A EUROPEAN JOURNAL OF CHEMICAL BIOLOGY

CHEMBIOCHEM

SYNTHETIC BIOLOGY & BIO-NANOTECHNOLOGY

Accepted Article

Title: Differential inhibition of APOBEC3 DNA-mutator isozymes by fluoro- versus non-fluoro-substituted 2'-deoxyzebularine embedded in single-stranded DNA

Authors: Maksim V. Kvach, Fareeda M. Barzak, Stefan Harjes, Henry A. M. Schares, Harikrishnan M. Kurup, Katherine F Jones, Lorraine Sutton, John Donahue, Richard T. D'Aquila, Geoffrey B. Jameson, Daniel A. Harki, Kurt L. Krause, Elena Harjes, and Vyacheslav Filichev

To be cited as: *ChemBioChem* 10.1002/cbic.201900505

Link to VoR: http://dx.doi.org/10.1002/cbic.201900505

A Journal of ChemPubSoc Europe

WILEY-VCH www.chembiochem.org

10.1002/cbic.201900505

Differential inhibition of APOBEC3 DNA-mutator isozymes by fluoro- *versus* non-fluoro-substituted 2'-deoxyzebularine embedded in single-stranded DNA

Maksim V. Kvach,[a] Fareeda M. Barzak,[a] Stefan Harjes,[a] Henry A. M. Schares,[c] Harikrishnan M. Kurup,[a, b] Katherine F. Jones,[d] Lorraine Sutton,[f] John Donahue,[f] Richard T. D'Aquila,[g] Geoffrey B. Jameson,[a, b] Daniel A. Harki,[c, d] Kurt L. Krause,[b, e] Elena Harjes,[a, b]* Vyacheslav V. Filichev[a, b]*

[a] Dr. M. V. Kvach, F. M. Barzak, Dr. S. Harjes, H. M. Kurup, Prof. Dr. G. B. Jameson, Dr. E. Harjes, Prof. Dr. V. V. Filichev

School of Fundamental Sciences, Massey University, Private Bag 11 222, Palmerston North 4442, New Zealand

E-mail: e.harjes@massey.ac.nz, v.filichev@massey.ac.nz

[b] H. M. Kurup, Prof. Dr. G. B. Jameson, Prof. K. L. Krause, Dr. E. Harjes, Prof. Dr. V. V. Filichev

Maurice Wilkins Centre for Molecular Biodiscovery, Auckland 1142, New Zealand;

[c] H. A. M. Schares, Prof. Dr. D. A. Harki

Department of Medicinal Chemistry, University of Minnesota, Minneapolis, Minnesota 55455, USA;

[d] K. F. Jones, Prof. Dr. D. A. Harki

Department of Chemistry, University of Minnesota, Minneapolis, Minnesota 55455, USA;

[e] Prof. K. L. Krause

Department of Biochemistry, University of Otago, Dunedin 9016, New Zealand;

[f] L. Sutton, Dr. J. Donahue

Division of Infectious Diseases, Department of Medicine, Vanderbilt University School of Medicine, Nashville, Tennessee, USA;

[g] Prof. Dr. R. T. D'Aquila

Division of Infectious Diseases and Northwestern HIV Translational Research Center, Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA

Twitter accounts:

Massey University: @MasseyUni

Maurice Wilkins Centre: @MWC_CoRE

Worldwide Cancer Research: @WorldwideCancer

Daniel Harki: @DanHarki

Katherine Jones: @_ChemKat

Henry Schares: @HenrySchares

University of Minnesota, Department of Medicinal Chemistry: @UMN_MedChem

University of Minnesota, Department of Chemistry: @UMNChemistry

Keywords: APOBEC3, cytidine deaminase, zebularine, enzyme activity, inhibitor, transition state analogue, nucleoside, nucleotide, 5-fluoro-2'-deoxyzebularine.

Abstract

The APOBEC3 (APOBEC3A-H) enzyme family is part of the human innate immune system that restricts pathogens by scrambling pathogenic single-stranded DNA (ssDNA) via deamination of cytosines to uracils. However, APOBEC3-mediated mutagenesis of viral and cancer DNA promotes its evolution, enabling disease progression and development of drug resistance. Therefore, APOBEC3 inhibition offers a new strategy to complement existing anti-viral and anti-cancer therapies by making such therapies effective for longer periods of time, thereby preventing the emergence of drug resistance. Here, we have synthesised 2'-deoxynucleoside forms of several known inhibitors of cytidine deaminase (CDA), incorporated them into oligodeoxynucleotides (oligos) in place of 2'-deoxycytidine in the preferred substrates of APOBEC3A, APOBEC3B, and APOBEC3G, and evaluated their inhibitory potential against these enzymes. An oligo with 5-fluoro-2'-deoxyzebularine (5FdZ) exhibited a 3.5 times better inhibition constant in comparison with the comparable 2'-deoxyzebularine (dZ) containing oligo against APOBEC3B. A similar trend of inhibition was

10.1002/cbic.201900505

observed for wild-type APOBEC3A. In contrast, use of 5FdZ in an oligo designed for APOBEC3G inhibition resulted in a less potent inhibitor than the dZ-containing oligo for both APOBEC3G$_{CTD}$ and full-length wild-type APOBEC3G.

Introduction

APOBEC3 (A3) enzymes are important components of the innate immune system that protect against pathogens by deaminating cytosine residues to uracils from the single-stranded DNA (ssDNA) of the invading viral genome (Figure 1A).[1] These A3 enzymes therefore restrict the spread of pathogenic genetic information. Conversely, A3 enzymes, in particular A3B, can mutate genomic DNA, especially in cancer cells, thereby contributing to cancer genome evolution, acquired drug resistance and poor survival prognosis in multiple cancers (including breast, bladder, cervix, lung, head and neck).[2] A3B inhibition presents a promising new strategy to complement existing anti-cancer therapies,[3] as A3B is a non-essential protein.[4]

Figure 1

A3 enzymes and cytidine deaminase (CDA) share a similar structural topology despite very low sequence identity and, more importantly, a structurally homologous zinc-containing active site. This active site includes a crucial Glu residue that functions as a general acid/base in the hydrolysis of cytosine.[5] Consequently, A3s and CDA share a similar mechanism of cytosine deamination. However, CDA accepts only individual nucleosides as substrates,[6] whereas A3 enzymes have a minimal ssDNA substrate of 2-3 nucleotides flanking the cytosine.[7]

To date, no selective small-molecule inhibitors of A3A or A3B have been reported. We recently developed the first rationally designed competitive inhibitor of A3 by incorporating a known inhibitor of CDA, 2'-deoxyzebularine (**dZ**, **Figure 1B**), into ssDNA oligonucleotides.[8] We demonstrated that **dZ** does not inhibit A3 enzymes as the free nucleoside, but becomes a low μM inhibitor, if and only if, it is incorporated into ssDNA. This key observation supports a mechanism in which ssDNA delivers **dZ** to the active site for inhibition.

We propose that the inhibitory potential of ssDNAs can be further improved by incorporating potent inhibitors of CDA (also an enzyme of pharmaceutical interest)[9] into ssDNA . Here, we consider several cytidine derivatives known to inhibit CDA, which we have incorporated into ssDNA as possible inhibitors of A3 enzymes (Figure 1B). 3-Deazauridine (the ribose analogue of **3dadU**) has been reported as a weak

inhibitor of human liver CDA (K_i = 100 μM).[10] 5-Fluorozebularine has been shown to be a more potent inhibitor of mouse kidney CDA in comparison with Zebularine (K_i = 0.3 μM versus 2.3 μM, respectively).[11] However, RNA molecules are not preferred substrates of A3 enzymes.[12]

Here we report the first syntheses of the 2'-deoxy forms of 3-deazauridine and 5-fluorozebularine (**3dadU** and **5FdZ**, respectively). We report also the incorporation of these nucleosides into ssDNA and their evaluation as A3 inhibitors using our previously described NMR-based[7a, 8, 13] and fluorescence-based[14] enzymatic assays. 3-Deaza-2'-deoxyzebularine (**3dadZ**, Figure 1B) has a CH instead of the N3 atom in comparison with **dZ** and in this way can be used to evaluate the importance of protonation of the N3 atom in **dZ** in its inhibitory mechanism. Our results indicate subtleties in inhibition of the cytosine deamination reaction among A3 enzymes, and support our general strategy of using known inhibitors of CDA to guide the design of ssDNAs as inhibitors of A3 enzymes.

