively 3' → 5' on single-stranded DNA. *Nat. Struct. Mol. Biol.* **13**, 392–399.
- (3) Rausch, J. W., Chelico, L., Goodman, M. F. and Le Grice, S. F. (2009) Dissecting APOBEC3G substrate specificity by nucleoside analog interference. *J. Biol. Chem.* **284**, 7047–7058.
- (4) Furukawa, A., Nagata, T., Matsugami, A., Habu, Y., Sugiyama, R., Hayashi, F., Kobayashi, N., Yokoyama, S., Takaku, H. and Katahira, M. (2009) Structure, interaction and real-time monitoring of the enzymatic reaction of wild-type APOBEC3G. *EMBO J.* **28**, 440–451.
- (5) Furukawa, A., Sugase, K., Morishita, R., Nagata, T., Kodaki, T., Takaori-Kondo, A., Ryo, A., and Katahira, M. (2014) Quantitative analysis of location- and sequence-dependent deamination by APOBEC3G using real-time NMR spectroscopy. *Angew. Chem. Int. Ed. Engl.* **53**, 2349–2352.
- (6) Nabel, C. S., Lee, J. W., Wang, L. C. and Kohli, R. M. (2013) Nucleic acid determinants for selective deamination of DNA over RNA by activation-induced deaminase. *Proc. Natl. Acad. Sci. USA* **110**, 14225–12230.
- (7) Chen, K. M., Harjes, E., Gross, P. J., Fahmy, A., Lu, Y., Shindo, K., Harris, R. S. and Matsuo, H. (2008) Structure of the DNA deaminase domain of the HIV-1 restriction factor APOBEC3G. *Nature* **452**, 116–119.
- (8) Holden, L. G., Prochnow, C., Chang, Y. P., Bransteitter, R., Chelico, L., Sen, U., Stevens, R. C., Goodman, M. F. and Chen, X. S. (2008) Crystal structure of the anti-viral APOBEC3G catalytic domain and functional implications. *Nature* **456**, 121–124.
- (9) Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612.
- (10) Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293.
- (11) Goddard, T. D. and Kneller, D. G. (2006) SPARKY 3. *San Francisco: University of California.*

Chapter 4

Deamination property of A3G_C coupled with sliding on the basis of the pH-dependence

4-1. Introduction

In Chapter 3, it was shown that phosphate backbone is important for A3G_C to slide on the DNA strand and thereby to exert the 3'→5' polarity. Subsequent analysis to investigate the dependence of salt concentration on 3'→5' polarity suggested that electrostatic intermolecular interaction tethers A3G_C onto ssDNA probably through interaction via phosphate backbone. In this chapter, we investigated the pH effects on the structure, deaminase activity, and 3'→5' polarity of A3G_C by using the real-time NMR monitoring method.

4-2. Results

4-2-1. The effects of pH and denaturant on the structure of A3G_C

Here we assessed the structure of A3G_C in buffers having different pHs using CD spectroscopy. Far-UV (200–260 nm) CD spectra were obtained at nine different pH values within the range of 2.5–12.7 (Figure 4-1A, B). The patterns of the spectra in the range of pH 4.0–10.5 were almost identical to that of the spectrum at pH 7.5, where A3G_C reportedly exerts deaminase activity (1–4), indicating that A3G_C is in the native form within this pH range. Indeed, the spectra in the range of pH 4.0–10.5 show strong negative intensity around the wavelength range of 210–230 nm, which is generally known as an indication of the presence of secondary structures (5). In Figure 4-1C, the variation in the molar ellipticity at 220 nm ($[\theta]_{220}$) of A3G_C at 20 different pH values in the range of 2.5–12.7 is showed. The average value of $[\theta]_{220}$ was $-5.1 \times 10^3 \text{ deg} \cdot \text{cm}^2 \cdot \text{dmol}^{-1}$ in the range of pH 4.0–10.5, where A3G_C has the native structure, as mentioned above. However, at pH values lower and higher than pH 4.0 and 10.5, respectively, the values of $[\theta]_{220}$ were higher (ca. $-4.0 \times 10^3 \text{ deg} \cdot \text{cm}^2 \cdot \text{dmol}^{-1}$), which implies that the content of the secondary structures was somehow lower.

Next, to evaluate whether or not A3G_C is in the native form in the region of $\text{pH} < 4.0$ and $\text{pH} > 10.5$ ($[\theta]_{220}$ value of ca. -4.0×10^{-3}), we conducted guanidine-induced denaturation of A3G_C at pH 7.5 (Figure 4-2A). The $[\theta]_{220}$ value was plotted against the concentration of guanidine hydrochloride (GdHCl) (Figure 4-2 B). Then, we fitted the plotted data to equation (4-1) in the Experimental Procedures section to obtain a guanidine-induced denaturation curve, which describes a two-state transition model, and determined the apparent unfolding free energy (ΔG_{app}) and cooperativity (m_{app}) values to be $3.3 \pm 0.3 \text{ kcal} \cdot \text{mol}^{-1}$ and $1.6 \pm 0.1 \text{ kcal} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$, respectively (Table 4-1). For reference, we also investigated guanidine-induced denaturation of A3G_C at pH 7.5 using fluorescence spectroscopy (Figure 4-2C) The fluorescence spectrum with zero GdHCl showed

a maximum at 335 nm, while at the highest GdHCl concentration (5.8 M), the maximum was at 348 nm (red-shifted), which strongly suggests that the tertiary structure of A3G_C has denatured almost completely (6) (Figure 4-2C). Here, fluorescence intensity at 335 nm was plotted against the concentration of GdHCl and fitted to the aforementioned equation (4-1), and the ΔG_{app} and m_{app} values were determined to be $3.2 \pm 0.4 \text{ kcal}\cdot\text{mol}^{-1}$ and $2.0 \pm 0.2 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{M}^{-1}$, respectively (Table 4-1). The $[\theta]_{220}$ value of the CD spectrum and the fluorescence intensity at 335 nm are generally appreciated to be probes to determine the contents of secondary and tertiary structures, respectively (6). Since, the ΔG_{app} values obtained on CD and fluorescence spectroscopies are the same within error, the unfolding of A3G_C's secondary and tertiary structures is highly cooperative.

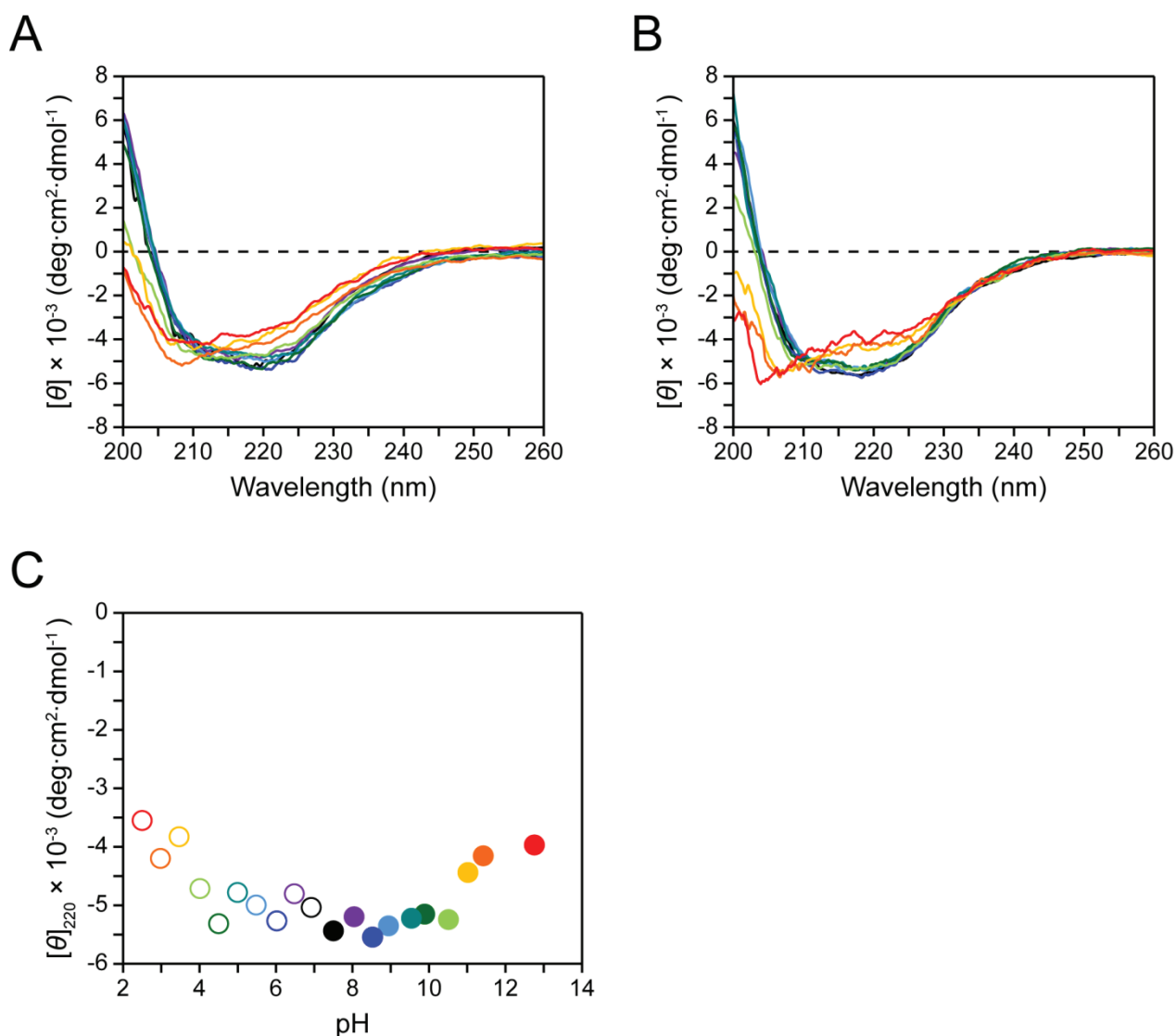


Figure 4-1.

The pH dependence of the structure of A3G_C by CD spectra.

(A–B) Far-UV (200–260 nm) CD spectra of A3G_C at pH 2.5–7.0 (A) and pH 7.5–12.7 (B) are shown. (C) The pH dependence of molar ellipticity at 220 nm. Colors correspond to those used in (A), (B).

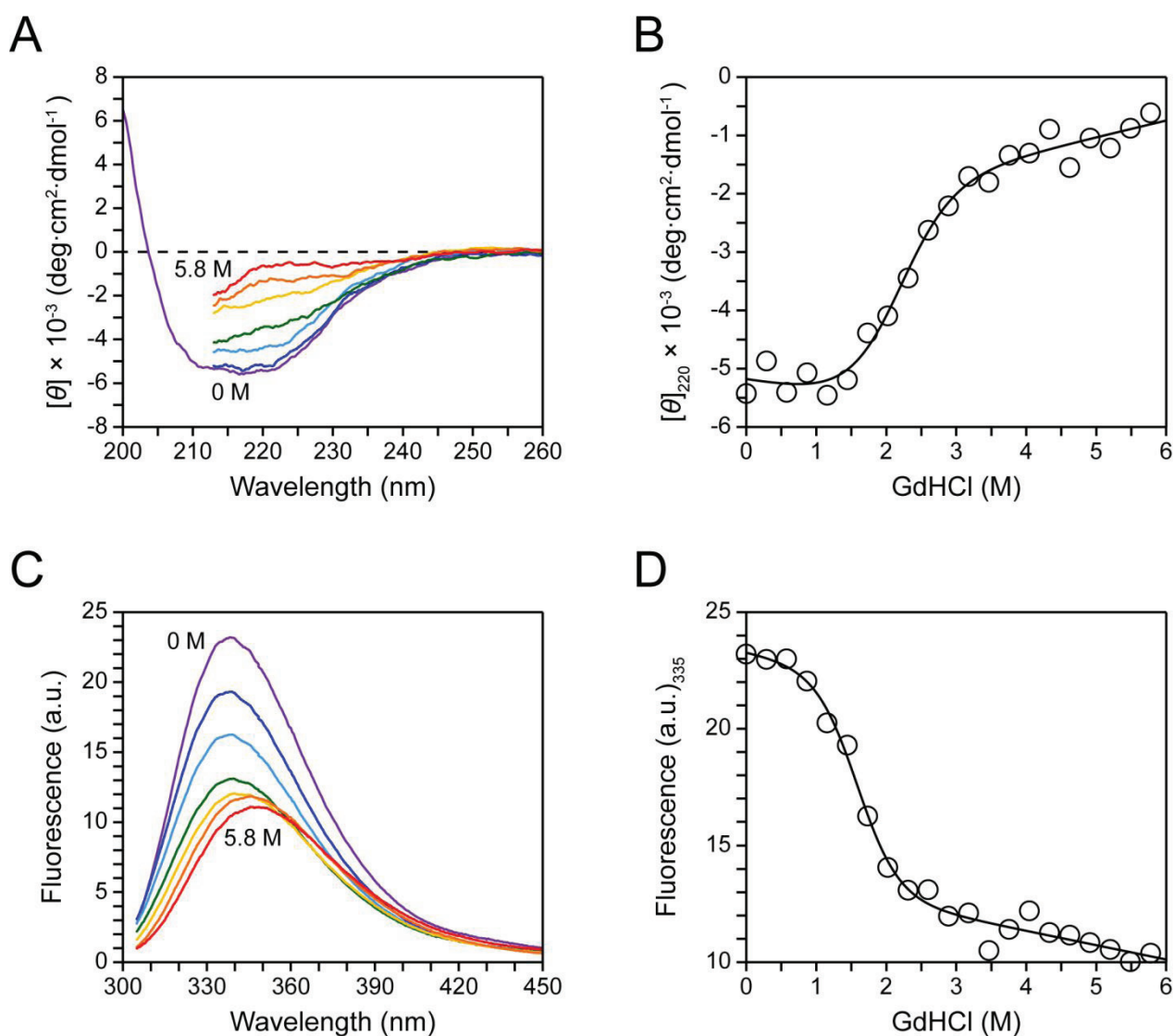


Figure 4-2

GdHCl-induced equilibrium unfolding transitions of A3G_C by CD and fluorescence spectra.

