# Cross-linked DNA with carbazole pseudo-nucleoside enables potent inhibition of the cancer DNA- and RNA-mutating enzyme APOBEC3A

Harikrishnan M. Kurup[1,2], Stefan Harjes[1], Yongdong Su,[1] Thomas B. J. Hall,[1] Geoffrey B. Jameson[1,2], Elena Harjes[1,2]*, Vyacheslav V. Filichev[1,2]*

[1]School of Natural Sciences, Massey University, Private Bag 11 222, Palmerston North 4442, New Zealand; [2]Maurice Wilkins Centre for Molecular Biodiscovery, Auckland 1142, New Zealand.

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

## Abstract

The human antiviral enzyme APOBEC3A has mutagenic activity in many human cancers, where it is used to fuel tumour evolution. Its single-stranded DNA and RNA C-to-U editing activities contribute to undesirable mutagenic outcomes in cancers. Inhibitors of APOBEC3A may therefore block mutagenesis and prevent tumour evolution and thus reduce detrimental outcomes such as drug resistance and metastasis. Here we demonstrate that use of a carbazole pseudo-nucleoside as a part of a covalent cross-link between distant nucleotides in DNA leads to a faster deaminating substrate and upon changing of dC to 2'-deoxy-5-fluorozebularine to a more potent inhibitor of APOBEC3A ($K_i$ = 29 ± 5 nM) in comparison with a DNA hairpin and previously described cross-linked DNAs. Truncation of the DNA sequence led to the first 4-mer cross-linked DNA inhibitor of APOBEC3A with nanomolar potency that is also stable against digestion by a phosphodiesterase. This provides a platform for a rational design of competitive inhibitors of APOBEC3 with properties of small-molecule drugs.

## Introduction

The enzymes of APOBEC3 family (A3A-A3H, except E) deaminate cytosine to uracil in single stranded DNA (ssDNA, **Figure 1A**) and form part of innate immune system protecting a host from viruses and retrotransposons.[1] However, two human enzymes, A3A and A3B, promote cancer genomic mutagenesis leading to development of drug resistance, metastasis, and poor disease outcomes.[2] Recent studies show that A3A, as the most active deaminase, contributes not only to A3 mutational signatures in genome but also in transcriptome of cancers[3] and drives carcinogenesis in mice.[4] As A3A and A3B are not essential for primary metabolism, it has been proposed that A3 inhibitors can be used to prevent or delay A3-mediated evolution of primary tumours into metastatic or drug-resistant tumours.[5]

We have previously developed the first selective A3 inhibitors[6] based on incorporation of cytosine-like 2'-deoxyzebularine (**dZ**, **Figure 1B**)[7] into short, linear ssDNA, achieving low micromolar inhibition constants ($K_i$). Structural studies by us and others[8] revealed that flexible ssDNA adopts a more rigid *U*-shaped loop upon binding to A3, projecting cytosine into the active-site pocket (**Figure 1A**). In addition, A3A and, to a lesser extent, the catalytically active C-terminal domain of A3B, $A3B_{CTD}$, have been reported recently to deaminate cytosine in short 3-4 nucleotide loops of DNA hairpins more rapidly than corresponding linear single-stranded DNA.[9] Importantly, DNA hairpins are both physiologically and pathologically relevant. Such structures are ubiquitous in nature, especially at inverted repeats, and cause stalling of replication, genomic instability, and predisposition to mutagenesis.[10] By using DNA hairpins in which the target dC at the apex of the short loops was changed to **dZ** derivatives, the first nanomolar inhibitors of A3A have been developed.[11] Whereas DNA hairpins carrying zebularine inhibited A3A in nM range, for A3B, A3F[11d] or A3G, at best, only weak high μM inhibition was observed.[11b]