Results and discussion

Synthesis of modified nucleosides, their DMT-protected phosphoramidites and corresponding oligos

The synthesis of modified nucleosides started from heterocycles **1a-c** and Hoffer's chlorosugar[15] (Scheme 1). For the synthesis of **3dadZ** (compounds **2-5a**) and its incorporation into DNA, we followed previously described procedures[16] with some modifications as described in the Supporting Information section.

Scheme 1

The pure β-anomer of **3dadU** (compound **2b**) was obtained using silyl modification of the classical Hilbert-Johnson reaction[17] by reacting silylated 2,4-dihydroxypyridine with Hoffer's chlorosugar in boiling $CHCl_3$ (see Supporting Information). Double recrystallisation from EtOH provided **2b** in 50% yield. Cleavage of the toluoyl protecting groups was accomplished in MeOH/NH_4OH providing nucleoside **3b**,[18] which was then converted into **4b** by selective installation of the 4,4'-dimethoxytrityl (DMT) group on the 5'-end of the nucleoside followed by benzoyl protection of the 4-hydroxyl group of the nucleobase (81% yield over three steps from **2b**). Phosphitylation of **4b** was performed under standard conditions using *N*,*N*-

di*iso*propylamino-2-cyanoethoxychlorophosphine and Et_3N in CH_2Cl_2 in 84% yield after silica gel column chromatography.

The synthesis of **5FdZ** as a free nucleoside has been performed in the past by enzymatic conversion of dC in the presence of heterocycle **1c** using trans-*N*-deoxyribosylase from *Lactobacillus acidophilus*.[19] Later, **5FdZ** was synthesised from 5-fluoro-2'-deoxyuridine in six steps.[20] We found both protocols to be unsatisfactory in terms of potential scalability, complex procedures and overall yield. As with the synthesis of **dZ**[8] and **3dadU**, we first used a Lewis acid free variant of the silyl-Hilbert-Johnson reaction for the preparation of **5FdZ** from silylated heterocycle **1c** and Hoffer's chlorosugar; this procedure failed. Instead, it was necessary to use freshly distilled $SnCl_4$ and low temperatures (– 35 °C)[21] to obtain 3',5'-bis-*O*-toluoyl protected **5FdZ** (**2c**) in a good yield, although this product was contaminated with the α-anomer (β/α = 9:1). Use of a slow stepwise gradient of acetone (0→20 %) in CH_2Cl_2 allowed isolation of pure β-anomer **2c** in 45% yield. Deprotection was performed using saturated NH_3 in MeOH providing **5FdZ**-nucleoside **3c**, for which an analytical sample was obtained after preparative TLC on silica gel. NMR analysis revealed that this compound exists in two forms, i.e. as an "open" nucleoside and as a "cyclic" nucleoside formed after intramolecular addition of 5'-OH to the double bond of the nucleobase (Figure 2A). Similar transformations have been described for several pyrimidines.[22] Using 2D NMR techniques we assigned signals in ^1H and ^{13}C NMR spectra for individual forms as reported in the Supporting Information and shown on Figure 2A. Appearance of an NH signal at 8.69 ppm in ^1H NMR and a significant shift in H6 and C6 signals in ^1H and ^{13}C NMR spectra as a result of a change in hybridisation at C6 suggest formation of a "cyclic" nucleoside. The ^1H-^{13}C Heteronuclear Multiple Bond Correlation (HMBC) spectrum, which shows three-bond correlations, was particularly helpful during the assignment (Figure 2B). The H6-C5' cross-peak, seen on the right-hand upper corner of Figure 2B, confirms existence of a three-bond linkage between H6 of the nucleobase and C5' of the sugar in the "cyclic" nucleoside. At the same time, the H6 proton cross-talks with all carbons of the nucleobase (C2, C4, C5) and with C1' carbon of the sugar moiety, which is possible only for an O5'-C6 "cyclic" nucleoside. We observed that the ratio between "open" and "cyclic" forms changes in different solvents. In D_2O, the "open" form predominantly exists, whereas in d_6-DMSO, CD_3CN and d_8-THF both forms are present. This indicates that "open" and "cyclic" forms of **5FdZ**-nucleoside are in dynamic equilibrium,

which complicates purification, but allows transformation of an equilibrium mixture of nucleosides to a single 5'-DMT-modified product **4c**. That is why "open" and "cyclic" forms of nucleoside **3c** without any purification after removal of toluoyl groups from **2c** were treated with DMT-Cl in pyridine; compound **4c** was obtained in 60% yield over two steps from **2c**. Finally, phosphitylation of **4c** gave phosphoramidite **5c** in 89% yield.

Figure 2

We incorporated the modified nucleosides at the location of dC in the preferred A3 substrate motifs. A3B and A3A prefer the TCA motif (oligo1, **Table 1**) whereas A3G preferentially deaminates the CCCA motif (oligo2, where the underlined C is deaminated first). Synthesis of DNA oligos was performed using an automated DNA synthesiser with an increased coupling time for phosphoramidites **5a-c** from 1.5 min for standard phosphoramidites to 5 min. Cleavage from the solid support and deprotection of phosphates and nucleobases were accomplished in concentrated aqueous NH_4OH for oligos possessing **3dadU** and **3dadZ**. Unfortunately, the same procedure led to degradation of **5FdZ-**containing oligos as evident from the reverse-phase HPLC profile in Figure 2C (red line). Deprotection by saturated NH_3 in MeOH was also unsuccessful (blue line, Figure 2C). We found that on-column deprotection of **5FdZ-oligo** in organic solvents[23] led to the least amount of by-products (black profile, Figure 2C). Here, **5FdZ-oligo** on the CPG-support was treated with 10% Et_2NH in acetonitrile for 5 min, followed by incubation of the support in ethylenediamine/toluene mixture for 2 hr at room temperature, allowing subsequent release of the deprotected oligo in H_2O. All oligos were purified by reverse-phase HPLC. Their compositions were confirmed by ESI-MS (see Supporting Information).

Table 1

Evaluation of oligos as inhibitors of A3 enzymes using NMR-based activity assay

To directly assess the inhibition of A3 enzymes, we used a previously described NMR-based activity assay wherein the DNA substrate deamination is monitored by ^1H NMR spectroscopy in the presence of enzyme with and without inhibitors.[7a, 8, 13] The NMR-based inhibition assay is a direct assay using just A3 enzymes;

it does not require a secondary enzyme, such as Uracil-DNA Glycosylase (UDG), which is used in many indirect assays. By introducing different inhibitors of cytidine deamination into the A3 recognition motif preferred by the particular A3 enzyme, we expect that the trend in inhibition for all A3 enzymes parallels to a greater than lesser extent that observed for CDA inhibition, because the active site and therefore the deamination mechanism are conserved. We evaluated the inhibitory activity of our modified DNAs using active A3 enzymes with reliable expression and stability over time, which have also been previously characterised in the NMR-based activity assays in our laboratory. This allows reliable determination of the inhibitory potential of modified oligos by comparing K_m values of the substrates to K_i values of inhibitors determined under identical conditions (enzyme and substrate concentrations, buffer and ionic strength). The enzymes chosen, $A3B_{CTD}$-QM-ΔL3-AL1swap, hereafter simplified to $A3B_{CTD}$-AL1, and GST-fused $A3G_{CTD}$, were recombinantly expressed and purified from *Escherichia coli*. To compare the inhibitory effect of oligonucleotides between $A3G_{CTD}$ and full-length A3G (flA3G), we have used flA3G that was purified from human cells grown planktonically[24] (see description of these enzymes and their purification in Supporting Information).