(A) CD spectra of A3G_C in concentration of GdHCl from 0 to 5.8 M (in 0, 1.4, 1.7, 2.3, 2.9, 3.8, 5.8 M) at pH 7.5. (B) Unfolding transition curves monitored by molar ellipticity at 220 nm in (A). (C) Fluorescence spectra of A3G_C in concentration of GdHCl from 0 to 5.8 M (in 0, 1.4, 1.7, 2.3, 2.9, 3.8, 5.8 M) at pH 7.5. Colors correspond to those used in (A). (D) Unfolding transition curves monitored by fluorescence intensity at 335 nm in (C).

Table 4-1.

Thermodynamic parameters for guanidine-induced denaturation of A3G_C obtained by CD and fluorescence spectroscopy.

CD		Fluorescence	
ΔG_{app} (kcal·mol ⁻¹)	m_{app} (kcal·mol ⁻¹ ·M ⁻¹)	ΔG_{app} (kcal·mol ⁻¹)	m_{app} (kcal·mol ⁻¹ ·M ⁻¹)
3.3 ± 0.3	1.6 ± 0.1	3.2 ± 0.4	2.0 ± 0.2

4-2-2. The effect of pH on the deaminase activity of A3G_C

Our real-time NMR monitoring method was applied to investigate the deaminase activity of A3G_C for a 10-mer ssDNA (s4-1), AAACCCGAAA (C to U conversion occurs at underscored cytosine, C6), at nine different pH values (pH 6.5, 7.0, 7.5, 8.0, 8.5, 9.5, 10.5, 11.4, and 12.7). In each experiment, the final concentrations of A3G_C and s4-1 were adjusted to 0.8 μ M and 200 μ M, respectively. A series of 2D TOCSY spectra for s4-1 were obtained at several reaction time points at each pH value. As an example, Figure 4-3A shows the spectra of s4-1 at pH 7.5, at time points 1.6 and 53 hours after addition of A3G_C to s4-1 (green and yellow signals, respectively). In Figure 4-3A, the H5–H6 correlation signals of C4 and U6 overlap, so the assignments were made with the aid of 2D ^1H – ^{13}C HSQC spectroscopy to discriminate between C and U (data not shown).

We collected the intensities of C6 signals, which were isolated from other signals, and followed their decreases during the reaction time course for each pH value. The time courses of the intensity decrease are plotted in Figure 4-3B. The plot for each pH was fit to aforementioned single exponential decaying function, equation (4-3), individually, and the rate constant, k , at each pH was obtained (Table 4-2). Since the k value can be considered as an index of activity, we defined the relative activity as the ratio of k at a particular pH to that at pH 6.5. In the pH range studied here, the relative activity was highest at pH 6.5 and decreased as the pH increased (Figure 4-3C).

We also applied the current analysis to the lower pH range of pH 2.5 to 6.0. Within this pH range, the time course of the intensity decrease of C6 signals did not show simple single exponential decay, but rather a sigmoidal-like curve. Our intention was to investigate the pH dependence of the 3'→5' polarity of the deaminase activity, therefore we decided to carry out the subsequent experiments in the pH range of 6.5–12.7 to make the analysis simple.

Table 4-2.

Relative activities of A3G_C for s4-1 at pHs ranging from 6.5 to 12.7.

pH	Relative activity (C6)
6.5	1.00 ± 0.05
7.0	0.64 ± 0.10
7.5	0.25 ± 0.02
8.0	0.091 ± 0.005
8.5	0.089 ± 0.004
9.5	0.015 ± 0.001
10.5	$(5.0 \pm 0.3) \times 10^{-3}$
11.4	$(6.4 \pm 0.5) \times 10^{-4}$
12.7	n.d.

n.d., not detected.

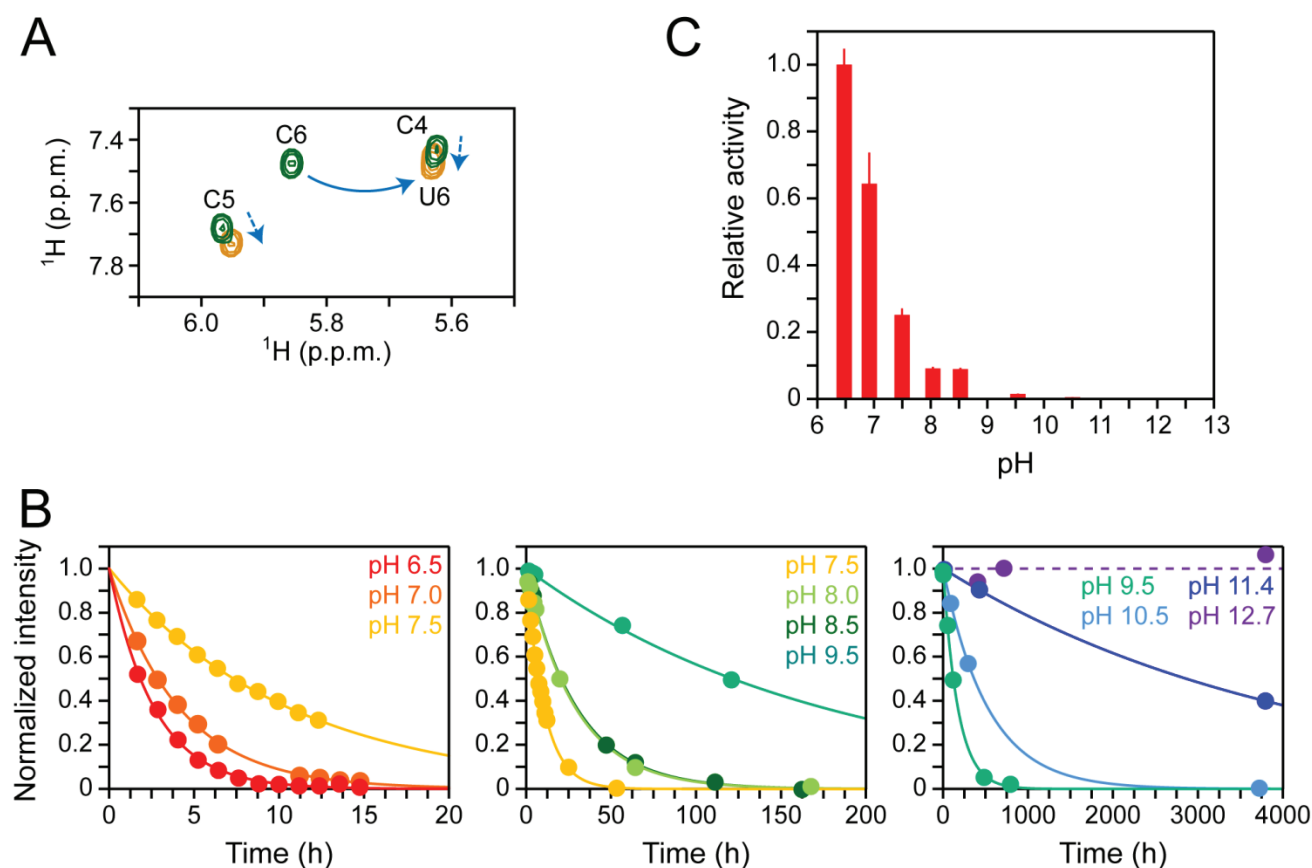


Figure 4-3.

Effect of pH on the deaminase activity of A3G_C for a 10-mer ssDNA as assessed by real-time NMR monitoring.

(A) Superpositioning of 2D TOCSY spectra of AAACCCGAAA (s4-1) at two reaction time points at pH 7.5 is presented as a representative. C to U conversion occurred at the underscored cytosine (C6). The H5–H6 correlation signals of cytosines and uracils are shown: green and yellow signals correspond to reaction time points 1.6 and 53 hours after addition of A3G_C, respectively. The signal of C6 is connected with that of U6 by a solid arrow. The signals of C4 and C5 are perturbed (broken arrows) due to the C6 to U6 conversion. (B) The time courses of the intensity changes were plotted for C6 of s4-1 at pH 6.5, 7.0, 7.5, 8.0, 8.5, 9.5, 10.5, 11.4, and 12.7. The data for each pH were fitted to a single exponential function individually, by which the rate constant, k , at each pH was obtained. (C) Relative activities for C6 are plotted against pH values. The relative activity at each pH was defined as the ratio of k at that particular pH to that obtained at pH 6.5.

4-2-3. The effect of pH on the 3'→5' polarity of A3G_C

We, next, used a 38-mer ssDNA substrate (s4-2), AAACCCGA₂₄CCCGTAA, to investigate the deamination reaction of A3G_C at nine different pH values (pH 6.5, 7.0, 7.5, 8.0, 8.5, 9.5, 10.5, 11.4, and 12.7). This s4-2 contains two CCC hotspots, whose third positions, C6 and C34 (underlined in the sequence), are converted to uracil by A3G_C. In each experiment, the final concentrations of A3G_C and s4-2 were adjusted to 0.8 μ M and 200 μ M, respectively. A series of 2D TOCSY spectra for s4-2 at several reaction time points were obtained at each pH value in the same way as for s4-1 in the previous sections. We collected the intensities of C6 and C34 signals, and plotted them against time (Figure 4-4A). The plots for the C6 and C34 signals at each pH were each fitted to equation (4-3) individually, by which the rate constants were obtained. The relative activities at different pHs were then calculated (Table 4-3), and are compared in Figure 4-4B. Here the relative activities for C6 and C34 at each pH were defined as the ratios of their k values to the k value obtained for C6 at pH 6.5. The pH dependence trends of the relative activity for C6 and C34 in s4-2 were the same in each case as the trend observed for C6 in s4-1; the k values being highest at around pH 6.5–7.0 and decreased as the pH increased (Figure 4-4B). Again, the activity of A3G_C decreased significantly above pH 9.5 and became negligible around pH 12.7. At all pH values, the relative activity for C6 was higher than that for C34.

Subsequently, the 3'→5' polarity of A3G_C's deaminase activity for s4-2 at each pH value was quantified as P , the ratio of the rate constant obtained for C6 to that obtained for C34 (Figure 4-4C and Table 4-3). P showed the highest value of ca. 4.8 at pH 6.5 and decreased as the pH increased, and then P reached a plateau with a value of ca. 3.5 at pH 8.0 and stayed almost the same to pH 11.4.

Table 4-3.

Relative activity and 3'→5' polarity values (P) of A3G_C for s4-2 at pHs ranging from 6.5 to 12.7.

pH	Relative activity (C6)	Relative activity (C34)	3'→5' polarity (P)
6.5	1.00 ± 0.14	0.21 ± 0.06	4.8 ± 0.6
7.0	1.03 ± 0.21	0.24 ± 0.06	4.3 ± 0.8
7.5	0.29 ± 0.04	0.073 ± 0.011	3.9 ± 0.4
8.0	0.063 ± 0.011	0.021 ± 0.004	3.0 ± 0.5
8.5	0.076 ± 0.008	0.022 ± 0.002	3.4 ± 0.3
9.5	0.018 ± 0.002	(5.6 ± 0.4)×10 ⁻³	3.3 ± 0.2
10.5	(6.6 ± 0.9)×10 ⁻³	(2.1 ± 0.2)×10 ⁻³	3.2 ± 0.4
11.4	(1.6 ± 0.2)×10 ⁻³	(4.7 ± 1.1)×10 ⁻⁴	3.5 ± 0.3
12.7	n.d.	n.d.	n.a.

n.d., not detected. n.a., not applicable.