We sought then to leverage the advantage conferred by hairpin structures, as substrates and inhibitors, by covalently linking the ends of the loop. By creating a cross-link 1 between modified nucleotides in the −2 and +1 positions (**Figure 1C**), strong inhibition was observed.[11b] Noting that A3A as well as A3B prefer a purine (dG or dA) in +1 position of the DNA sequence,[3b, 3c, 5a, 12] we hypothesised that the use of a pseudo-nucleoside with a large hydrophobic residue, especially one that is potentially more electron-rich, instead of a canonical nucleobase, will provide a better binder to A3 enzymes. We therefore positioned a carbazole nucleoside at position +1, as an alkyne derivative $\mathbf{dH^{E1}}$, and at position −2 cytosine dC, as an azide derivative, 4-azidomethyl-5-methyl-dC ($\mathbf{dC^{N3}}$). A new triazole cross-link 2,

which relative to cross-link 1, reverses positions of azide and alkyne, was created using Cu(I)-catalysed azide-alkyne cycloaddition (CuAAC) between carbazole nucleoside **dH$^{E1}$** at position +1 and **dC$^{N3}$** at position –2 (**Figure 1D**).

Two APOBEC3 constructs were used: wild-type A3A (wtA3A) and the catalytically active C-terminal domain of A3B (A3B$_{CTD}$) where four mutations were made to enhance solubility and stability (QM), along with truncation of a long loop common to both A3A and A3B (ΔL3), and replacement of loop 1 of A3B by loop 1 of A3A (AL1swap) – together summarised as A3B-QM-ΔL3-AL1swap, which is denoted here as A3A-mimic due to sequence and functional features mapping more closely to A3A than to A3B.[6c, 8d, 13] Most experiments were done with A3A-mimic as it is soluble and can be obtained in larger quantity. The results were confirmed, using the wild-type A3A.

The new cross-link 2 substrate is deaminated much faster by A3A-mimic in comparison with DNA hairpins and cross-link 1 described recently.[11a] The best substrate was converted to the most potent inhibitor to date of wild-type A3A (wtA3A) with $K_i$ of 29 ± 5 nM. Moreover, with an eye to enhancing potential druggability of our inhibitors, we removed nucleotides flanking the cross-link, and obtained the minimal inhibitor composed of just four nucleotides. This inhibitor maintained nanomolar potency against A3A.

## Results and Discussion

*Design and synthesis of cross-link 2 substrate and inhibitors*

We decided to use a carbazole nucleoside **dH$^{E1}$** (**Figure 1D**) because several carbazole derivatives have been successfully incorporated into DNA in the past,[14] some of which were used for modulation of thermal stability of DNA duplexes relying on noncovalent π-stacking interactions and/or intercalation with nucleic acids or with other polyaromatic compounds.[14] In addition, carbazole is a well-known scaffold in medicinal chemistry that is found also in natural products.[15]

As a partner of **dH$^{E1}$** for cross-linking, we synthesised 5'-*O*-DMT-protected **dC$^{N3}$** carrying azidoethyl in the 4$^{th}$ position of 5-methyl-dC as a 3'-*O*-phosphoramidite[11a, 16] but also as the more stable 3'-*O*-H-phosphonate, which is also compatible with standard phosphoramidite chemistry in an automatic DNA synthesiser.[17] A substrate having dC in position 0 of the 9-mer sequence and inhibitors having **dZ** and **FdZ** instead of dC$^0$ were synthesised as linear oligonucleotides carrying **dH$^{E1}$** and **dC$^{N3}$** in positions +1 and –2, respectively. After cleavage

from the solid support, deprotection and isolation, pure oligos were cross-linked using CuACC which was monitored by reverse-phase HPLC. Cross-linked products had shorter retention times, which is consistent with previous observations;[11a, 18] and identity was confirmed with ESI-MS. **Figure 1E** shows the sequence abbreviations used. Full details of synthesis and characterisation of products by NMR spectroscopy and mass spectrometry, as well as representative structures of each cross-linked product are provided in Supporting Information.

Use of **FdZ** in the sequence also required several adjustments in the synthetic protocol. To prevent hydration of the terminal alkyne[19] in **dH$^{E1}$** during oligo deprotection using 10% $Et_2NH$ in acetonitrile followed by ethylenediamine/toluene,[6b] we protected the alkyne in **dH$^{E1}$** with tri*iso*propylsilyl (TIPS), which is stable to DNA synthesis[19a, 19c] and during oligo deprotection. To avoid $HO^-$-mediated degradation of **FdZ** during TIPS deprotection, we included *o*-nitrophenol as a weak acid ($pK_a$ = 7.22)[20] to a 1 M tetra-*n*-butylammonium fluoride (TBAF) solution in THF, which led to successful cleavage of TIPS and isolation of the desired **FdZ**-containing oligos.