Oligos containing **3dadZ** and **3dadU** instead of the target dC in the preferred TCA-recognition motif for $A3B_{CTD}$-AL1 had no effect on the initial speed of deamination by $A3B_{CTD}$-AL1 (**Figure 3**). These oligos fail to inhibit A3 enzymes under experimental conditions. These data are in line with previous results that **3dadU**, as an individual ribose-based single nucleoside, is a very weak inhibitor of human liver CDA (K_i = 100 μM).[10] Although higher concentrations of **3dadU**-oligo might result in inhibition of $A3B_{CTD}$-AL1, such concentrations would provide a weak basis for the development of modified **3dadU**-oligos as inhibitors compared to our current strategy. On the other hand, inhibition of $A3B_{CTD}$-AL1 by **5FdZ-oligo** was more profound in comparison with **dZ-oligo** under identical conditions. Previously, we confirmed that **dZ-oligo** is a competitive inhibitor of this enzyme.[8] By monitoring the reaction in the presence of inhibitor at various concentrations, we obtained the inhibition constant (K_i) for **5FdZ-oligo** (2.1 ± 0.8 μM, see Supporting Information), which was 3.5 times lower than the K_i of **dZ**-**oligo** (7.5 ± 1.7 μM). The overall inhibition effect was improved from 30-fold (**dZ**) to nearly 100-fold (**5FdZ**) if we compare the apparent inhibition constants (K_i) of **dZ** and **5FdZ**-containing oligos with K_m of the ssDNA substrate 5'-ATTT-C-ATTT (K_m = 200 μM). This

means that **5FdZ**-containing oligos can potentially be used in cells in the low μM range to inhibit A3A and A3B. Due to the presence of the electron withdrawing F, the heterocycle in **5FdZ** is more activated towards the nucleophilic addition of H_2O than **dZ** (Figure 4), which is evident from existence of **5FdZ** in equilibrium between "open" and "closed" forms (see Figure 2A). This likely explains why **5FdZ** is a better inhibitor than **dZ** of $A3B_{CTD}$-AL1 once placed in ssDNA. Formation of reversible covalent adducts with the enzyme is also possible. Similar adducts have been described between zebularine and DNA methyltransferases.[25]

Our observations parallel those reported earlier for CDA that 5-fluorozebularine is a better inhibitor than zebularine and 3-deazauridine (K_i = 0.3 μM, 2.3 μM[11] and 100 μM,[10] respectively). These results indicate that the structure of the nucleoside used instead of dC in the preferred ssDNA substrate determines the inhibitory potential of the oligos and that the trend of $A3B_{CTD}$-AL1 inhibition correlates with K_i values reported earlier for individual nucleosides against CDA. This result also opens the possibility to further improve inhibition by introducing other inhibitors of cytidine deamination into ssDNA sequences.

Figure 3

The fact that **3dadZ** does not inhibit $A3B_{CTD}$-AL1 highlights the importance of protonation of N3 in **dZ** by the conserved glutamic acid present in the active site of A3s (and CDA). This protonation makes C4 in **dZ** more electrophilic and more susceptible to the nucleophilic attack by OH^-/H_2O coordinated to the Zn^{2+}, which converts **dZ** into a tetrahedral transition-state analogue of cytidine deamination (Figure 4).[26] This mechanism is not realized in the case of **3dadZ**, because the C=C double bond of **3dadZ** is inactive towards water addition. Moreover, the nucleobase of **3dadU** is planar and does not mimic the tetrahedral geometry of C4 in the transition state of cytidine deamination.

Figure 4

Next, having two active A3G enzymes, the C-terminal domain ($A3G_{CTD}$) with the wild-type sequence and full-length A3G (flA3G), we decided to test if inhibition of $A3G_{CTD}$ is a good model to investigate inhibition of two-domain enzymes such as flA3G. Our studies were performed with two oligos, an A3G-preferred, **CC5FdZ-oligo** in which the dC that is first deaminated by A3G was changed to **5FdZ**, and the previously

reported inhibitor **CCdZ-oligo**.[8] Our data show that inhibition of A3G deaminase activity by targeting only the catalytically active C-terminal domain, $A3G_{CTD}$, accurately translates to the overall inhibition of flA3G (Figure 5A). This is consistent with the fact that the N-terminal domain of A3G completely lacks deaminase activity.[27]Accordingly, the specificity of ssDNA binding to the full-length A3B and A3G lies in the C-terminal domains, and the inactive N-terminal domains enhance ssDNA deamination efficiency at the C-terminal domain and regulate processivity of enzymes.[27a, 28]

Interestingly, the **CC5FdZ-oligo** (K_i = 71 ± 14 μM, Figure 5B) did not improve inhibition of $A3G_{CTD}$ in comparison with **CCdZ-oligo** (K_i = 53 ± 10 μM),[8] which is in contrast to the trend observed above for $A3B_{CTD}$-AL1. Nonetheless, that both **CCdZ-oligo** and **CC5FdZ-oligo** are inhibitors supports our strategy of targeting the catalytically active C-terminal domain of A3 enzymes using our DNA-based inhibitors as a means to inhibit full-length enzymes.

Figure 5

Evaluation of oligos as inhibitors of A3A using a fluorescence-based activity assay

To investigate how a **5FdZ-**containing oligo inhibits another wild-type A3 expressed in human cells, we purified A3A from HEK293T and used it to perform our previously published fluorescence-based activity assay.[14] The deamination of a fluorescently labeled substrate oligonucleotide was monitored in the presence of **dZ**- and **5FdZ-**containing oligo competitors (Figure 6). The assay was developed previously to evaluate small molecule inhibitors of A3A, A3B and A3G under identical settings. In the current work, we used an 18-mer oligo with 5'-…TATCCCA…-3' in the middle of the sequence as the enzyme substrate (see details in the Supporting Information). The CCC-motif is a preferred sequence for deamination by A3G[28-29] but this sequence is also readily deaminated by A3A and A3B;[30] therefore, the oligo is a pan-A3 substrate. The results clearly show that for A3A, the **5FdZ-**containing oligo is a more potent inhibitor with an IC_{50} of 0.16 ± 0.01 μM than the equivalent **dZ** containing oligonucleotide sequence with an IC_{50} of 0.39 ± 0.03 μM. These data are consistent with the $A3B_{CTD}$-AL1 data, as A3A and $A3B_{CTD}$ share all the residues surrounding

the target cytosine in x-ray structures.[31] Control assays with the dC- and dU-containing oligos can be found in the Supporting Information.

Figure 6

Plausible explanation of differences in inhibition of A3A/A3B$_{CTD}$ and A3G by dZ and 5FdZ-containing oligos

The fact that the trend of inhibition by **dZ** and **5FdZ**-containing oligos is varied for A3A/A3B$_{CTD}$ and A3G should not be completely unexpected as A3 family members differ strongly in their preferences to deaminate methylated cytosine in ssDNA.[32] The selectivity of different A3 family members for non-methylated against 5-methylated cytosine (**5MedC**) can be significantly changed by swapping loop 1 and loop 7 between the enzymes.[5b, 33] This property speaks for subtle control of the active site specificity for deamination of **5MedC**, as the swapped amino acids are not in direct contact with the target cytosine. Similar effects can be relevant to interaction of A3 enzymes with 5-fluoro-2'-deoxycytidine (**5FdC**) and **5FdZ**. Recently, we have also observed that 5-methyl-2'-deoxyzebularine is a worse inhibitor than **dZ** in the context of an oligo designed to inhibit A3G$_{CTD}$.[8]

Nevertheless, we have compared active sites of A3A, A3B and A3G to find possible differences in amino acids in proximity to the target cytosine. Such differences may explain the preferences of A3 enzymes towards various substrates and inhibitors.