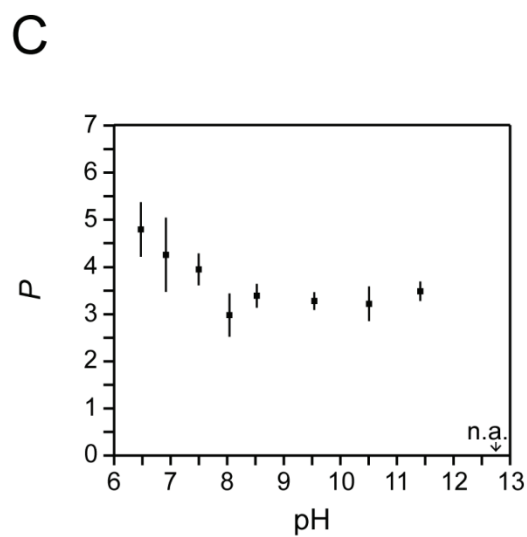
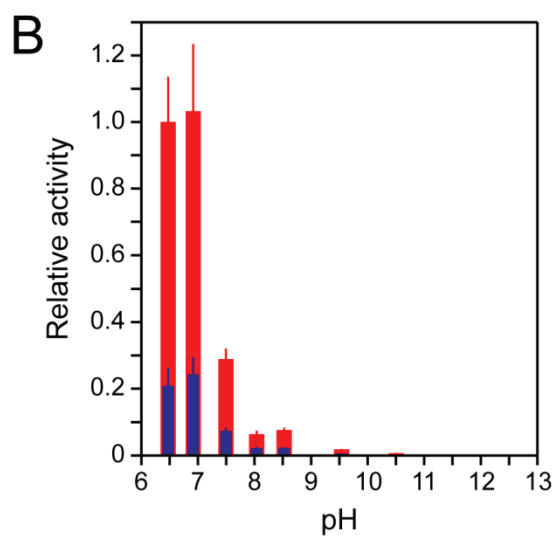
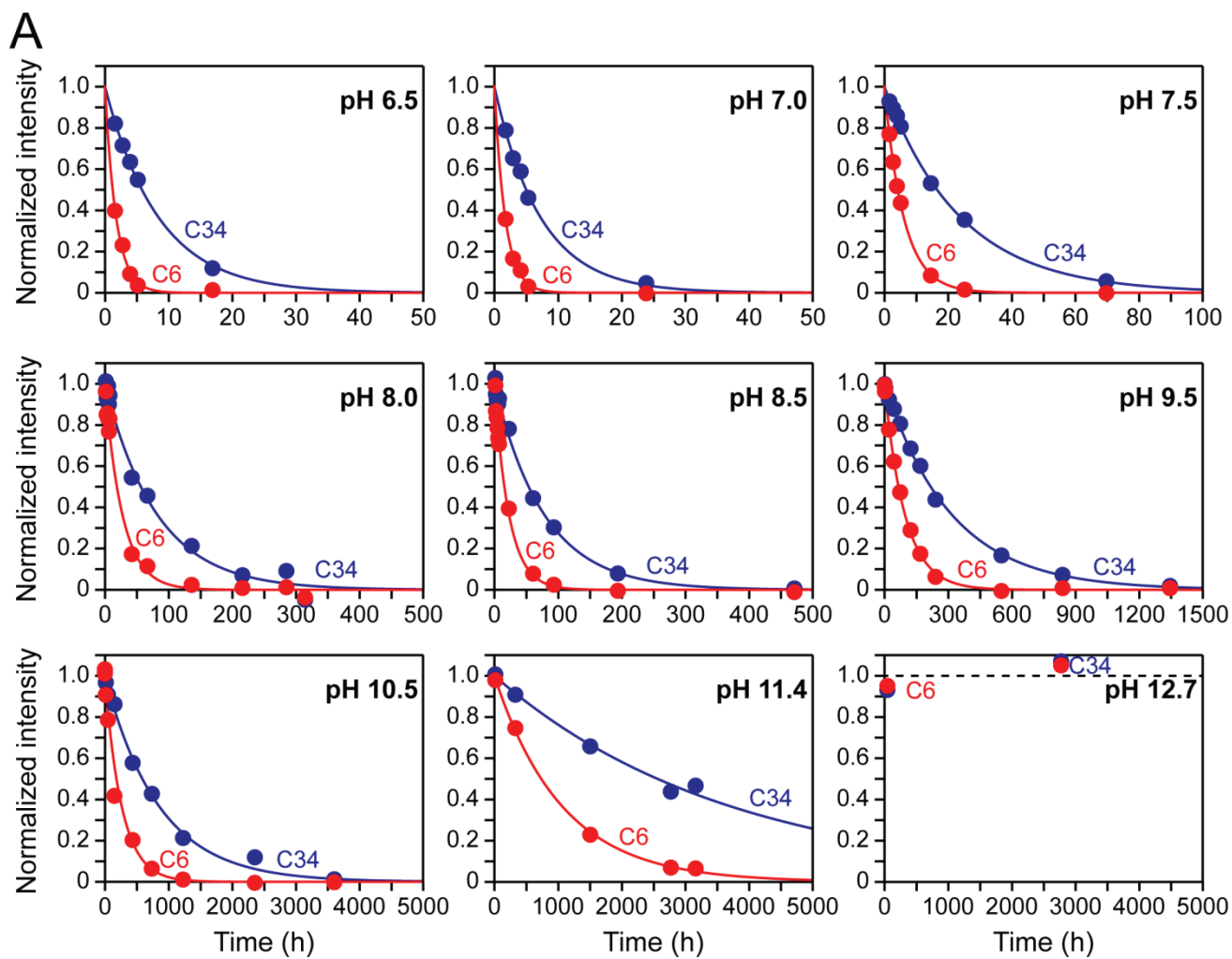


Figure 4-4.

Real-time NMR monitoring of deamination reactions for ssDNA having two CCC hotspots at various pHs.

(A) The time courses of the intensity were plotted for C6 and C34 of AAACCCGA₂₄CCCGTAA (s4-2) at pH 6.5, 7.0, 7.5, 8.0, 8.5, 9.5, 10.5, 11.4 and 12.7. C6 and C34, which are converted to U, are underscored in the sequence. The data of C6 and C34 for each pH were fitted to a single exponential function individually, by which the rate constant, k , at each pH was obtained. (B) Relative activities for C6 (red) and C34 (blue) are plotted against pH values. The relative activities for C6 and C34 at each pH were defined as the ratio of their k s at each pH to that obtained for C6 at pH 6.5. (C) 3'→5' polarity values (P) are plotted against pH values. P at each pH is defined as the ratio of k obtained for C6 to that obtained for C34.

4-3. Discussion

4-3-1. The origin of low deaminase activity in higher pH range

We then estimated the fraction of the folded state for A3G_C at different pH values using equation (4-2). As a result, in the range of pH 4.0–10.5, the fraction of the folded state was found to be 90 %, while it was 58 % in the range of pH 11.0–12.7. The deaminase activity of A3G_C turned out to be higher at lower pH and decreased as the pH increased within the range of pH 6.5–11.4. ssDNA is negatively charged because it has phosphate moieties, therefore we suppose that this trend is due to a reduction of the stability of the complex between A3G_C and ssDNA due to the weakened electrostatic interaction with an increase in pH value. We also noticed that the activity of A3G_C was significantly decreased above pH 9.5 and become negligible around pH 12.7. According to the analysis in the previous section, the fraction of the folded state of A3G_C in the range of pH 9.5–12.7 was between 92 % (pH 9.5) and 50 % (pH 12.7). Therefore, we assume that the low deaminase activity in this pH range is due not only to the denaturation of A3G_C, but also to decreased stability of the complex between A3G_C and ssDNA.

4-3-2. The difficulty of evaluating deaminase activity in lower pH range

Within the lower pH range of pH 2.5 to 6.0, the time course of the intensity decrease of C6 signals showed a sigmoidal-like curve. The reasons for this phenomenon are not apparent, however, the protonation/deprotonation equilibria of histidine and/or cytosine may play roles. Previously, it was shown that the deaminase activity of A3G_C becomes maximum at pH 5.5 (3). Since this pH is close to the pK_a value of the imidazole of histidine (ca. 6.0 in general), it was suggested that the protonation of histidine 216 is the key to the promotion of the deaminase activity (3). As for cytosine, we observed that the chemical shift values of the H5–H6 correlations were significantly perturbed at different pH values within the range of 2.5–6.0, which is supposed to be related to the protonation of

cytosine. In those reason, we conclude that it is difficult to evaluate the deaminase activity of A3G_C by single-exponential function as well as higher pHs, however those phenomena may be important to characterize the deaminase mechanism of A3G.

4-3-3. The key interaction for 3'→5' polarity

Previously, Furukawa and coworkers showed that A3G_C does not exert 3'→5' polarity for an ssDNA substrate whose two CCC hotspots are linked by a double-stranded DNA stretch (4). Then, we revealed that the electrostatic intermolecular interaction between A3G_C and ssDNA is the key for sliding, by demonstrating that the lower the NaCl concentration the higher the 3'→5' polarity in Chapter 3. These findings indicated that the 3'→5' polarity occurs because A3G_C slides along the phosphate backbone of the ssDNA stretch between the two CCC hotspots in Chapter 3. In this chapter, we investigated the pH dependence of the 3'→5' polarity. The results indicated that the lower the pH, the higher the 3'→5' polarity in the range of pH 6.5–8.0. These results resemble the effect of NaCl in Chapter 3. This can be explained by that the electrostatic intermolecular interaction between A3G_C and ssDNA is strengthened at lower pH, thereby the abortion of sliding is repressed, resulting in the high 3'→5' polarity.

4-4. Conclusion

Here, we investigated the effects of pH on the deamination reaction and 3'→5' polarity of A3G_C. A3G_C was firstly shown to retain its native structure fully or to some extent at the pH values studied (pH 4.0–12.7) by CD spectroscopy. Then real-time NMR spectroscopy was applied to a 10-mer ssDNA containing a single CCC, by which it was shown that the deaminase activity of A3G_C decreases as the pH increases. This was explained by a reduction in the stability of the complex between A3G_C and ssDNA due to the weakened electrostatic interaction with an increase in the pH value. Lastly, a 38-mer ssDNA having two CCC hotspots connected by a long poly-adenine linker was used to show that the 3'→5' polarity of A3G_C increases as the pH decreases in the range of 6.5–8.0. This finding can be rationalized that stronger electrostatic intermolecular interaction between A3G_C and ssDNA at lower pH represses the abortion of sliding, resulting in high 3'→5' polarity.

4-5. Experimental Procedures

Preparation of the protein and DNA substrates

A3G_C was expressed purified as described previously in Chapter 2. The obtained A3G_C was dialyzed against a storage buffer comprising 20 mM Tris-HCl, pH 7.5, 30 mM NaCl, 5 mM DTT, and 10 μ M ZnCl₂, and then stored until use at 4°C. All substrate ssDNAs, a 10-mer ssDNA (AAACCCGAAA; s4-1) and a 38-mer ssDNA (AAACCCGA₂₄CCCGTAA; s4-2) were purchased from Fasmac Co., Ltd.

Circular dichroism (CD) and fluorescence spectroscopy

Samples were prepared by mixing A3G_C in storage buffer with several measurement buffers having different pHs (1:19 ratio). The final concentration of A3G_C was adjusted to 6 μ M. Different buffer systems were used to obtain different pH solutions: Glycine-HCl (pH 2.5–3.5), Acetic acid-Sodium acetate (pH 4.0–5.5), Bis-Tris-HCl (pH 6.0–7.0), Tris-HCl (pH 7.5–8.5), CHES-NaOH (pH 9.0–9.5), CHAPS-NaOH (pH 10.0–11.0), and Triethylamine-HCl (pH 11.4–12.7). The concentrations of these buffers were adjusted to 20 mM, and 30 mM NaCl, 1 mM DTT and 10 μ M ZnCl₂ were added to obtain the measurement buffers. Equilibrium denaturation was carried out using different concentrations of guanidine hydrochloride (GdHCl) in the range of 0 to 5.8 M at pH 7.5. All CD and fluorescence measurements were carried out at 25°C.

CD spectra in the far-UV region (200–260 nm) using an optical path cell of 0.1 cm were obtained with a Jasco J-720A spectrometer (Japan Spectroscopic Co.). All spectra were background-corrected, smoothed, and converted to molar ellipticity $[\theta]$ (deg·cm²·dmol⁻¹) (5).