Potentially to further enhance substrate activity and inhibitor potency, we elongated the cross-links to bring additional conformational freedom for the cross-link. For the synthesis of dZ-oligos containing **dH$^{E2}$** and **dH$^{E3}$** with elongated cross-links, we introduced 3-iodo-carbazole nucleoside into DNA that was subjected to the on-column Sonogashira reaction[21] using 1,3-diethynylbenzene or 1,7-octadiyne, respectively (see Supporting Information for further details).

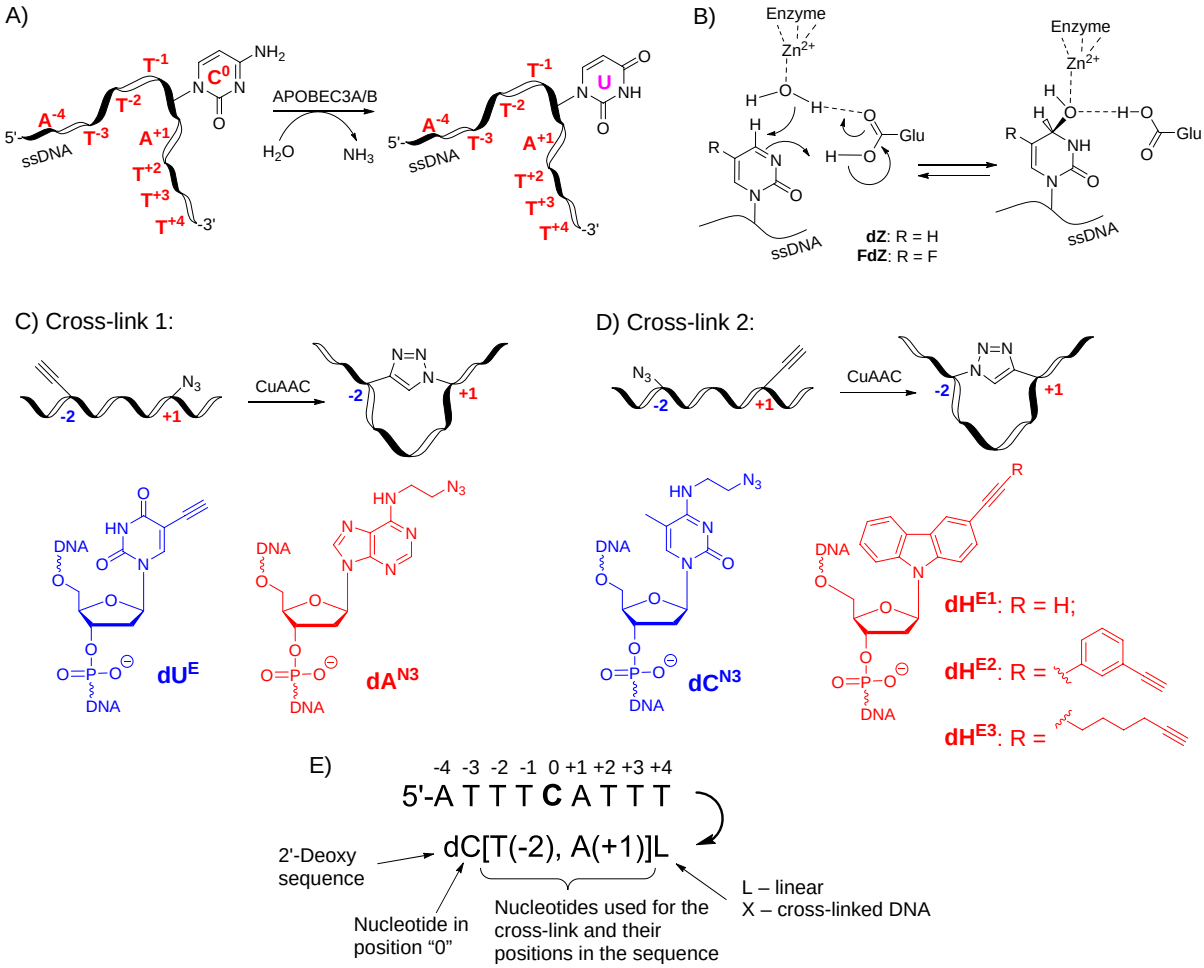



**Figure 1.** (A) Deamination of dC to dU in *U*-shaped linear ssDNA by A3A and A3B enzymes. $C^0$ represents the target cytosine; nucleobases (thymine T and adenine A) at 5' relative to $C^0$ are shown as $T^{-1}$, $T^{-2}$, and so forth and those at 3' relative to $C^0$ as $A^{+1}$, $T^{+2}$, and so forth. (B) Structures of A3 inhibitors, 2'-deoxyzebularine (**dZ**) and 5-fluoro-2'-deoxyzebularine (**FdZ**), incorporated into ssDNA (left) and their hydrated forms (right) observed in the active site of A3A (FdZ)[11b, 11c] and A3G (dZ)[22] forming 4-(*R*)-hydroxy-3,4-dihydro-2'-deoxyzebularine, and its 5-fluoro analogue, and thus mimicking an intermediate of C-to-U conversion. (C) Cu(I)-catalysed azide-alkyne cycloaddition (CuAAC) forms a 1,4-disubstituted 1,2,3-triazole cross-link 1 between nucleotides $\mathbf{dU^E}$ and $\mathbf{dA^{N3}}$ placed in positions –2 and +1 of ssDNA, respectively. (D) Cross-link 2 formed by modified carbazole pseudo-nucleotides ($\mathbf{dH^{En}}$, n = 1, 2, 3) in position +1 and 4-azidoethyl-2'-deoxycytidine ($\mathbf{dC^{N3}}$) in position –2. (E) DNA sequence used for modifications and abbreviation of linear and