Structural analysis of the A3B$_{CTD}$-AL1 complex with ssDNA[31] and sequential alignment of A3A, A3B$_{CTD}$, and A3G$_{CTD}$ revealed that not only the zinc-coordinating residues, but most of the residues in the active site close to the target cytosine are well conserved between these proteins. However, one residue in the substrate-binding pocket differs: an isoleucine, Ile 279/Ile 96, of A3B$_{CTD}$/A3A is threonine (Thr283) in the corresponding position of A3G$_{CTD}$ sequence. The sidechain of Ile279 is ~4.4 Å distant from the NH_2 group of the target cytosine in the inactive, substrate-binding E255A mutant of A3B$_{CTD}$ (pdb-id: 5DT5). On the other hand, the Thr283 hydroxyl group makes hydrogen bonds to a neighbouring Thr and to the main chain that forms part of the substrate/inhibitor binding pocket (pdb-id: 3V4J). The latter interaction may reduce the

conformational flexibility needed to accommodate a substituent larger than hydrogen in the 5-position of cytosine. We note that in the A3B-AL1 structure the bound cytidine is tightly supported by Tyr313 (Tyr315 in $A3G_{CTD}$), which in turn is buttressed by a conserved Trp (Trp285 for $A3G_{CTD}$) on the loop that for $A3G_{CTD}$ is locked in place by hydrogen bonding to Thr283. Interestingly, mouse CDA (pdb-id: 2FR5) has Ile87 in a similar position in the three-dimensional structure as Ile279 (Figure S3 in the Supporting Information). Thus, the Ile versus Thr substitution may play a role in the differences seen for **5FdZ** compared to **dZ** inhibition of $A3G_{CTD}$ and $A3B_{CTD}$-AL1. AID, mouse APOBEC1 and mouse APOBEC3 deaminate **5FdC** as well as **5MedC** with reduced speed compared to dC.[32b, 34] These results were explained by steric effects, as F and Me are larger than H. As discussed above, the active site around the target cytosine is very similar for $A3A/A3B_{CTD}$ and $A3G_{CTD}$, with the exception of Ile/Thr. The dynamics of the active site may allow better accommodation of **5FdZ** in $A3A/A3B_{CTD}$ than in $A3G_{CTD}$. In any event, the substrate and inhibitor binding and the deamination mechanism vary subtly between A3s and CDAs. Examples of highly homologous enzymes with significantly different transition states are well-established.[35]

Conclusion

The structure of modified nucleosides **dZ** and **5FdZ** embedded in the otherwise identical DNA sequence determines the inhibitory effects on human A3A, $A3B_{CTD}$, and $A3G_{CTD}$ as well as full-length A3G. On the other hand, the 2'-deoxyribose derivative of 3-deazauridine, a previously described weak inhibitor of CDA, cannot inhibit A3 upon its incorporation into ssDNA under the conditions tested. Our results indicate that some correlation exists between CDA and A3 inhibition, when CDA inhibitors replace the deamination-susceptible cytidine in the ssDNA sequence. Our results also highlight the importance of protonation of the N3 atom in **dZ** for its inhibitory mechanism. Noteworthy differences in inhibition profiles among different A3 enzymes observed here point to possibilities of obtaining highly specific A3inhibitors, supporting our approach towards developing oligonucleotide-based A3 inhibitors using chemically-modified nucleosides whose structure can stall enzymatic cytosine deamination.[36] Future work will continue to focus on the chemical optimisations of our ssDNA-based A3 inhibitors and their evaluation *in vitro* and *in vivo*.

Nucleotides flanking the target **dZ** and **5FdZ** can be further modified to improve inhibitory potential and enhance the lifetime of oligonucleotides in biological media.

Supporting Information

Experimental details for the synthesis of **3dadZ**, **3dadU** and **5FdZ** nucleosides, their DMT-protected phosphoramidites and modified oligos, for protein expression and purification, and for the NMR-based kinetic assays; examples of calculation of inhibition of A3-enzymes by **FdZ**-containing oligos; fluorescence-based A3A deamination assays; and sequence alignment of proteins used in this study.

Author information

Corresponding authors

*E-mail: e.harjes@massey.ac.nz, v.filichev@massey.ac.nz

ORCID

Geoffrey B. Jameson 0000-0003-4839-0784

Vyacheslav V. Filichev 0000-0002-7383-3025

Elena Harjes 0000-0002-3643-9432

Daniel A. Harki 0000-0001-5950-931X

Maksim V. Kvach 0000-0003-3592-0336

Kurt L. Krause 0000-0002-3525-1112

Funding: VVF, EH, GBJ, MVK and SH are grateful for the financial support provided by the Worldwide Cancer Research (grant 16-1197), Palmerston North Medical Research Foundation, Massey University Research Fund (MURF 2015, 7003 and RM20734) and School of Fundamental Sciences, Massey University. DAH acknowledges the United States National Institutes of Health (P01-CA234228, R01-GM110129 and R01-GM118000) for financial support. KFJ acknowledges the National Science Foundation Graduate

Research Fellowship Program for support. KLK thanks the Health Research Council, Otago University and the Thrash Foundation.

Notes:

The authors declare the following competing financial interest(s): DAH is a co-founder, shareholder, and a consultant of ApoGen Biotechnologies, Inc.

Acknowledgements:

NMR and mass-spectrometry facilities at Massey University and the assistance of Dr. P. J. B. Edwards and Mr. D. Lun are gratefully acknowledged. We are also thankful for the assistance in preparation of full-length A3G provided by Michele Krause and Chisu Song (Otago University). We thank Prof. Reuben S. Harris (University of Minnesota, USA) and members of his cancer research program for helpful discussions. We thank anonymous reviewers for the extensive and constructive feedback given to this manuscript.

References:

[1] a) R. S. Harris, J. P. Dudley, *Virology* **2015**, *479-480*, 131-145; b) M. H. Malim, P. D. Bieniasz, *Cold Spring Harb. Perspect. Med.* **2012**, *2*, a006940.

[2] a) S. Venkatesan, R. Rosenthal, N. Kanu, N. McGranahan, J. Bartek, S. A. Quezada, J. Hare, R. S. Harris, C. Swanton, *Ann. Oncol.* **2018**, *29*, 563-572; b) A. M. Sieuwerts, S. Willis, M. B. Burns, M. P. Look, M. E. Meijer-Van Gelder, A. Schlicker, M. R. Heideman, H. Jacobs, L. Wessels, B. Leyland-Jones, K. P. Gray, J. A. Foekens, R. S. Harris, J. W. M. Martens, *Horm. Cancer* **2014**, *5*, 405-413; c) E. K. Law, A. M. Sieuwerts, K. LaPara, B. Leonard, G. J. Starrett, A. M. Molan, N. A. Temiz, R. I. Vogel, M. E. Meijer-van Gelder, F. C. G. J. Sweep, P. N. Span, J. A. Foekens, J. W. M. Martens, D. Yee, R. S. Harris, *Sci. Adv.* **2016**, *2*, e1601737; d) P. J. Vlachostergios, B. M. Faltas, *Nat. Rev. Clin. Oncol.* **2018**, *15*, 495-509; e) L. Galluzzi, I. Vitale, *Trends Genet.* **2017**, *33*, 491-492; f) J. Zou, C. Wang, X. Ma, E. Wang, G. Peng, *Cell Biosci.* **2017**, *7*, 29.

[3] M. E. Olson, R. S. Harris, D. A. Harki, *Cell Chem. Biol.* **2018**, *25*, 36-49.

[4] J. M. Kidd, T. L. Newman, E. Tuzun, R. Kaul, E. E. Eichler, *PLoS Genet* **2007**, *3*, e63.