Fluorescence spectra were collected with a Jasco FP-8500 spectrometer (Japan Spectroscopic Co.) using an optical path cell of 1 cm. The spectra slit width of 5 nm was used for excitation and emission. The excitation wavelength was 295 nm and the fluorescence intensity at the emission wavelength in 300–450 nm was monitored.

The thermodynamic parameters for guanidine-induced denaturation curves were obtained by least squares fitting using two-state equations (7,8).

$$s_{\text{obs}} = \frac{(a_N + b_N[\text{GdHCl}]) + (a_U + b_U[\text{GdHCl}])\exp\left(\frac{-\Delta G_{\text{app}} + m_{\text{app}}[\text{GdHCl}]}{RT}\right)}{1 + \exp\left(\frac{-\Delta G_{\text{app}} + m_{\text{app}}[\text{GdHCl}]}{RT}\right)} \quad (4-1)$$

The data were fitted using Equation (1), where s_{obs} is the molar ellipticity or fluorescence intensity observed; a_N (a_U) and b_N (b_U) the intercept and slope of the base line at low (N) and high (U) GdHCl concentrations, respectively; R the gas constant; T the absolute temperature in K; $[\text{GdHCl}]$ the GdHCl concentration; and ΔG_{app} and m_{app} the free energy and cooperativity of folding calculated with Equation (4-1), respectively. The errors were estimated using a single-elimination Jack-knife procedure (9).

The fraction of the folded state (f_N) at each pH was calculated using the equation,

$$f_N = (Y - Y_U)/(Y_N - Y_U) \quad (4-2)$$

, where Y is the $[\theta]$ value at 220 nm ($[\theta]_{220}$) at each pH. Y_N and Y_U , corresponding to a_N and a_U indicated above, respectively, are $[\theta]_{220}$ for the folded and unfolded states at pH 7.5 and $[\text{GdHCl}] = 0$, respectively ($Y_N = -5.2 \times 10^3$, $Y_U = -2.5 \times 10^3$) (8).

NMR spectroscopy

NMR spectra were recorded at 25°C with either a DRX600 or Avance III 600 spectrometer, both being equipped with a cryogenic probe and a Z-gradient (Bruker Biospin). Samples were prepared by mixing A3G_C in storage buffer with several measurement buffers having different pH. The measurement buffers were prepared as in the previous section, except that DTT was adjusted to 5 mM and 5% deuterium oxide. The concentration of A3G_C was set at 0.8 μ M, while that of the ssDNA substrates was adjusted to 200 μ M. After mixing A3G_C and ssDNA together, a series of 2D TOCSY with mixing a time of 20 ms were recorded at different reaction time points, by which the deamination reaction of cytosines of ssDNA could be monitored in real-time. Water signal suppression was achieved with 3-9-19 watergate. Resonance assignments of the substrate ssDNAs were made using 2D TOCSY with an aid of 2D ^1H - ^{13}C HSQC spectroscopy and also by comparing the chemical shifts with the assignments for ATTCCCAATT, AAACCCGAAA and AAACCCGA₂₄CCCGTAA, previously studied (2 and Chapter 3). Spectra were processed with NMRPipe (10), and analyzed using SPARKY (11).

The intensities of the peaks of either cytosine or uracil were plotted against time and analyzed as previous section. The data were fit to a single exponential decaying function, using the following equations, by which apparent deamination rate constants were derived:

$$I(t)_C = I_s \exp(-kt) + I_b \quad (4-3)$$

, where I_b corresponds to the baseline intensity of the spectrum; $I_s + I_b$ the signal intensity at time point zero for cytosine; and k the apparent deamination rate constant, which was always obtained through linear least-squares analysis. The value for the apparent deamination rate constant was used as an index of activity.

We defined the P with the following equation:

$$P = \frac{k_{(C6)}}{k_{(C34)}} \quad (4-4)$$

, where $k_{(C6)}$ and $k_{(C34)}$ are the apparent deamination rate constants at positions C6 and C34 of a 38-mer ssDNA (s4-2). P can be regarded as a measure of the magnitude of 3'→5' polarity. The error was estimated as follows. Firstly, the error of the NMR signal intensity was obtained from the base level noise of each NMR spectrum. Then the error of the rate constant was calculated from data sets constructed by Monte Carlo simulation using the error of the NMR signal intensity. Finally, the errors of the relative activity and the P were obtained by means of error propagation calculation.

4-6. References

- (1) Holden, L. G., Prochnow, C., Chang, Y. P., Bransteitter, R., Chelico, L., Sen, U., Stevens, R. C., Goodman, M. F and Chen, X. S. (2008) Crystal structure of the anti-viral APOBEC3G catalytic domain and functional implications. *Nature* **456**, 121–124.
- (2) Furukawa, A., Nagata, T., Matsugami, A., Habu, Y., Sugiyama, R., Hayashi, F., Kobayashi, N., Yokoyama, S., Takaku, H. and Katahira, M. (2009) Structure, interaction and real-time

- monitoring of the enzymatic reaction of wild-type APOBEC3G. *EMBO J.* **28**, 440–451.
- (3) Harjes, S., Solomon, W. C., Li, M., Chen, K. M., Harjes, E., Harris, R. S. and Matsuo, H. (2013) Impact of H216 on the DNA binding and catalytic activities of the HIV restriction factor APOBEC3G. *J. Virol.* **87**, 7008–7014.
 - (4) Furukawa, A., Sugase, K., Morishita, R., Nagata, T., Kodaki, T., Takaori-Kondo, A., Ryo, A., and Katahira, M. (2014) Quantitative analysis of location- and sequence-dependent deamination by APOBEC3G using real-time NMR spectroscopy. *Angew. Chem. Int. Ed. Engl.* **53**, 2349–2352.
 - (5) Kelly, S. M., Jess, T. J. and Price, C. N. (2005) How to study proteins by circular dichroism. *Biochim. Biophys. Acta* **175**, 119–139.
 - (6) Royer, C. A. (2006) Probing protein folding and conformational transitions with fluorescence. *Chem. Rev.* **106**, 1769–1784.
 - (7) Pace, C. N. (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.* **131**, 266–280.
 - (8) Scholtz, J. M., Grimsley, G. R. and Pace, C. N. (2009) Solvent denaturation of proteins and interpretations of the m value. *Methods Enzymol.* **466**, 549–565.
 - (9) Harris, C. D. (1998) Nonlinear least-squares curve fitting with microsoft excel solver. *J. Chem. Educ.* **75**, 119–121.
 - (10) Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293.
 - (11) Goddard, T. D. and Kneller, D. G. (2006) SPARKY 3. *San Francisco: University of California.*

Chapter 5

Real-time monitoring of the deamination reaction of A3G_C in presence of the Vif protein complex

5-1. Introduction

It is reportedly said that the Vif protein destroys A3G mainly through the ubiquitin proteasomal pathway (1–3). A3G_N is supposed to be key for this phenomenon because A3G_N, not A3G_C, is the target of Vif (4–6). Recently, the crystal structure of the Vif protein complex was determined. Interestingly, Britan-Rosich and coworkers reported that a 15-mer fragment of Vif peptide directly inhibits the deaminase activity of full length A3G (7). In particular, 25–39 and 105–119 of Vif drastically reduced the deaminase activity of A3G (7). Recently the Vif protein was successfully purified by co-expression with CBF β , EloB and EloC following co-purification with the N terminal domain of Cul5 (8–10). Guo and coworkers also presented the crystal structure of the Vif-CBF β -EloB-EloC-Cul5 complex (Vif-complex) and found that the Vif protein retains the novel unique fold (Figure 5-1A) (10). The Vif protein is composed of two domains, N-terminal domain and C-terminal domain (1–3). Both domains exhibit strong RNA binding abilities with dissociation constants in the nanomolar range, which enables co-oligomerization with RNA (11). Here we examine purification of the Vif-complex and deaminase activity of A3G_C in the presence of Vif.

5-2. Results

5-2-1 Constructs of the Vif protein complex

Firstly, we constructed a Vif protein complex. We took the crystal structure of the Vif complex as a model for our protein construct (Figure 5-1A) (10). This structure is composed of HIV-1 Vif (1-192), CBF β (1-170), EloB (1-102), EloC (17-112), and Cul5 (12-386), a 1:1:1:1:1 complex (10). Here we used HIV-1 Vif (1-192), CBF β (1-182), EloB (1-118), EloC (17-112), and Cul5 (12-386), as shown in Figure 5-1B. The difference between the two constructs is the protein length, we used full length CBF β and EloB (in the crystal structure, Vif is tagged but not CBF β) (10). A schematic representation of the Vif protein purification strategy is shown in Figure 5-1C (details are given in the Experimental Procedures section). The Vif-CBF β -EloB-EloC complex was co-expressed in *Escherichia coli* and purified, and then mixed with Cul5 individually purified. The mixture was further purified by liquid chromatography. The concentration of the purified Vif protein complex was quantified as the UV absorbance at 280 nm and the band strength on SDS-PAGE. The results of SDS-PAGE are shown in Figure 5-1D. These findings indicate that we obtained a purified Vif-CBF β -EloB-EloC-Cul5 complex, the ratio of components being 1:1:1:1:1.

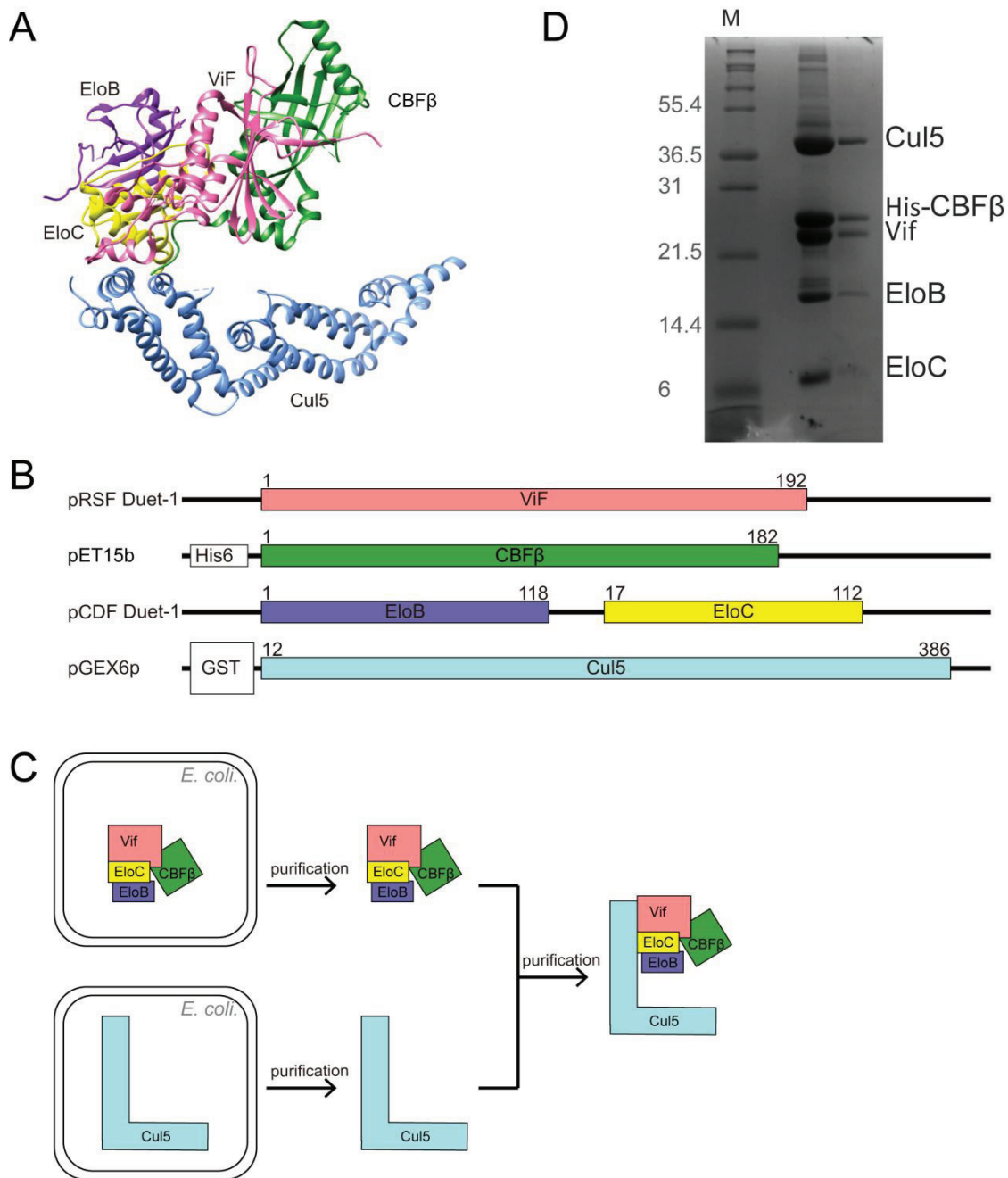


Figure 5-1.