cross-linked oligos. For example, abbreviation $dC[C^{N3}(-2),H^{E1}(+1)]L$ indicates a linear DNA sequence containing cytosine at position 0, in which $\mathbf{dC^{N3}}$ is present in position –2 instead of T and $\mathbf{dH^{E1}}$ is in position +1 instead of dA, relative to the parent linear dATTTCATTT sequence. Letter X at the end of the name, i.e. $dC[C^{N3}(-2),H^{E1}(+1)]X$ means that the oligonucleotide is cross-linked by CuAAC. Truncated oligonucleotides, 4-mer $FdZ[C^{N3}(-2),H^{E1}(+1)]X$ and 3-mer $FdZ[C^{N3}(-1),H^{E1}(+1)]X$, are cross-linked oligonucleotides obtained from the linear $dC^{N3}$ T FdZ $dH^{E1}$ and $dC^{N3}$ FdZ $dH^{E1}$, respectively. Examples of chemical representation of structures are provided in Supporting Information.

*Enzymatic activity of cross-link 2 substrates*

The enzyme used in our experiments is the well-characterised and very active A3A-mimic (A3B$_{CTD}$-QM-ΔL3-AL1swap) , which is derived from the C-terminal domain of A3B (A3B$_{CTD}$, see protein sequence in Supporting Information.).[6a] The same enzyme was used in the crystal structure of its complex with linear ssDNA (except that for the X-ray structure, the active-site Glu255 was mutated to Ala to prevent deamination of the substrate, dC-containing ssDNA).[8d] However, loop 1 has substantial differences in length and amino-acid residues between A3A and A3B$_{CTD}$, and the A3B$_{CTD}$-QM-ΔL3-AL1swap has the loop 1 of A3A. Specifically, the presence of His29 from A3A instead of the possibly corresponding Arg in A3B$_{CTD}$, endows this A3B construct with the substrate specificity of A3A,[6c, 11c, 13] and thus A3B$_{CTD}$-QM-ΔL3-AL1swap is denoted here as A3A-mimic.