[5] a) T. P. Ko, J. J. Lin, C. Y. Hu, Y. H. Hsu, A. H. Wang, S. H. Liaw, *J. Biol. Chem.* **2003**, *278*, 19111-19117; b) Y. Fu, F. Ito, G. Zhang, B. Fernandez, H. Yang, X. S. Chen, *Biochem. J.* **2015**, *471*, 25-35.

[6] C. B. Yoo, S. Jeong, G. Egger, G. Liang, P. Phiasivongsa, C. Tang, S. Redkar, P. A. Jones, *Cancer Res.* **2007**, *67*, 6400-6408.

[7] a) S. Harjes, W. C. Solomon, M. Li, K.-M. Chen, E. Harjes, R. S. Harris, H. Matsuo, *J. Virol.* **2013**, *87*, 7008-7014; b) M. Liu, A. Mallinger, M. Tortorici, Y. Newbatt, M. Richards, A. Mirza, R. L. M. van Montfort, R. Burke, J. Blagg, T. Kaserer, *ACS Chem. Biol.* **2018**.

[8] M. V. Kvach, F. M. Barzak, S. Harjes, H. A. M. Schares, G. B. Jameson, A. M. Ayoub, R. Moorthy, H. Aihara, R. S. Harris, V. V. Filichev, D. A. Harki, E. Harjes, *Biochemistry* **2019**, *58*, 391-400.

[9] a) R. Z. Mahfouz, A. Jankowska, Q. Ebrahem, X. Gu, V. Visconte, A. Tabarroki, P. Terse, J. Covey, K. Chan, Y. Ling, K. J. Engelke, M. A. Sekeres, R. Tiu, J. Maciejewski, T. Radivoyevitch, Y. Sauntharajah, *Clin. Cancer Res.* **2013**, *19*, 938-948; b) M. R. Savona, O. Odenike, P. C. Amrein, D.

P. Steensma, W. A. Harb, A. E. DeZern, L. C. Michaelis, S. Faderl, J. N. Lowder, P. Taverna, A. Oganesian, R. Dua, R. Nawabi, M. Azab, G. Garcia-Manero, *Blood* **2015**, *126*, 1683-1683.
[10] G. G. Chabot, J. Bouchard, R. L. Momparler, *Biochem. Pharmacol.* **1983**, *32*, 1327-1328.
[11] J. J. Barchi, A. Haces, V. E. Marquez, J. J. McCormack, *Nucleosides Nucleotides* **1992**, *11*, 1781-1793.
[12] C. S. Nabel, J. W. Lee, L. C. Wang, R. M. Kohli, *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 14225-14230.
[13] A. Furukawa, T. Nagata, A. Matsugami, Y. Habu, R. Sugiyama, F. Hayashi, N. Kobayashi, S. Yokoyama, H. Takaku, M. Katahira, *EMBO J.* **2009**, *28*, 440-451.
[14] M. Li, S. M. D. Shandilya, M. A. Carpenter, A. Rathore, W. L. Brown, A. L. Perkins, D. A. Harki, J. Solberg, D. J. Hook, K. K. Pandey, M. A. Parniak, J. R. Johnson, N. J. Krogan, M. Somasundaran, A. Ali, C. A. Schiffer, R. S. Harris, *ACS Chem. Biol.* **2012**, *7*, 506-517.
[15] a) M. Hoffer, *Chem. Ber.* **1960**, *93*, 2777-2781; b) V. Rolland, M. Kotera, J. Lhomme, *Syn. Commun.* **1997**, *27*, 3505-3511; c) M. Kögler, R. Busson, S. De Jonghe, J. Rozenski, K. Van Belle, T. Louat, H. Munier-Lehmann, P. Herdewijn, *Chem. Biodiv.* **2012**, *9*, 536-556.
[16] a) F. Seela, U. Bindig, *Liebigs Ann. Chem.* **1989**, 895-901; b) R. H. Durland, T. S. Rao, R. Revankar, J. H. Tinsley, M. A. Myrick, D. M. Seth, J. Rayford, P. Singh, K. Jayaraman, *Nucleic Acids Res.* **1994**, *22*, 3233-3240; c) M. Urban, N. Joubert, M. Hocek, R. E. Alexander, R. D. Kuchta, *Biochemistry* **2009**, *48*, 10866-10881.
[17] M. P. Kotick, C. Szantay, T. J. Bardos, *J. Org. Chem.* **1969**, *34*, 3806-3813.
[18] B. L. Currie, R. K. Robins, M. J. Robins, *J. Heterocycl. Chem.* **1970**, *7*, 323-329.
[19] T. B. Oeyen, S. G. Laland, *Biochim. Biophys. Acta, Nucleic Acids Protein Synth.* **1969**, *182*, 567-569.
[20] D. Cech, A. Holy, *Collect. Czech. Chem. Commun.* **1977**, *42*, 2246-2260.
[21] K. Ebenryter-Olbinska, J. Karolak-Wojciechowska, E. Sochacka, *Carbohydr. Res.* **2014**, *392*, 7-15.
[22] a) P. K. Chang, *J. Org. Chem.* **1965**, *30*, 3913-3915; b) A. R. Katritzky, M. Kingsland, O. S. Tee, *Chem. Commun.* **1968**, 289-290; c) T. C. Thurber, L. B. Townsend, *J. Org. Chem.* **1976**, *41*, 1041-1051; d) J. Cadet, L.-S. Kan, S. Y. Wang, *J. Am. Chem. Soc.* **1978**, *100*, 6715-6720; e) P. S. Liu, V. E. Marquez, J. S. Driscoll, R. W. Fuller, J. J. McCormack, *J. Med. Chem.* **1981**, *24*, 662-666; f) S. M. N. Efange, E. M. Alessi, H. C. Shih, Y. C. Cheng, T. J. Bardos, *J. Med. Chem.* **1985**, *28*, 904-910; g) G. W. M. Visser, A. T. Bijma, J. A. R. Dijksman, G. C. M. Gorree, M. Van Walsum, J. D. M. Herscheid, *Nucl. Med. Biol.* **1989**, *16*, 351-357; h) V. E. Marquez, K. V. B. Rao, J. V. Silverton, J. A. Kelley, *J. Org. Chem.* **1984**, *49*, 912-919.
[23] Glen Research, **2013**, Deprotection Supplement, Deprotection - Volumes 1-5, http://www.glenresearch.com/Technical/Deprotection.pdf
[24] Y. Durocher, S. Perret, A. Kamen, *Nucleic Acids Res* **2002**, *30*, E9.
[25] L. Zhou, X. Cheng, B. A. Connolly, M. J. Dickman, P. J. Hurd, D. P. Hornby, *J. Mol. Biol.* **2002**, *321*, 591-599.
[26] a) S. Xiang, S. A. Short, R. Wolfenden, C. W. Carter, *Biochemistry* **1995**, *34*, 4516-4523; b) C. H. Borchers, V. E. Marquez, G. K. Schroeder, S. A. Short, M. J. Snider, J. P. Speir, R. Wolfenden, *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 15341-15345.
[27] a) F. Navarro, B. Bollman, H. Chen, R. Konig, Q. Yu, K. Chiles, N. R. Landau, *Virology* **2005**, *333*, 374-386; b) S. U. Siriwardena, T. A. Guruge, A. S. Bhagwat, *J. Mol. Biol.* **2015**, *427*, 3042-3055.
[28] M. A. Carpenter, E. Rajagurubandara, P. Wijesinghe, A. S. Bhagwat, *DNA repair* **2010**, *9*, 579-587.
[29] a) A. Rathore, M. A. Carpenter, O. Demir, T. Ikeda, M. Li, N. M. Shaban, E. K. Law, D. Anokhin, W. L. Brown, R. E. Amaro, R. S. Harris, *J. Mol. Biol.* **2013**, *425*, 4442-4454; b) R. M. Kohli, S. R. Abrams, K. S. Gajula, R. W. Maul, P. J. Gearhart, J. T. Stivers, *J. Biol. Chem.* **2009**, *284*, 22898-22904; c) R. M. Kohli, R. W. Maul, A. F. Guminski, R. L. McClure, K. S. Gajula, H. Saribasak, M. A. McMahon, R. F. Siliciano, P. J. Gearhart, J. T. Stivers, *J. Biol. Chem.* **2010**, *285*, 40956-40964; d) M. D. Stenglein, M. B. Burns, M. Li, J. Lengyel, R. S. Harris, *Nat. Struct. Mol. Biol.* **2010**, *17*, 222-229.
[30] I.-J. L. Byeon, C.-H. Byeon, T. Wu, M. Mitra, D. Singer, J. G. Levin, A. M. Gronenborn, *Biochemistry* **2016**, *55*, 2944-2959.
[31] K. Shi, M. A. Carpenter, S. Banerjee, N. M. Shaban, K. Kurahashi, D. J. Salamango, J. L. McCann, G. J. Starrett, J. V. Duffy, O. Demir, R. E. Amaro, D. A. Harki, R. S. Harris, H. Aihara, *Nat Struct Mol Biol* **2017**, *24*, 131-139.