Design of the Vif complex.

(A) Structure of the Vif complex (10). Vif, CBF β , EloB, EloC, and Cul5 are colored pink, green, purple, yellow, and light blue, respectively. Molecular graphics and analyses were performed with the UCSF Chimera package (12). (B, C) Design of the Vif protein complex purification system. The details are shown in the Experimental Procedures section. Schematic representation of the Vif-complex expression plasmid (B), and the purification strategy for the Vif complex (C) are shown. (D) The results of SDS-polyacrylamide gel electrophoresis of the purified Vif complex. The left column is the stock solution (20 μ M) and right column is 11 times diluted one. The gel was stained with Coomassie brilliant blue.

5-2-2 Deamination assaying in the presence of the Vif protein complex

We next evaluated the deaminase activity of A3G_C in presence of the Vif complex. We used two CCC hotspot containing ssDNA AAACCCGA₂₄CCCGTAA (s5, corresponds to s3-3 and s4-2) as a substrate. Figure 5-2A, B show examples of real-time NMR monitoring of the deamination reaction using 2D TOCSY spectra in the absence and presence of the Vif complex (0, 0.8, 11 μ M). Relative activity is plotted at each Vif concentration in Figure 5-2C. It is clear that the deaminase activity toward target C decreases as the Vif concentration increases and the deamination activity was lost at 22 μ M Vif protein. Interestingly, the *P* value also decreased as the Vif concentration increased (Figure 5-2D).

5-3. Discussion

5-3-1 Vif mediated deamination inhibition

It was shown that the Vif protein complex directly inhibits the deaminase activity of A3G_C. Surprisingly, Vif shut out deaminase activity approximately 1/10 of substrate DNA and 25 fold of A3G_C concentration. This kind of inhibition has also been found using denatured full length Vif and 15-mer peptide fragments derived from Vif decreased the activity of full length A3G (7). In their study, Britan-Rosich and coworkers found that at least two kinds of fragment that inhibit deaminase activity. They mentioned that the Vif protein interacts with the ES complex of A3G, which results in a decrease of deaminase activity (7). A3G_N is supposed to possess a binding interface for Vif (4–6). On the other hand, A3G_C may also be able to bind to Vif on mammalian protein-protein interaction trap (MAPPIT) analysis (5). Their suggestion indicated that the binding affinity of A3G_C may be the result of inhibition of deaminase activity. 3'→5' polarity also decreased as the Vif concentration increased, which indicates the Vif complex also blocks sliding of A3G_C. The Vif protein retains strong binding affinity with RNA, which may cause their co-oligomerization (11). The Vif protein has strong positive charge on its molecular surface (10). Thus, it is possible that the Vif complex competitively binds to ssDNA and blocks sliding of A3G_C. Anyway, the Vif protein destroys A3G not only through its proteasomal degradation but also through direct interference with deamination. This finding will be useful to investigating novel anti-HIV-1 drugs using A3G because this kind of drug is required to rescue A3G from Vif binding as well as help it exert deaminase activity. Our real time NMR monitoring method can be used for screening such chemicals that may be new drugs.

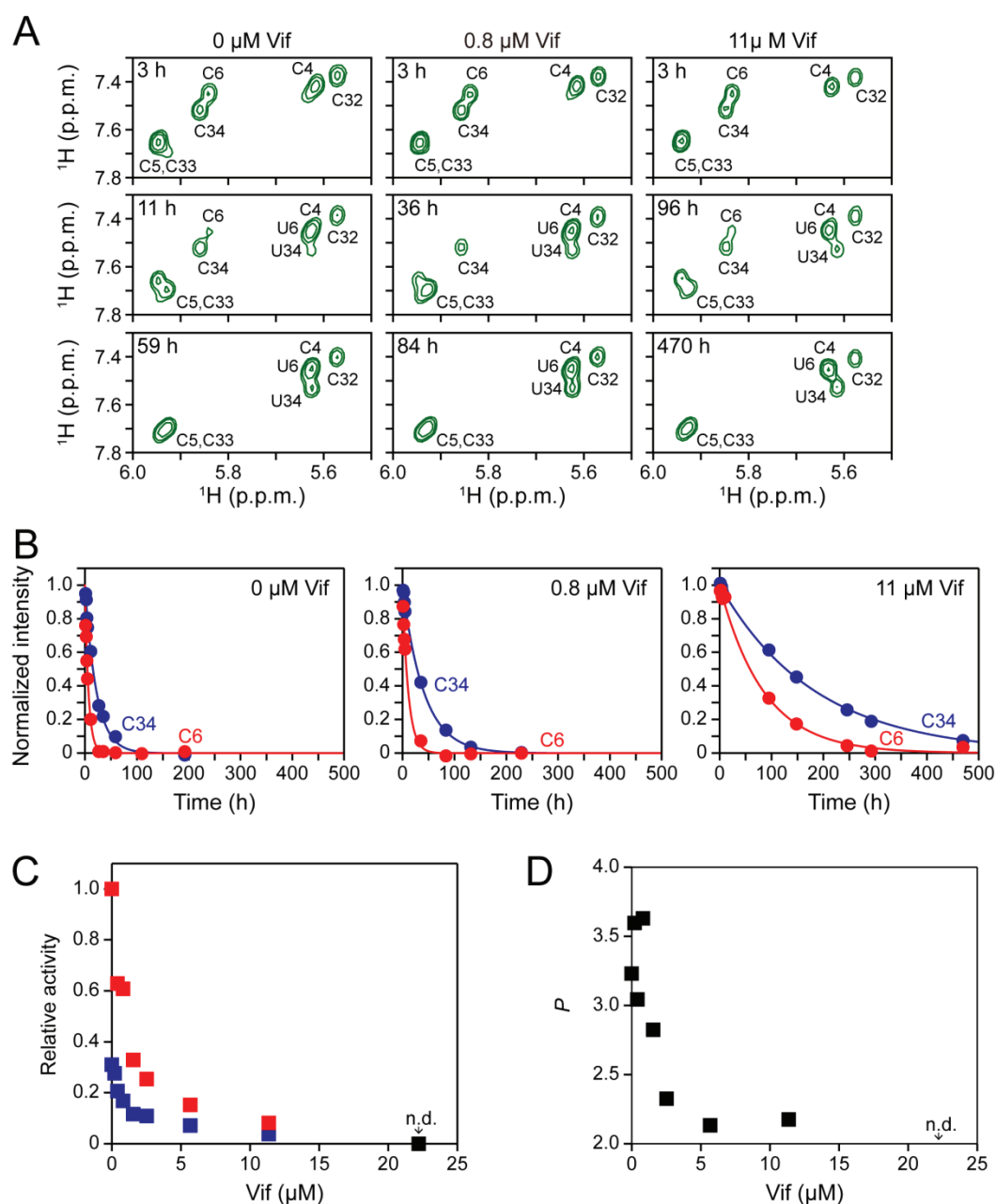


Figure 5-2.

Real-time monitoring of deamination reaction of A3G_C in the presence of the Vif complex.

(A) An example of real-time monitoring of the deamination reaction using 2D TOCSY spectra in the presence of three different Vif complex concentrations (0, 0.8, and 11 μM). The sequence of the substrate ssDNA is AAACCCGA₂₄CCCGTAA (s5). The time point of each spectrum is shown in the panel. (B) The time course of the signal intensity of target C shown in (A). (C) Relative activities for C6 (red) and C34 (blue) are plotted against the Vif complex concentration. The relative activities for C6 and C34 at each pH were defined as the ratio of their k_s at each Vif complex concentration to that obtained for C6 in absence of the Vif complex. (D) 3'→5' polarity values (P) are plotted against the Vif complex concentration. The P at each pH is defined as the ratio of k obtained for C6 and C34.

5-4. Conclusion

We evaluated the deaminase activity of A3G_C in the presence of the Vif protein complex. As the concentration of Vif complex increased, the deaminase activity as well as 3'→5' polarity decreased and finally disappeared completely. These results indicate that the Vif complex may interact with A3G_C and ssDNA. This binding affinity of the Vif complex may block the sliding of A3G_C. This finding indicates that our real-time NMR monitoring method is a powerful tool for screening chemicals for anti HIV-1 drugs that help A3G to exert deaminase activity in the presence of the Vif protein.

5-5. Experimental Procedures

Preparation of A3G_C and DNA substrates

A3G_C was expressed and purified as described previously in Chapter 2. The obtained A3G_C was dialyzed against a storage buffer comprising 20 mM Tris-HCl, pH 7.5, 30 mM NaCl, 5 mM DTT, and 10 μ M ZnCl₂, and then stored until use at 4°C. All substrate ssDNAs were purchased from Fasmac Co., Ltd.

Preparation of the Vif protein complex

The genes of full length Vif, CBF β , EloB, EloC and Cul5 were kindly donated by Dr. Takaori-Kondo and his coworkers. Full length Vif (residues 1–192) was cloned into the pRSF Duet-1 vector. Full length CBF β isoform 2 (residues 1–182) from human was cloned into pET15b with a N-terminal His6-affinity tag and a thrombin cleavage site. Full length EloB (residues 1–118) and truncated EloC (residues 17–112) were cloned into the pCDF Duet-1 plasmid. The N terminal domain of Cul5 (residues 12–386) was cloned into the pGEX6p vector with a prescission protease cleavage site. The constructs used in this chapter and their purification strategy are summarized in Figure 5-1. Vif, CBF β , EloB and EloC were co-expressed, and Cul5 was separately expressed in *Escherichia coli* BL21(DE3) cells. The proteins were cultured in LB medium and over-expressed for 18 hours at 16°C by induction with 0.1 mM IPTG. Harvested cells were lysed in 20 mM Tris-HCl (pH 7.5) with 150 mM NaCl, and then clarified by sonication and centrifugation at 3,000 g for 20 min. The Vif, CBF β , EloB and EloC complex was first purified using Ni Sepharose resin (GE Healthcare). The resin was washed with a solution comprising 20 mM Tris-HCl (pH 7.5), 500 mM NaCl and 30 mM imidazole, and the bound proteins were eluted with a buffer comprising 20 mM Tris-HCl (pH 7.5), 500 mM NaCl and 500 mM imidazole. The Cul5 supernatant was loaded onto glutathione sepharose 4B resin (GE Healthcare). The resin was washed and the bound proteins were cleaved with precision protease (GE Healthcare) overnight at 4°C. The cleaved Cul5 was eluted from the resin with lysis buffer. The Vif, CBF β , EloB and EloC complex, and Cul5 were mixed and dialyzed more than three times against buffer comprising 20 mM Tris-HCl (pH 8.0), 30 mM NaCl, 10 μ M ZnCl₂, 5 mM DTT. The complex was subjected to further purification by anion exchange using Q sepharose HP resin (GE Healthcare) with a linear NaCl gradient, 30–1000 mM, at 4°C. The elute

was loaded onto Ni Sepharose resin (GE Healthcare) equilibrated with a solution comprising 20 mM Tris-HCl (pH 7.5), 500 mM NaCl, 10 μ M ZnCl₂, 5 mM DTT and 30 mM imidazole. The Vif protein complex was eluted with a linear imidazole gradient, 30–500 mM. The purified proteins were dialyzed more than three times against storage buffer comprising 20 mM Tris-HCl (pH 7.7), 30 mM NaCl, 10 μ M ZnCl₂ and 5 mM DTT. Finally, the obtained Vif complex was passed through a syringe filter with a 0.45 μ m pore size PES membrane, and then stored until use at 4°C.

NMR spectroscopy

NMR spectra were recorded at 25°C with a DRX600 spectrometer equipped with a cryogenic probe and a Z-gradient (Bruker Biospin). A final concentration of 200 μ M each DNA substrate was dissolved in 20 mM Tris-HCl (pH 7.7), 30 mM NaCl, 10 μ M ZnCl₂, 5 mM DTT, and 5% deuterium oxide. A final concentration of 0.8 μ M A3G_C was used. The Vif complex was added to DNA dissolved in the measurement buffer, and then A3G was mixed with the solution. The concentrations of the proteins were confirmed by combining the UV absorbance at 280 nm and the band strength on SDS-PAGE. The following analyses were performed as described in Chapter 3.