Using the 9-mer oligonucleotide 5'-AT$_3$CAT$_3$ as the standard linear substrate of A3A-mimic, we compared the deamination of cross-linked oligos using our previously described real-time NMR assay, which reports directly on consumption of substrate and formation of product.[6] A marked 2.3 times increase in deamination rate was detected for cross-linked oligonucleotide dC[C$^{N3}$(–2),H$^{E1}$(+1)]X connecting the –2 and +1 positions in the DNA sequence (**Figure 2**). To our surprise, this carbazole-containing oligonucleotide was deaminated even faster (1.4 times) than a native dC-hairpin-1, (GC)$_2$TTC(GC)$_2$ (**Figure 2**), and our original cross-link 1[11a] both of which were deaminated 1.5 times faster than a linear substrate under similar conditions.

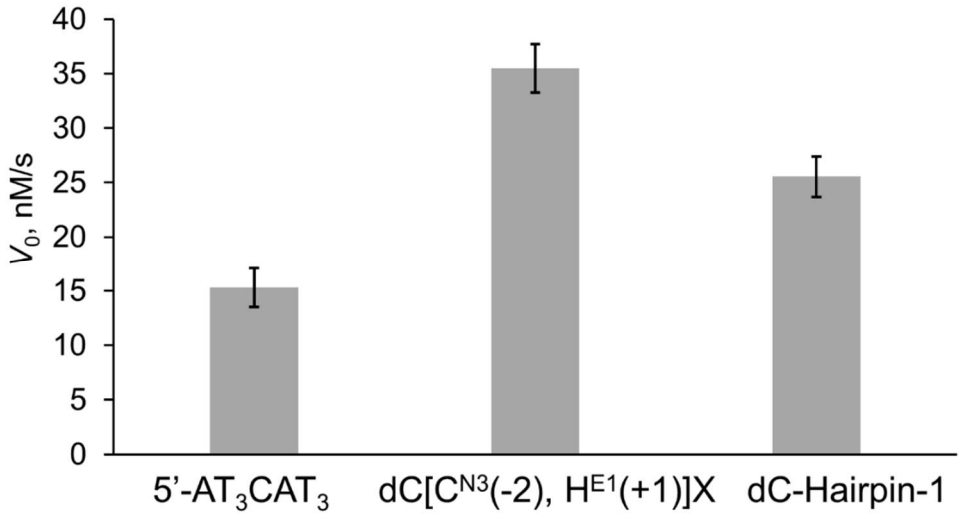



**Figure 2.** Cross-linked oligonucleotide and a dC-hairpin-1, (GC)$_2$TTC(GC)$_2$, in comparison with the linear substrate of A3A-mimic. Substrate concentrations are 400 µM, enzyme concentration is 100 nM in 50 mM Na-phosphate (pH 6.0), 200 mM NaCl, 2 mM β-mercaptoethanol, 200 µM 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) and 10 % D$_2$O at 25 °C. Error bars are estimated standard deviations from triplicate measurements.

*Enzyme inhibition by cross-link 2 inhibitors*

The increased rate of deamination of our carbazole-containing cross-linked oligonucleotide prompted us to investigate the inhibitory potential of these constructs by changing dC to dZ or FdZ at position 0. To evaluate inhibitory potential of cross-linked oligonucleotides against linear oligonucleotides characterised earlier[6a, 6b] (dZ-linear denotes 5'-$AT_3$**dZ**$AT_3$, FdZ-linear denotes 5'-$AT_3$**FdZ**$AT_3$), we monitored residual activity of A3A-mimic and wild-type A3A (denoted wt-A3A) on the unmodified oligo (5'-$T_4$CAT) in the presence of known concentrations of inhibitors using the NMR assay (**Figure 3**).