[32] a) M. A. Carpenter, M. Li, A. Rathore, L. Lackey, E. K. Law, A. M. Land, B. Leonard, S. M. D. Shandilya, M.-F. Bohn, C. A. Schiffer, W. L. Brown, R. S. Harris, *J. Biol. Chem.* **2012**, *287*, 34801-34808; b) C. S. Nabel, H. J. Jia, Y. Ye, L. Shen, H. L. Goldschmidt, J. T. Stivers, Y. Zhang, R. M. Kohli, *Nat. Chem. Biol.* **2012**, *8*, 751-758; c) R. Suspene, M.-M. Aynaud, J.-P. Vartanian, S. Wain-Hobson, *PLoS One* **2013**, *8*, e63461; d) K. Kamba, T. Nagata, M. Katahira, *PLoS One* **2015**, *10*, e0124142; e) E. K. Schutsky, C. S. Nabel, A. K. F. Davis, J. E. DeNizio, R. M. Kohli, *Nucleic Acids Res* **2017**, *45*, 7655-7665; f) F. Ito, Y. Fu, S. A. Kao, H. Yang, X. S. Chen, *J. Mol. Biol.* **2017**, *429*, 1787-1799.

[33] P. Wijesinghe, A. S. Bhagwat, *Nucleic Acids Res* **2012**, *40*, 9206-9217.

[34] G. Rangam, K. M. Schmitz, A. J. Cobb, S. K. Petersen-Mahrt, *PLoS One* **2012**, *7*, e43279.

[35] a) P. C. Kline, V. L. Schramm, *Biochemistry* **1993**, *32*, 13212-13219; b) P. C. Kline, V. L. Schramm, *Biochemistry* **1995**, *34*, 1153-1162; c) A. Lewandowicz, V. L. Schramm, *Biochemistry* **2004**, *43*, 1458-1468.

[36] F. M. Barzak, S. Harjes, M. V. Kvach, H. M. Kurup, G. B. Jameson, V. V. Filichev, E. Harjes, *Org. Biomol. Chem.* **2019**, in press, doi: 10.1039/c9ob01781j.

10.1002/cbic.201900505

Figures:

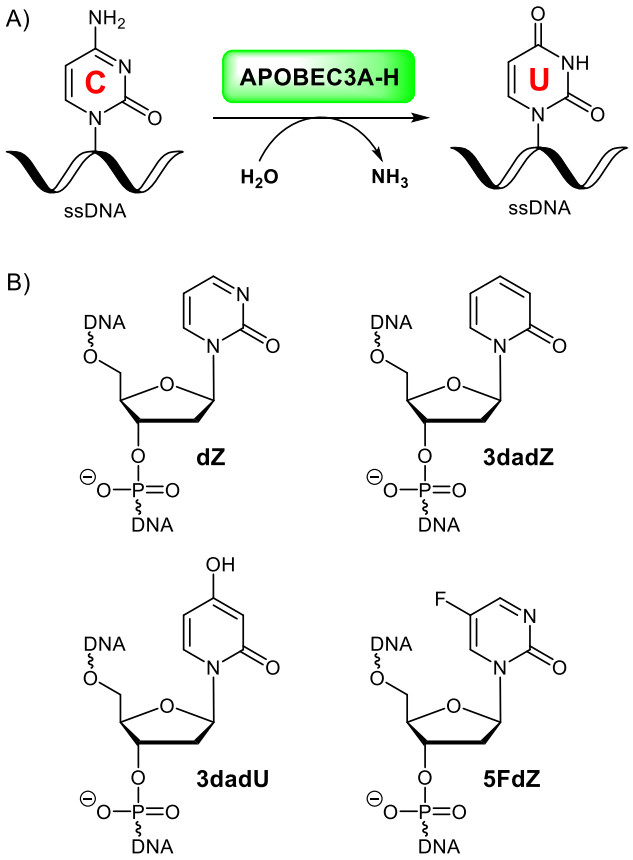


Figure 1. A) Deamination of dC in ssDNA by APOBEC3 enzymes; B) Modified nucleotides as potential inhibitors of cytidine deamination.