5-6. References

- (1) Aydin, H., Taylor, M. W. and Lee, J. E. (2014) Structure-guided analysis of the human APOBEC3-HIV restrictome. *Structure* **22**, 668–684.
- (2) Batisse, J., Guerrero, S., Bernacchi, S., Sleiman, D., Gabus, C., Darlix, J. L., Marquet, R., Tisné, C. and Paillart, J. C. (2012) The role of Vif oligomerization and RNA chaperone activity in HIV-1 replication. *Virus Res.* **169**, 361–376.
- (3) Desimmie, B. A., Delviks-Frankenberry, K. A., Burdick, R. C., Qi, D., Izumi, T. and Pathak, V. K. (2014) Multiple APOBEC3 restriction factors for HIV-1 and one Vif to rule them all. *J. Mol. Biol.* **426**, 1220–1245.
- (4) Mehle, A., Wilson, H., Zhang, C., Brazier, A. J., McPike, M., Pery, E. and Gabuzda, D. (2007) Identification of an APOBEC3G binding site in human immunodeficiency virus type 1 Vif and inhibitors of Vif-APOBEC3G binding. *J. Virol.* **81**, 13235–13241.
- (5) Uyttendaele, I., Lavens, D., Catteuw, D., Lemmens, I., Bovijn, C., Tavernier, J. and Peelman, F. (2011) Random mutagenesis MAPPIT analysis identifies binding sites for Vif and Gag in both cytidine deaminase domains of Apobec3G. *PLoS One* **7**, doi: 10.1371/journal.pone.0044143.
- (6) Letko, M., Booiman, T., Kootstra, N., Simon, V. and Ooms, M. (2015) Identification of the HIV-1 Vif and human APOBEC3G protein interface. *Cell Rep.* **13**, 1789–1799.
- (7) Britan-Rosich, E., Nowarski, R. and Kotler, M. (2011) Multifaceted counter-APOBEC3G mechanisms employed by HIV-1 Vif. *J. Mol. Biol.* **410**, 1065–1076.
- (8) Zhou, X., Evans, S. L., Han, X., Liu, Y. and Yu, X. F. (2012) Characterization of the interaction of full-length HIV-1 Vif protein with its key regulator CBF β and CRL5 E3 ubiquitin ligase components. *PLoS One* **7**, doi: 10.1371/journal.pone.0033495.

- (9) Kim, D. Y., Kwon, E., Hartley, P. D., Crosby, D. C., Mann, S., Krogan, N. J. and Gross, J. D. (2013) CBF β stabilizes HIV Vif to counteract APOBEC3 at the expense of RUNX1 target gene expression. *Mol. Cell* **49**, 632–644.
- (10) Guo, Y., Dong, L., Qiu, X., Wang, Y., Zhang, B., Liu, H., Yu, Y., Zang, Y., Yang, M. and Huang, Z. (2014) Structural basis for hijacking CBF- β and CUL5 E3 ligase complex by HIV-1 Vif. *Nature* **505**, 229–233.
- (11) Sleiman, D., Bernacchi, S., Xavier, Guerrero, S., Brachet, F., Larue, V., Paillart, J. C. and Tisne, C. (2014) Characterization of RNA binding and chaperoning activities of HIV-1 Vif protein. Importance of the C-terminal unstructured tail. *RNA Biol.* **11**, 906–920.
- (12) Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612.

Chapter 6

Deamination activity of A3G_C toward 5-methylcytidine or 5-hydroxymethylcytidine

6-1. Introduction

Among the DNA modifications, methylation of cytidine at its C-5 position is known as a major epigenetic mechanism linked to gene regulation in development and in tumorigenesis (1–5), while DNA demethylation is one of the crucial processes for genome reprogramming during early embryogenesis (1–7). Several pathways for DNA demethylation have been suggested, summarized in Figure 6-1, one of which involves 5-methylcytidine (5mC) to 5-hydroxymethylcytidine (5hmC) conversion by Ten-eleven translocation (Tet) proteins (1–9), deamination of 5hmC by several deaminases (APOBEC family proteins) (2–4, 10), and subsequent by base excision repair (BER), such as replacement of a 5-hydroxymethyluracil-guanine mispair by MBD4 or thymine-DNA glycosylase (TDG) (11, 12). APOBEC family proteins were hypothesized to participate in the demethylation pathways via deamination of 5mC and 5hmC, many of the APOBEC family proteins were shown to be poor enzymes for deamination of 5mC and 5hmC (13, 14). Among the APOBEC family proteins, it was shown recently in vitro, using a DNA glycosylase-based assay, and in vivo that APOBEC3A (A3A) exhibits deamination activity toward C and 5mC (15, 16), but that A3G exhibits activity only toward C (15, 16). Wijesinghe *et al.* showed that A3G exhibits no detectable deamination activity toward 5mC using an in vivo method, the so-called Kanamycin-resistance reversion assay, which is a genetically modified *Escherichia coli*-based method (15). Carpenter *et al.* used DNA glycosylase-based methods to investigate the activity for 5mC to T conversions by A3A and A3G prepared from human cells (16). It was shown that although A3A can deaminate 5mC, A3G cannot deaminate 5mC within the experimental time (120 min). Here, we applied our real-time NMR method to detect, if any, the deamination activity of A3G_C toward 5mC or 5hmC. Because the real-time NMR method monitors the perturbation of the signal of 5mC or 5hmC itself, we expected to observe any low activity.

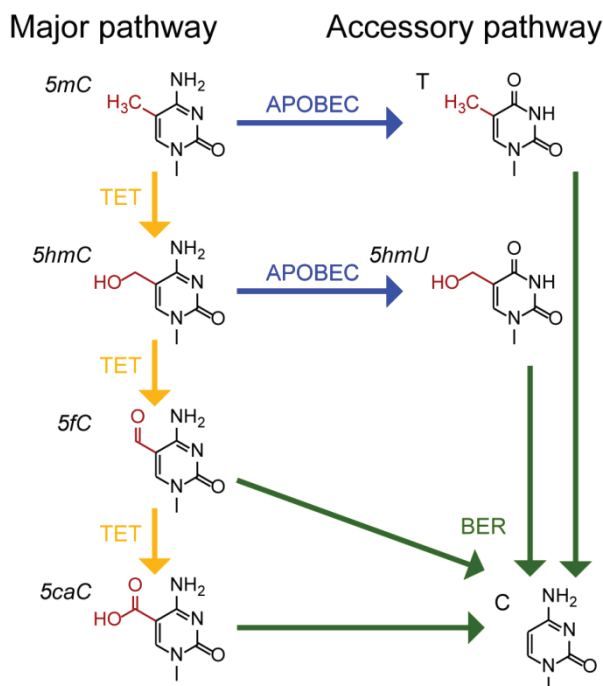


Figure 6-1.

Proposed pathways for 5mC demethylation.

5-methylcytosine (5mC) is converted into C by chemical modification following to DNA miss-math repair by DNA repairing system using base edition repair (BER) (1–9). 5mC are oxidized by ten-eleven translocation (TET) proteins and converted into 5hmC, 5fC, and then 5CaC (1–9). APOBEC family proteins are reportedly involved in this pathway by deaminating 5mC and 5-hydroxymethylcytosine (5hmC) into Thymine (T) or 5-hydroxymethyluracil (5hmU), respectively.

6-2. Results

6-2-1. Deaminase activity toward 5mC

Substrates containing 5mC or 5hmC at the deamination target site were subjected to real-time monitoring to investigate the activity of A3G_C toward them by direct observation. We used 5'-AAAAC5mCGAAA-3' (s6-1_5mC) as a substrate, whose standard sequence is 5'-AAAACCGAAA-3' (s6-1_C) and which contains two cytidines, one of which is the target (underlined). Figure 6-2A presents the ^1H - ^{13}C HSQC spectrum for the methyl groups of 5mC and thymine (T), in which the 5mC6 and T6 signals are labeled, respectively. The spectrum at 1,900 hours after the addition of A3G_C (Figure 6-2A, middle panel) clearly shows that 5mC of s6-1_5mC is converted to thymidine to some extent, since a T6 signal is present, as seen for a control DNA, s6-1_T (Figure 6-1A, right panel). This is the first time to detect the deamination activity of A3G toward 5mC. It is assumed difficult for the in vivo method to detect such weak activity.

6-2-2. Deaminase activity toward 5hmC

Next, we investigated the reactivity of substrates containing 5hmC for the deaminase reaction. 5'-AAAAC5hmCGAAA-3' (s6-1_5hmC) was subjected to real-time monitoring (Figure 6-2B). The ^1H - ^{13}C HSQC spectra for the methylene groups of 5hmC and 5-hydroxymethyluracil (5hmU) are presented in Figure 6-2. The spectrum at 1,700 hours after the addition of A3G_C (Figure 6-2B, middle panel) was exactly the same as that at time zero (Figure 6-2B, left panel), which is different from that of a control DNA, s6-1_5hmU (Figure 6-2B, right panel). The measurement was continued as long as 3,000 hours, however, no change in the spectral pattern was observed. Thus, our results explicitly showed that A3G_C has failed to deaminate 5hmC. We then tested deaminase activity of 3G_C toward 5'-ATTCC5hmCGATT-3' (s6-2), of which contains CC5hmC as a hotspot. 3G_C deaminated fifth cytosine of s6-2 into uracil and no deamination activity were detected toward 5hmC into 5hmU (data not shown).

6-3. Discussion

6-3-1. The effects on 5' modification for deaminase activity of A3G_C

Figure 6-2C compares the relative activity values for s6-1_C, s6-1_5mC, and s6-1_5hmC. The difference in structure between cytidine, 5mC, and 5hmC is the steric bulkiness of the base moiety, which increases in that order. Nabel and coworkers synthesized a series of substrates with unnatural C5 substituents of varied steric and electronic character, and investigated the reactivity of human activation-induced deaminase (AID) which is required for class-switch recombination, somatic hypermutation B-cell tolerance (17–19), and mouse APOBECs (1, 2, and 3) by means of a DNA glycosylase-based assay (14). They concluded that deamination decreases with increasing steric bulkiness at the 5 position of cytidine. We assume that this is also the case for 3G_C, in other words, 3G_C's activity also decreases with increasing steric bulkiness at the 5 position of cytidine.

6-4. Conclusion

We attempted to detect the activity of A3G_C toward an ssDNA containing 5mC or 5hmC at the targeted cytidine site by direct observation. We showed for the first time that A3G_C can convert 5mC into thymidine, although it cannot convert 5hmC into 5hmU. Additionally A3G failure to deaminate 5hmC within CCC hotspot, CC5hmC, instead of generating CU5hmC. Altogether we have observed A3G_C's deamination reaction directly and unambiguously, and thus we have successfully showed that the combination of a real-time NMR method and a chemically modified nucleotide substrates is a powerful strategy for analyzing deamination activity.

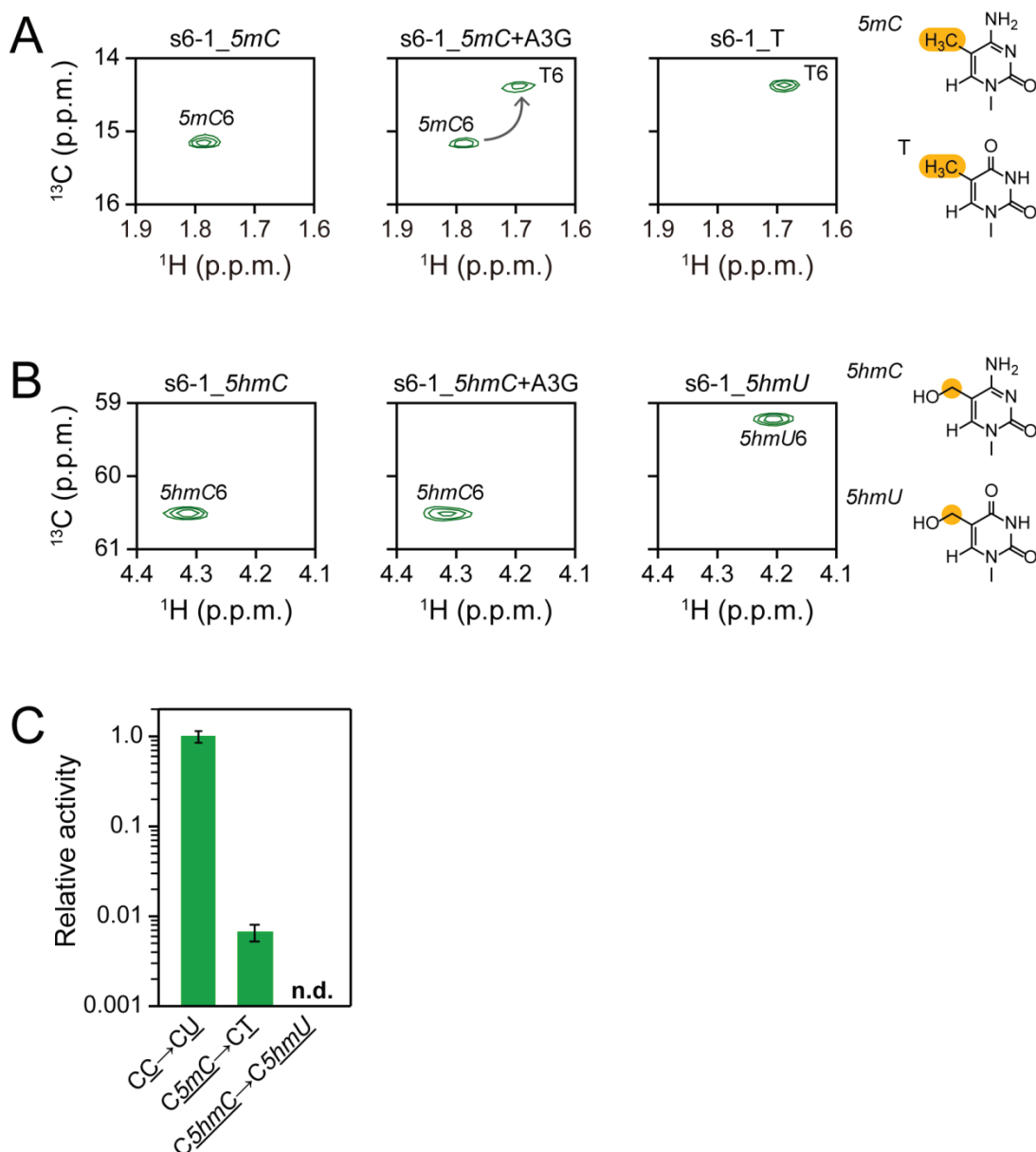


Figure 6-2.