The fastest-deaminated substrate dC[$C^{N3}$(-2),$H^{E1}$(+1)]X based on the carbazole-containing cross-link exhibited the most powerful A3A inhibition studied here having dZ or FdZ nucleotides instead of dC. Notably, the FdZ-containing oligonucleotide FdZ[$C^{N3}$(–2),$H^{E1}$(+1)]X showed higher inhibitory potential than dZ-containing oligo dZ[$C^{N3}$(–2),$H^{E1}$(+1)]X, which is in line with our previous observations[6b] that FdZ-linear oligonucleotides are more powerful inhibitors of A3A-mimic and wt-A3A than dZ-linear oligonucleotides (**Figure 3**). Significantly, the previously characterised FdZ-hairpin DNAs[11c] were not as powerful inhibitors as cross-linked FdZ[$C^{N3}$(–2),$H^{E1}$(+1)]X, which also correlates with the lower deamination rate for dC-hairpin-1 in comparison with dC[$C^{N3}$(-2),$H^{E1}$(+1)]X (**Figure 2**). This reinforces our observation[11c] that these dZ- and FdZ-containing inhibitors are functioning as quasi-transition-state inhibitors.

Encouraged by these results and with an eye to increasing druggability potential for A3 inhibitors, we synthesised short 3-mer and 4-mer cross-linked DNAs carrying carbazole and dZ or FdZ, but with the nucleotide tails truncated. Although the 4-mer FdZ[$C^{N3}$(–2),$H^{E1}$(+1)]X was a less potent inhibitor than the cross-linked 9-mer oligo, its inhibitory potential was very close to that for the FdZ-hairpins. A further decline in inhibition of C-to-U deamination was observed upon shortening of the DNA length to the 3-mer (–1,+1 link), for which, surprisingly given the tightness of the loop, inhibition was still slightly greater than linear ssDNA inhibitors. The 4-mer non-cross-linked inhibitor had inhibitory potential very similar to that for the FdZ linear inhibitor; the non-cross-linked 3-mer inhibitor showed negligible inhibition, paralleling observations on other linear ssDNA species that at least four nucleotides are needed for enzymatic activity.[23]

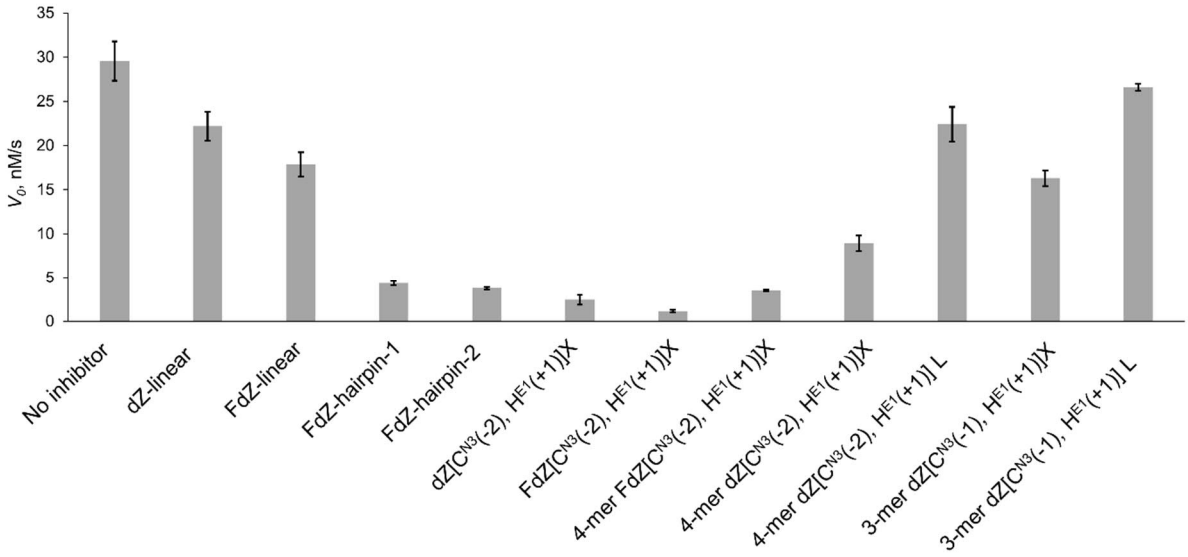



**Figure 3.** Inhibition of A3A-mimic-catalysed deamination of 5'-$T_4$CAT by dZ/FdZ-containing linear and cross-linked oligos.