10.1002/cbic.201900505

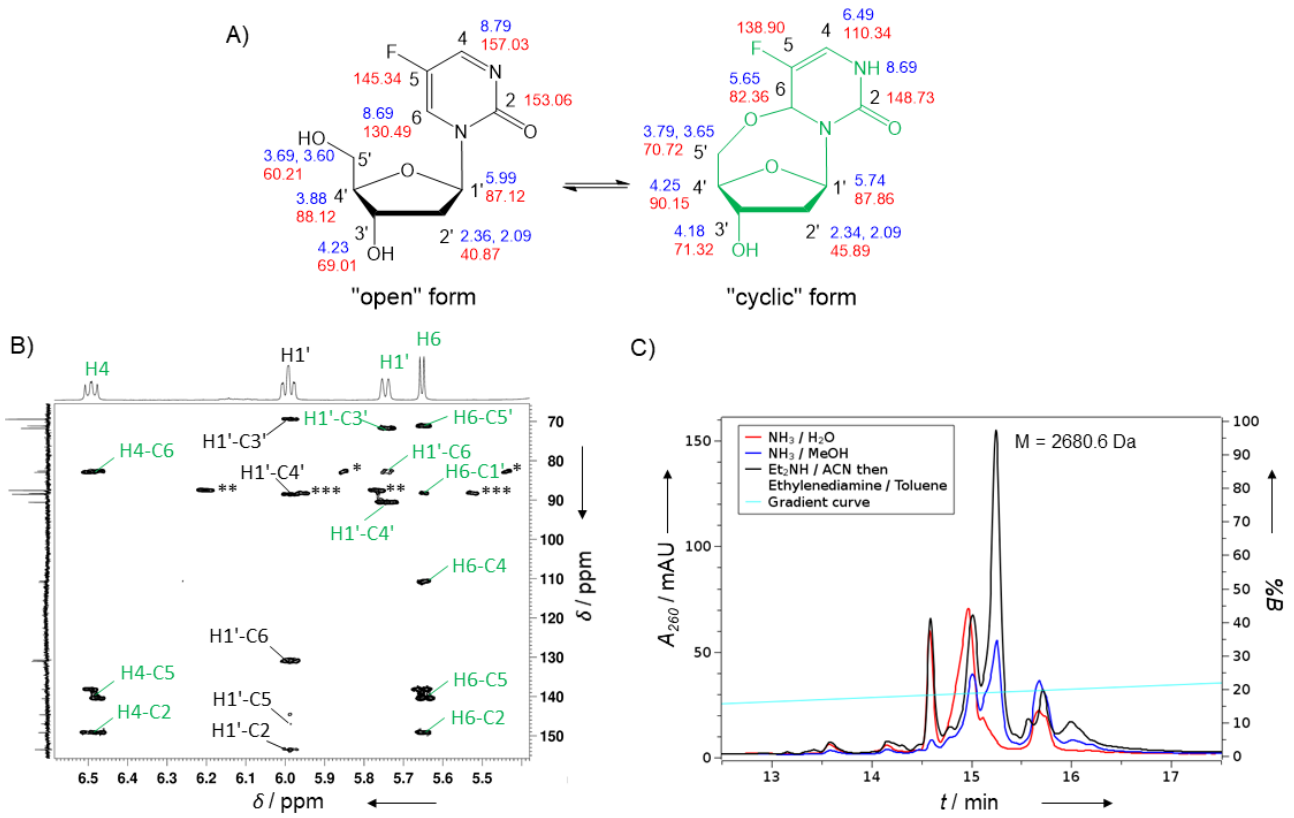


Figure 2. A) NMR assignment of "open" (structure in black) and "cyclised" (structure in green) forms of **5FdZ**. Chemical shifts (δ in ppm) are shown for ^{1}H in blue and for ^{13}C in red; B) ^{1}H-^{13}C HMBC spectrum recorded in DMSO-d_6 at rt showing 2-4 bond correlations and coexistence of "open" (black labels) and "cyclised" (green labels) forms of **5FdZ**. *, **, and *** are single bond correlations of H6-C6 (closed), H1'-C1' (open) and H1'-C1' (closed), respectively. C) RP-HPLC profiles of **5FdZ-oligo** cleaved from the support and deprotected under different conditions. The major peak isolated after treatment with Et_2NH/CH_3CN followed by ethylenediamine/toluene gave the expected mass in ESI-MS of 2680.6 Da (Calcd. for [M]: 2680.5 Da).

10.1002/cbic.201900505

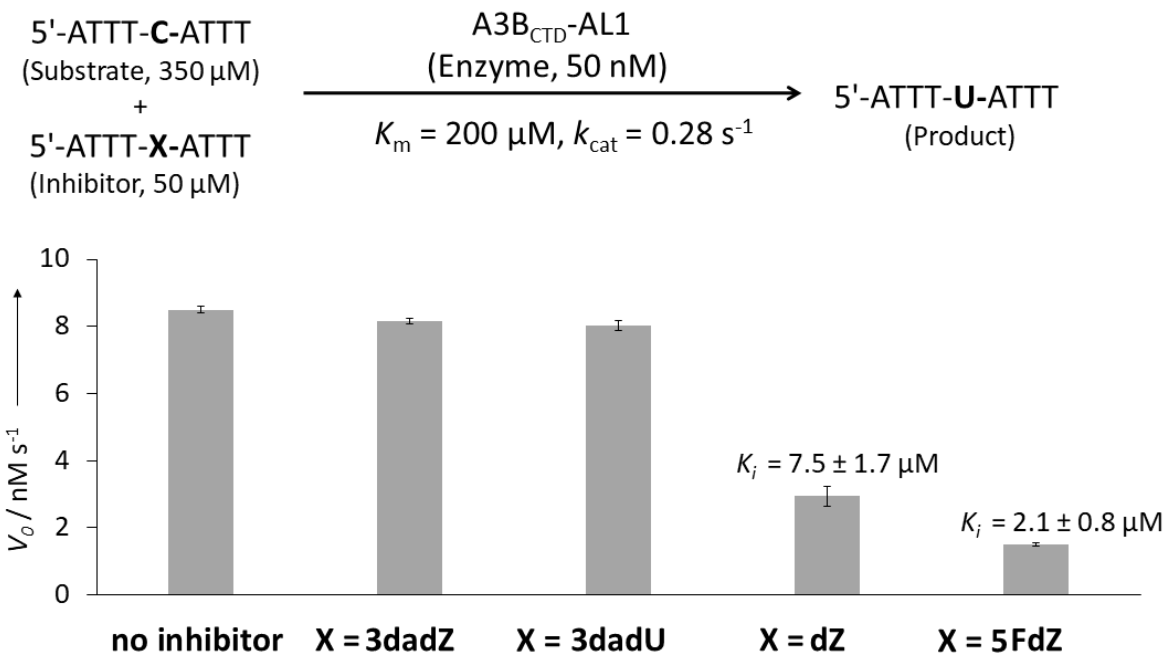


Figure 3. Inhibition of A3B$_{CTD}$-QM-ΔL3-AL1swap-catalysed deamination of the substrate (5'-ATTT-**C**-ATTT) by oligos containing modified nucleosides. The bold C is the target 2'-deoxycytidine deaminated by the enzyme. Determination of K_i for **5FdZ**-oligo can be found in the Supporting information. K_i (mean and standard deviations) for the **dZ**-oligo was reported previously[8] and is provided for comparison; all experiments were repeated multiple times in the same laboratory and the same time interval. Plotted are mean values. The error bars report SD.

10.1002/cbic.201900505

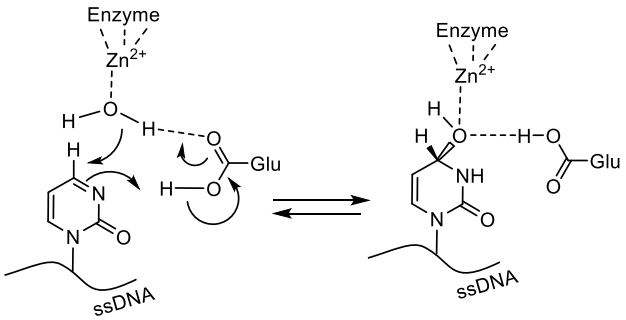


Figure 4. Proposed conversion of **dZ** into its hydrate and formation of a transition state analogue of cytosine deamination.

10.1002/cbic.201900505

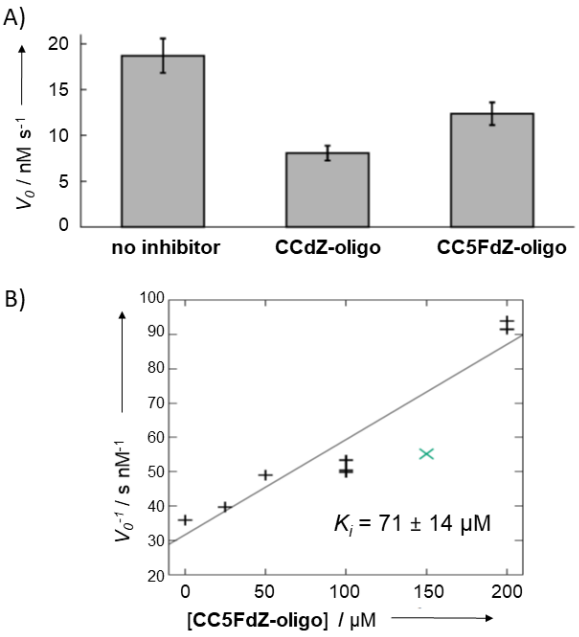


Figure 5. A) The speed of full-length A3G-catalysed deamination of the substrate (5'- ATTCCCAATT, 500 µM) in the absence of inhibitor (no inhibitor) and in the presence of **CCdZ**- and **CCFdZ**-oligos (50 µM). Conditions: 100 mM NaCl, 50 mM sodium phosphate buffer, pH 6.0, 10 % D_2O, 1 mM citrate supplemented with 50 µM TMSPS as standard. Error bars represent SD. B) The Dixon plot of inverse speed of deamination against inhibitor concentration for $A3G_{CTD}$ catalysed deamination of 5'-ATTCCCAATT (320 µM, underlined C is deaminated) in the presence of **CC5FdZ-oligo**. The green point was ignored by fitting as an outlier according to the Q-test (with 95% confidence).