Real-time monitoring of deamination reactions of substrates containing 5-methylcytidine or 5-hydroxymethylcytidine at the target position.

(A) ^1H – ^{13}C HSQC spectrum of an ssDNA substrate, AAAAC5mCGAAA (left), and that recorded 1,900 hours after addition of A3G_C (center). The ^1H – ^{13}C HSQC spectrum of AAAACTGAAA is shown as a reference (right). (B) ^1H – ^{13}C HSQC spectrum of an ssDNA substrate (top), AAAAC5hmCGAAA (left), and that recorded 1,700 hours after addition of A3G_C (center). The ^1H – ^{13}C HSQC spectrum of AAAAC5hmUGAAA is shown as a reference (right). (C) Comparison of the relative activity between the reactions of three substrates, AAAACCGAAA, AAAAC5mCGAAA, and AAAAC5hmCGAAA.

6-5 Experimental Procedures

Preparation of the protein and DNA substrates

A3G_C was expressed and purified as described previously in Chapter 2. The obtained A3G_C was dialyzed against a storage buffer comprising 20 mM Tris-HCl, pH 7.5, 30 mM NaCl, 5 mM DTT, and 10 μ M ZnCl₂, and then stored until use at 4°C. All substrate ssDNAs, a 10-mer ssDNA used in this chapter, were purchased from Fasmac Co., Ltd. The sequence of substrate DNAs are listed in table 6-1.

Table 6-1. Substrate ssDNAs used in this chapter.

Substrate name	Sequence	Length	Position of target
s6-1_C	AAAACCGAAA	10	6 (C)
s6-1_5mC	AAAAC5mCGAAA	10	6 (5mC)
s6-1_5hmC	AAAAC5hmCGAAA	10	6 (5hmC)
s6-1_T	AAAAC TGAAA	10	6 (5hmC)
s6-1_5hmU	AAAAC5hmUGAAA	10	6 (5hmC)
s6-2	ATTCC5hmCGATT	10	5 (C) and 6 (5hmC)

NMR spectroscopy

All NMR spectra were recorded at 25°C with a DRX600 spectrometer equipped with a cryogenic probe and a Z-gradient (Bruker Biospin). A final concentration of 200 μ M each DNA substrate was dissolved in 20 mM Tris-HCl (pH 7.5), 30 mM NaCl, 10 μ M ZnCl₂, 5 mM DTT, and 5% deuterium oxide. A final concentration of 20.0 or 70.0 μ M A3G_C was used. The following analyses were performed as described in Chapter 3.

6-6. References

- (1) Jaenisch, R. and Bird, A. (2003) Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat. Genet.* **33**, 245–254.
- (2) Bhutani, N., Burns, D. M., Blau, H. M. (2011) DNA demethylation dynamics. *Cell* **146**, 866–872.
- (3) Teperek-Tkacz, M., Pasque, V., Gentsch, G., Ferguson-Smith, A. C. (2011) Epigenetic reprogramming: is deamination key to active DNA demethylation? *Reproduction* **142**, 621–632.
- (4) Chen, Z. X. and Riggs, A. D. (2011) DNA methylation and demethylation in mammals. *J. Biol. Chem.* **286**, 18347–18353.
- (5) Nabel, C. S., Manning, S. A. and Kohli, R. M. (2012) The curious chemical biology of cytosine: deamination, methylation, and oxidation as modulators of genomic potential. *ACS Chem. Biol.* **7**, 20–30.

- (6) Hill, P. W., Amouroux, R. and Hajkova, P. (2014) DNA demethylation, Tet proteins and 5-hydroxymethylcytosine in epigenetic reprogramming: an emerging complex story. *Genomics* **104**,324–33.
- (7) Schübeler D. (2015) Function and information content of DNA methylation. *Nature* **517**, 321–326.
- (8) Williams, K., Christensen, J. and Helin, K. (2011) DNA methylation: TET proteins-guardians of CpG islands? *EMBO Rep.* **13**, 28–35.
- (9) Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L. M., Liu, D. R., Aravind, L. and Rao, A. (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930–935.
- (10) Guo, J. U., Su, Y., Zhong, C., Ming, G. L. and Song, H. (2011) Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell* **145**, 423–434.
- (11) Zhu, J. K. Active DNA demethylation mediated by DNA glycosylases. (2009) *Annu. Rev. Genet.* **43**, 143–166.
- (12) Cortellino, S., Xu, J., Sannai, M., Moore, R., Caretti, E., Cigliano, A., Le Coz, M., Devarajan, K., Wessels, A., Soprano, D., Abramowitz, L. K., Bartolomei, M. S., Rambow, F., Bassi, M. R., Bruno, T., Fanciulli, M., Renner, C., Klein-Szanto, A. J., Matsumoto, Y., Kobi, D., Davidson, I., Alberti, C., Larue, L. and Bellacosa, A. (2011) Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell* **146**, 67–79.
- (13) Larijani, M., Frieder, D., Sonbuchner, T. M., Bransteitter, R., Goodman, M. F., Bouhassira, E. E., Scharff, M. D. and Martin, A. (2005) Methylation protects cytidines from AID-mediated deamination. *Mol. Immunol.* **42**, 599–604.
- (14) Nabel, C. S., Jia, H., Ye, Y., Shen, L., Goldschmidt, H. L., Stivers, J. T., Zhang, Y. and Kohli, R. M. (2012) AID/APOBEC deaminases disfavor modified cytosines implicated in DNA demethylation. *Nat. Chem. Biol.* **8**, 751–758.
- (15) Wijesinghe, P. and Bhagwat, A. S. (2012) Efficient deamination of 5-methylcytosines in DNA by human APOBEC3A, but not by AID or APOBEC3G. *Nucleic Acids Res.* **40**, 9206–9017.
- (16) Carpenter, M. A., Li, M., Rathore, A., Lackey, L., Law, E. K., Land, A. M., Leonard, B., Shandilya, S. M., Bohn, M. F., Schiffer, C. A., Brown, W. L. and Harris, R. S. Methycytosine and normal cytosine deamination by the foreign DNA restriction enzyme APOBEC3A. *J. Biol. Chem.* **287**, 34801–34808.
- (17) Harris, R. S. and Liddament, M. T. (2004) Retroviral restriction by APOBEC proteins. *Nat. Rev. Immunol.* **4**, 868–867.
- (18) Kumar, R., DiMenna, L. J., Chaudhuri, J. and Evans, T. (2014) Biological function of activation-induced cytidine deaminase (AID). *Biomed. J.* **37**, 269–283.
- (19) Meyers, G., Ng, Y. S., Bannock, J. M., Lavoie, A., Walter, J. E., Notarangelo, L. D., Kilic, S. S.,

Aksu, G., Debré, M., Rieux-Laucat, F., Conley, M. E., Cunningham-Rundles, C., Durandy, A. and Meffre, E. (2011) Activation-induced cytidine deaminase (AID) is required for B-cell tolerance in humans. *Proc. Natl. Acad. Sci. USA*. **108**, 11554–11559.

Chapter 7

General conclusion

A3G is an anti-HIV factor that inactivates HIV-1 by deaminating ssDNA reverse-transcribed from the viral RNA genome. The C-terminal domain of A3G (A3G_C) is responsible for the deamination. Although several structures of A3G_C in its free form have been reported, the structure of the A3G_C-ssDNA complex has yet to be elucidated because of the weak interaction. The deamination activity of A3G_C is highly sequence-specific; a triplet cytosine (CCC hot spot) is the most favorable target. Full length A3G as well as A3G_C exhibit a deamination bias called 3'→5' polarity; a CCC hotspot located closer to the 5'-end is deaminated more effectively than one less close to the 5'-end. This phenomenon has been explained by that A3G slides along ssDNA with no directional preference and that the catalytic activity depends on the direction of sliding. Here, we developed a real-time NMR monitoring method and characterized the mechanism of deamination by A3G.

In Chapter 2, we demonstrated the methodology of a real-time NMR monitoring method involving 2D TOCSY spectra, and that it is a powerful tool for evaluating the deaminase activity of A3G_C. The deamination activity of A3G_C can be evaluated by fitting the time course of NMR signal intensity with a single exponential function even if the substrate concentration is much higher than that of the enzyme.

In Chapter 3, we characterized the deaminase activity of A3G_C toward DNA oligonucleotides containing a nucleotide analog at a single or multiple positions. Firstly, A3G_C was shown to require the 5 nucleotide recognition sequence, the hotspot being positioned at the center. Secondly, analyses involving DNA oligonucleotides having two CCC hotspots connected by linker, which is composed of either DNA, RNA or abasic DNA, suggested that the sugar and base moieties are not involved in the sliding of A3G_C but exert the 3'→5' polarity. Thirdly, catalytic analysis of different NaCl concentrations indicated that the activity is bell-shaped against NaCl concentration, while the lower the NaCl concentration, the higher the 3'→5' polarity. These results suggest that the electrostatic interaction between A3G_C and the phosphate backbone within ssDNA is crucial for sliding.

These results are confirmed in Chapter 4. In this chapter, we investigated the pH effects on the structure, deaminase activity and 3'→5' polarity of A3G_C. A3G_C was shown to retain the native structure in the pH4.0–pH 10.5 range by inspecting the pattern of CD spectra. Deamination assaying against 10-mer ssDNA containing a single CCC hot spot, of various pHs revealed that A3G retains its deamination activity in the pH 6.5–pH 12.7 range and the deaminase activity increases as the pH increases in this range. The analyses involving 38-mer ssDNA having two CCC hotspots connected by a linker showed similar pH preference of deamination activity to that of the 10-mer DNA. Similar to the results reported in Chapter 3, the effect of NaCl concentration, i.e. the 3'→5' polarity, is higher at the lower pH. This suggests that A3G_C dissociates from ssDNA during the

sliding at higher pH due to the weakened electrostatic interaction, while A3G_C continues to slide without abortion at lower pH due to a strengthened electrostatic interaction although the sliding speed itself is supposed to be slower.

In Chapter 5, we purified the Vif-CBF β -EloB-EloC-Cul5 complex and then monitored the deamination reaction of A3G_C in real-time using NMR in presence of the Vif protein complex. As the concentration of the Vif complex increased, the deaminase activity as well as 3'→5' polarity decreased and finally disappeared completely. These results indicate that the Vif complex may be interact with A3G_C and ssDNA. We demonstrated that real-time NMR monitoring method is even powerful in presence of other molecule such as inhibitor.

In Chapter 6, we applied this method to estimate the deamination activity of A3G_C toward 5mC or 5hmC, which is a cytosine derivative supposed to be involved in epigenetic gene regulation. It was demonstrated that A3G_C can deaminate 5mC for the first time and no detectable deaminase activity was observed toward 5hmC. In this chapter we demonstrated that this method may be useful for evaluating other chemical reactions such as the 5mC oxidation catalyzed by the TET protein.

In this study, we mainly used 2D TOCSY or ^1H - ^{13}C HSQC spectra with no labeling with stable isotope with spin quantum number 1/2 for experiments. To achieve higher temporal resolution, one can introduce ^{13}C -labeled cytosines to the substrate ssDNA and obtain a series of 2D ^1H - ^{13}C HSQC spectra. Our real-time NMR monitoring can be applied to other enzymes or complicated systems. This method is also useful for characterizing the mechanism of the enzyme and may be useful for discovering anti-HIV-1 drugs that inactivate Vif and help A3G to exert its deaminase activity.