10.1002/cbic.201900505

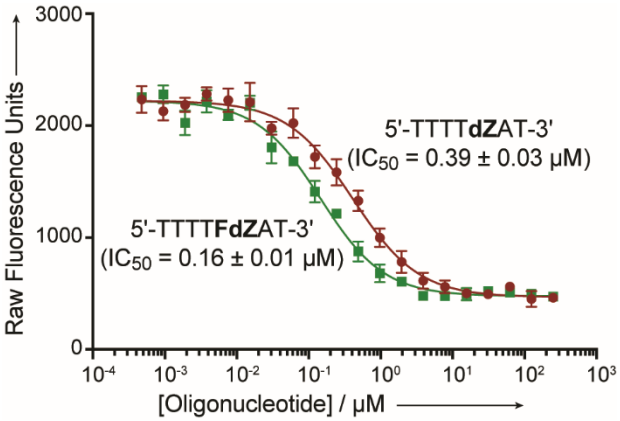
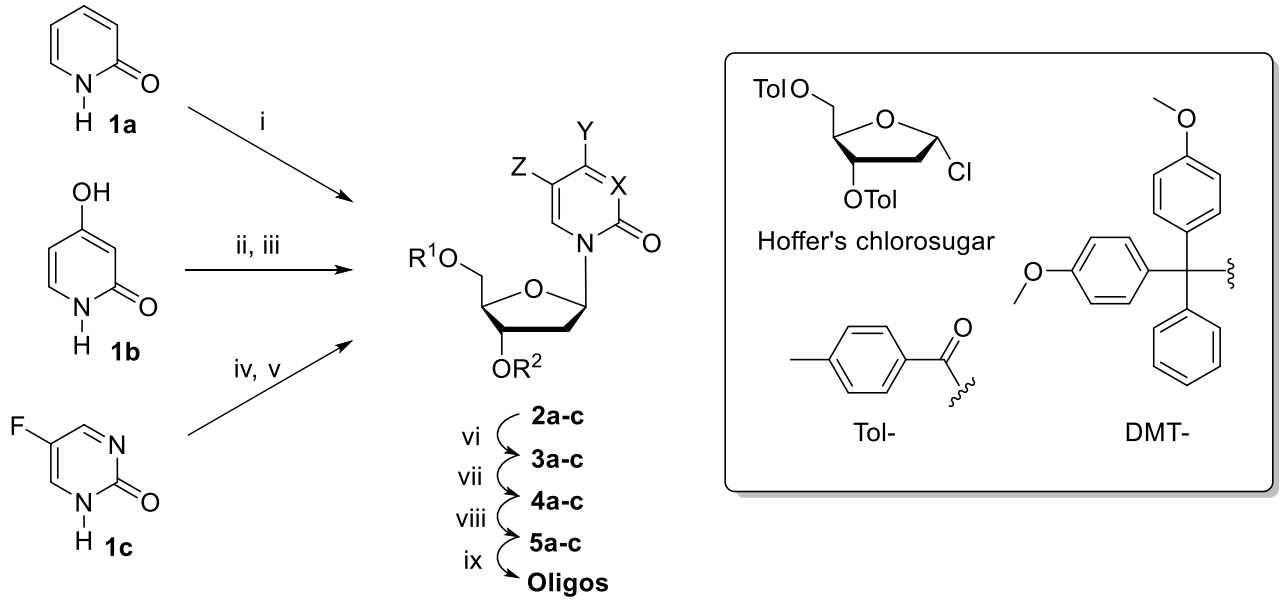


Figure 6. A) Inhibition of human APOBEC3A-catalysed deamination of a fluorescently labelled oligonucleotide reporter by **dZ**- and **5FdZ**-containing oligonucleotide competitors. Representative graphical data are shown. Individual replicates and the sequence of the fluorescent oligonucleotide reporter are provided in the Supporting Information. IC_{50} values are the mean + SEM for four biological replicates.

10.1002/cbic.201900505



#	X	Y	Z	R^1	R^2	Yield
2a	CH	H	H	Tol	Tol	11%
2b	CH	OH	H	Tol	Tol	50%
2c	N	H	F	Tol	Tol	45%
3a	CH	H	H	H	H	ND
3b	CH	OH	H	H	H	quant.
3c	N	H	F	H	H	ND
4a	CH	H	H	DMT	H	58% from **2a**
4b	CH	OBz	H	DMT	H	81% from **2b**
4c	N	H	F	DMT	H	60 % from **2c**
5a	CH	H	H	DMT	$P(NiPr)_2OCH_2CH_2CN$	93%
5b	CH	OBz	H	DMT	$P(NiPr)_2OCH_2CH_2CN$	84%
5c	N	H	F	DMT	$P(NiPr)_2OCH_2CH_2CN$	89%

Scheme 1. *i)* Hoffer's chlorosugar, KOH, CH_3CN, 15 min; *ii*) Hexamethyldisilazane (HMDS), $(NH_4)_2SO_4$ (cat), reflux 1h; *iii*) Hoffer's chlorosugar, $CHCl_3$, distill., 15 min; *iv*) HMDS, $(NH_4)_2SO_4$ (cat), reflux 1h; *v*) Hoffer's chlorosugar, $SnCl_4$, 1,2-dichloroethane, -35 °C; *vi*) for **3a** and **3b**: 28% aq. ammonia, MeOH, 48 h; for **3c**: satd. ammonia in MeOH, 48 h; *vii*) for **4a** and **4c**: 4,4'-dimethoxytritylchloride (DMTCl), pyridine, 0°C→r.t., overnight; for **4b**: 4,4'-dimethoxytritylchloride, pyridine, 0°C→r.t., overnight then Bz_2O, pyridine, 0°C→r.t., overnight *viii)*, *N*,*N*-di*iso*propylamino-2-cyanoethoxychlorophosphine, Et_3N, CH_2Cl_2; *ix*) DNA synthesis and purification.

10.1002/cbic.201900505

Table 1. Oligonucleotides used in the current study.

Name	Sequence 5'→3'
Oligos used in NMR-based activity assay:	
oligo1	ATTT-**C**-ATTT
oligo2	ATTCC-**C**-AATT
dZ-oligo[a]	ATTT-**dZ**-ATTT
3dadZ-oligo	ATTT-**3dadZ**-ATTT
3dadU-oligo	ATTT-**3dadU**-ATTT
5FdZ-oligo	ATTT-**5FdZ**-ATTT
CC5FdZ-oligo	ATTCC-**5FdZ-**AATT
Oligos used in fluorescence based activity assay:	
T4-dZ-oligo[a]	TTTT-**dZ**-AT
T4-5FdZ-oligo	TTTT-**5FdZ**-AT

a) Prepared as in reference [8].

10.1002/cbic.201900505

Graphical abstract:

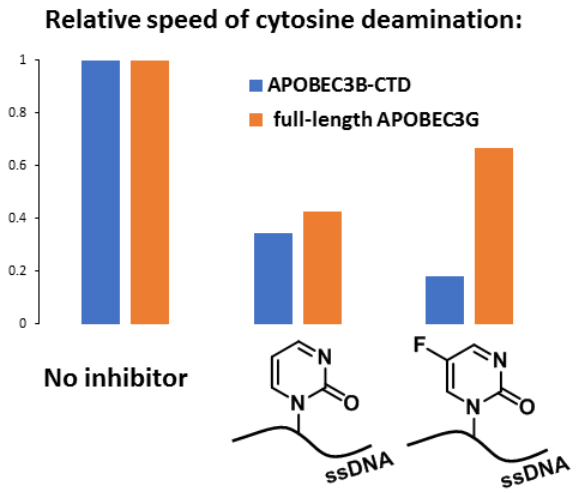


Table of content text:

In comparison with 2'-deoxyzebularine, incorporation of 5-fluoro-2'-deoxyzebularine into single-stranded DNA instead of the target dC yielded improved inhibition of APOBEC3A and APOBEC3B but not of full-length APOBEC3G.