List of publications

Publications that constitute this thesis

1. Catalytic analysis of APOBEC3G involving real-time NMR spectroscopy reveals nucleic acid determinants for deamination.
Keisuke Kamba, Takashi Nagata and Masato Katahira
PLoS One, (2015) volume 10 (Issue 4), e0124142
(Chapter 3, 6)
2. Characterization of the deamination coupled with sliding along DNA of anti-HIV factor APOBEC3G on the basis of the pH-dependence of deamination revealed by real-time NMR monitoring.
Keisuke Kamba, Takashi Nagata and Masato Katahira
Frontiers in Microbiology, (2016) volume 7 (Issue 587), doi: 10.3389/fmicb.2016.00587
(Chapter 2, 4)
3. Vif complex castrates A3G by direct inhibition of deaminase activity as well as its sliding.
Keisuke Kamba, Takashi Nagata and Masato Katahira
In preparation.
(Chapter 5)

Presentations

International Conference

1. The stability and folding/unfolding of staphylococcal nuclease at the residue level.
○Shun Terauchi, Keisuke Kamba, Yoshiharu Mori, Yoshitake Sakae, Takashi Nakamura, Yuko Okamoto, Kunihiro Kuwajima and Kosuke Maki
The IGER International Symposium on Science of Molecular Assembly and Biomolecular Systems 2014, March 12–13, 2014, Nagoya, Japan
2. Functional mechanism of an anti-HIV protein, APOBEC3G, as revealed by real-time NMR monitoring of its enzymatic reaction.
Keisuke Kamba, Ayako Furukawa, Kenji Sugase, Ryo Morishita, Takashi Nagata and ○Masato Katahira
The XXVIth International Conference on Magnetic Resonance in Biological Systems, August 24–29, 2014, Dallas, Texas, USA
3. Deamination mechanism of an anti-HIV factor APOBEC3G coupled with sliding on single-stranded DNA as analyzed by real-time NMR monitoring method.
○Keisuke Kamba, Ayako Furukawa, Kenji Sugase, Takashi Nagata and Masato Katahira
The 41st International Symposium on Nucleic Acids Chemistry, November 5–7, 2014, Fukuoka, Japan
4. Real-time NMR monitoring of enzymatic reaction of anti-HIV protein, structure of anti-prion RNA aptamer and wood biomass analysis.
Keisuke Kamba, Tsukasa Mashima, Hideyasu Okamura, Ayako Furukawa, Kenji Sugase, Hiroshi Nishimura, Takashi Watanabe, Takashi Nagata, and ○Masato Katahira
The 5th Japan-Taiwan NMR symposium, September 28–30, 2014, Hokkaido, Japan
5. Real-time NMR methods reveals deamination mechanism by human antiviral factor APOBEC3G.
○Keisuke Kamba, Takashi Nagata and Masato Katahira
6th Asia-Pacific NMR Symposium, August 13–16, 2015, Hong Kong, China
6. Nucleic acid determinants of deamination by human anti-HIV factor, APOBEC3G, as revealed by real-time NMR method.
○Keisuke Kamba, Takashi Nagata and Masato Katahira
The 42nd International Symposium on Nucleic Acids Chemistry, September 23–25, 2015, Hyogo, Japan

Domestic Conference

1. NMR 法による APOBEC3C デアミナーゼ活性に関する研究
○岩岡諒、北村紳悟、神庭圭佑、古川亜矢子、岡村英保、杉浦互、永田崇、岩谷靖雅、片平正人
第 50 回日本生物物理学会年会、2012 年 9 月 22–24 日、愛知県
2. ヒト抗ウイルス因子 APOBEC3G の脱アミノ化反応の NMR リアルタイムモニタリング
○神庭圭佑、古川亜矢子、小瀧努、永田崇、片平正人
第 35 回日本分子生物学会年会、2012 年 12 月 11–14 日、福岡県
3. 宿主遺伝子発現を制御するヒト抗ウイルス因子 APOBEC3G の脱アミノ化反応の NMR 実時間計測
○神庭圭佑
京都大学グローバル COE プログラム・地球温暖化時代のエネルギー科学拠点
平成 24 年度年次報告会、2013 年 1 月 16 日、京都府
4. スタフィロコッカル・ヌクレアーゼの残基ごとの安定性とフォールディング/アンフォールディングの研究
○寺内駿、神庭圭佑、森義治、榮慶丈、中村敬、岡本祐幸、桑島邦博、槇互介
2012 年度生物物理学会中部支部講演会、2013 年 2 月 19 日、愛知県
5. 核磁気共鳴法とレプリカ交換分子動力学シミュレーションによるスタフィロコッカル・ヌクレアーゼの安定性とフォールディング/アンフォールディングの研究
○寺内駿、神庭圭佑、森義治、榮慶丈、中村敬、岡本祐幸、桑島邦博、槇互介
第 68 回日本物理学会年次大会、2013 年 3 月 26–29 日、広島県
6. NMR 実時間計測法により抗ウイルス因子 APOBEC3G のシチジン脱アミノ化反応の配列依存性を評価する
○神庭圭佑、小瀧努、永田崇、片平正人
第 12 回日本蛋白質科学会年回、2013 年 6 月 12–14 日、鳥取県
7. スタフィロコッカル・ヌクレアーゼの安定性とフォールディング/アンフォールディングの研究
○寺内駿、神庭圭佑、森義治、榮慶丈、中村敬、岡本祐幸、桑島邦博、槇互介
第 51 回日本生物物理学会年会、2013 年 10 月 28–30 日、京都府
8. 抗ウイルス因子 APOBEC3G の DNA 認識機構の解明
～NMR 実時間計測法によるアプローチ～
○神庭圭佑、永田崇、片平正人
第 52 回 NMR 討論会、2013 年 11 月 12–14 日、石川県

9. ヒト抗ウイルス因子 APOBEC3G の基質認識機構及び Vif による APOBEC3G 阻害機構の解明に向けて
○神庭圭佑、永田崇、片平正人
第 36 回日本分子生物学会年会、2013 年 12 月 3–6 日、兵庫県
10. 核磁気共鳴法と分子動力学シミュレーションによるスタフィロコッカル・ヌクレアーゼの水素結合形成の研究
○寺内駿、神庭圭佑、森義治、榮慶丈、中村敬、岡本祐幸、桑島邦博、槇互介
第 69 回日本物理学会年次大会、2014 年 3 月 27–30 日、神奈川県
11. 抗 HIV 因子 APOBEC3G の基質認識及びスライディング機構の実時間 NMR 解析
○神庭圭佑、永田崇、片平正人
第 52 回日本生物物理学会年会、2014 年 9 月 25–27 日、北海道
12. 実時間 NMR 法による抗 HIV 因子 APOBEC3G のスライディングにカップルした脱アミノ化機構の解析
○神庭圭佑、永田崇、片平正人
第 37 回日本分子生物学会年会、2014 年 11 月 25–27 日、神奈川県
13. ヒト抗 HIV 因子 APOBEC3G と 1 本鎖 DNA の相互作用機構
○神庭圭佑
新学術領域研究「理論と実験による柔らかな分子系の機能の科学」
第 2 回公開シンポジウム、2014 年 11 月 28–29 日、大阪府
14. 効率よくウイルス DNA に変異を導入する APOBEC3G の脱アミノ化機構
○神庭圭佑、永田崇、片平正人
第 53 回日本生物物理学会年会、2015 年 9 月 13–15 日、石川県
15. リアルタイム NMR 法の新たな展開
ー抗 HIV タンパク質 APOBEC3G の認識ヌクレオチド、DNA 上のスライディング及びエピジェネティクスとの関連に関する新知見ー
○神庭圭佑、永田崇、片平正人
第 54 回 NMR 討論会、2015 年 11 月 6–8 日、千葉県
16. NMR による活性モニタリングに基づいたヒト抗 HIV 因子 APOBEC3G の脱アミノ化及び動作機構の解明
○神庭圭佑
新学術領域研究「理論と実験による柔らかな分子系の機能の科学」第 4 回全体合宿会議、2015 年 11 月 24–26 日、福岡県

17. Analysis of the biochemical characteristics of APOBEC3F by real-time NMR

○Wan li, Keisuke Kamba, Takashi Nagata and Masato Katahira

第 38 回日本分子生物学会年会・第 88 回日本生化学会大会合同大会、
2015 年 12 月 1-4 日、兵庫県

18. NMR による実時間酵素活性アッセイにより抗ウイルス因子 APOBEC3G の脱アミノ化機構を明らかにする

○神庭圭佑、永田崇、片平正人

第 38 回日本分子生物学会年会・第 88 回日本生化学会大会合同大会、
2015 年 12 月 1-4 日、兵庫県

List of Honors

1. 新学術領域研究「理論と実験による柔らかな分子系の機能の科学」
若手研究者海外派遣プログラム 採択
Functional mechanism of an anti-HIV protein, APOBEC3G, as revealed by real-time NMR monitoring of its enzymatic reaction.
The XXVIth International Conference on Magnetic Resonance in Biological Systems,
August 24–29, 2014, Dallas, Texas, USA
2. 日本核磁気共鳴学会・平成 27 年度第 1 回若手研究者渡航費助成金 採択
Real-time NMR methods reveals deamination mechanism by human antiviral factor APOBEC3G.
6th Asia-Pacific NMR Symposium, August 13–16, 2015, Hong Kong, China
3. 第 54 回 NMR 討論会 若手ポスター賞・太陽日酸賞
リアルタイム NMR 法の新たな展開
ー抗 HIV タンパク質 APOBEC3G の認識ヌクレオチド、DNA 上のスライディング及び
エピジェネティクスとの関連に関する新知見ー
2015 年 11 月 7 日、千葉県
4. 新学術領域研究「理論と実験による柔らかな分子系の機能の科学」第 4 回全体合宿会議
若手招待講演
NMR による活性モニタリングに基づいたヒト抗 HIV 因子 APOBEC3G の脱アミノ化
及び動作機構の解明
2015 年 11 月 26 日、福岡県
5. 平成 27 年度京都大学エネルギー理工学研究所表彰 学生賞
実時間 NMR 法の開発及びヒト抗 HIV 因子 APOBEC3G の相互作用機構の解明
2016 年 3 月 4 日、京都府

Acknowledgment

This work would not have been possible without the advice and assistance of many others. First and foremost, I would like to express my sincere gratitude to supervisor, Prof. Masato Katahira, for providing me this precious study opportunity as a Ph. D student in his laboratory. I was grateful to him for giving me a chance to attend and present my works at many conferences. I have improved as a scientist with his guidance, valuable suggestions, responsible discussions and kind assistance throughout this work. I especially would like to express my deep appreciation to Associate Professor Takashi Nagata for his elaborated guidance and individual discussion that make my research of great achievement. I was able learn a lot as a scientist including experimental techniques and writing, thanks to his help.

I am deeply grateful to Associate Professor Tsutomu Kodaki for helps suggestions and discussion which encourage me and give me many ideas. I sincerely thank Assistant Professor Tsukasa Mashima for his kindly overall support for experiment such as nucleic acid handling, NMR measurement and analyzing NMR spectra.

I am very grateful to Prof. Akifumi Takaori-Kondo, Dr. Keisuke Shindo, Kayoko Nagata for their valuable cooperation in my experiments with donating us Vif and its related proteins. I am also very grateful to Dr. Ayako Furukawa for kindly give me advice about properties of A3G and real-time NMR monitoring experiments. I sincerely appreciated Masayuki Saimura and Kyoto University Research Administration Office (KURA) for their technical support.

I am sincerely grateful to Prof. Takashi Morii and Prof. Masahiro Kinoshita as a second reader of this thesis and their valuable comments on this thesis.

I thank the past and present member of Katahira lab. As a member of the Katahira lab, I am grateful for all my colleagues. I have shared my time in the lab with Hideyasu Okamura, Keiko Kondo, Ryo Iwaoka, Yudai Yamaoki, Huyen Nguyen, Lin Meng-I, Mamiko Iida, Wan Li, Wan Hasnidah WAN OSMAN, Akihiro Kamiya, Yuko Sakurai, Hideaki Tagashira, Hiro-o Yoshida, Yoshihisa Kagiya, Hironao Sakai, Masato Tsumura, Ayaka Kiyoshi, Kenshi Takagi, Kazuma Nagata, Yuki Ueyama, Syunsuke Ozawa, Yushiro Kose Ryosuke Fushitani, who have helped to my research for suggestion, cooperation and making good lab environment.

Finally I would like to express my appreciation to my parents, Chiaki Kamba and Mayumi Kamba, my grandparents, Kou Kamba, Sumie Kamba, and my brother Shotaro Kamba for their encouragements and supports.

Keisuke Kamba

Laboratory of Bioenergy
Department of Fundamental Energy Science
Graduate School of Energy Science
Kyoto University

February, 2016