Initial rate of deamination was measured under the following experimental conditions: 400 µM of 5'-$T_4$CAT, 4 µM of dZ/FdZ-containing oligos, 200 nM of A3A-mimic in a 50 mM sodium phosphate buffer (pH 6.0) containing 100 mM NaCl, 2.5 mM β-mercaptoethanol, 50 µM 3-(trimethylsilyl)-2,2,3,3-tetradeuteropropionic acid (TSP) at 25 °C. Error bars are estimated standard deviations from triplicate measurements. FdZ-hairpin-1: $(GC)_2$-TTFdZ-$(GC)_2$, FdZ-hairpin-2: AAGCT-TTFdZ-AGCTT.

To quantitatively characterise our strongest cross-linked inhibitors, the inhibition constants were calculated by several methods assuming a competitive mode of inhibition[6a] based on experiments in which the concentration of the individual inhibitor was varied, and the residual deamination initial rate of a linear substrate was measured for A3A-mimic and wt-A3A (**Table 1**). Further characterisation of our best inhibitor FdZ[$C^{N3}$(–2),$H^{E1}$(+1)]X was conducted by observing its inhibition of the deamination of dC-hairpin-1, $(GC)_2TTC(GC)_2$, which is a more rapidly deaminated substrate than linear ssDNA by A3A-mimic (**Figure 2**) and also by wild-type A3A.[11c] The kinetic data, measured over time until exhaustion of the substrate, were analysed by global regression using Lambert's W function, which provides superior estimates of observed $K_m$ and $V_{max}$ than non-linear regression analysis of initial rate data or more conventional Lineweaver-Burk/Dixon plots.[24]

Under the given experimental conditions, dC-hairpin-1 is a better substrate for A3A-mimic and wtA3A than the linear oligo 5'-$AT_3CAT_3$ because their $K_m$[11c] are in the low µM range while $k_{cat}$ are very similar. The specificity constant defined by $k_{cat}/K_m$ reinforces that the A3A-mimic and wt-A3A are, respectively, ~25 and 42 times more efficient at deaminating the dC-hairpin than the linear 5'-$AT_3CAT_3$. In the presence of a low concentration of inhibitor, $V_{max}$ was unchanged (**Table 1**), which indicates that FdZ- and dZ-oligos are competitive inhibitors

of A3A-mimic, consistent with the binding of FdZ and dZ in the active site of A3 enzymes observed in crystal structures.[11b, 11c, 22]

**Table 1.** The inhibition constants ($K_i$, μM) for dZ/FdZ-containing oligo inhibitors of A3A-mimic and wt-A3A obtained by various calculation methods.

| Enzyme | A3A-mimic [a] | | wt-A3A [b] |
|---|---|---|---|
| Oligo name | Lineweaver-Burk plot for the substrate/Dixon plot for the inhibitor | Non-linear regression analysis for both substrate and inhibitor | Lambert's W function |
| dZ-linear[11a] | 7.5 ± 1.7 | 7.0 ± 1.3 | -[d] |
| FdZ-linear[11a] | 2.1 ± 0.8 | 4.4 ± 0.7 | 2.4 ± 0.9 |
| dZ[$U^E$(–2),$A^{N3}$(+1)]X[11a] | 0.66 ± 0.14 | 0.69 ± 0.14 | 0.36 ± 0.12 |
| FdZ-hairpin-1[11b, 11c] | 0.28 ± 0.05 | 0.35 ± 0.06 | 0.117 ± 0.015 |
| FdZ[$C^{N3}$(–2),$H^{E1}$(+1)]X | 0.075 ± 0.015 | 0.100 ± 0.016 (0.0068 ± 0.0014)[c] | 0.029 ± 0.005 |
| 4-mer FdZ[$C^{N3}$(–2),$H^{E1}$(+1)]X | 0.25 ± 0.05 | 0.28 ± 0.05 | 0.18 ± 0.13[e] |
| dZ[$C^{N3}$(–1),$H^{E2}$(+1)]X | -[d] | -[d] | No inhibition[f] |
| dZ[$C^{N3}$(–2),$H^{E2}$(+1)]X | -[d] | -[d] | 1.6 ± 0.3 |
| dZ[$C^{N3}$(–1),$H^{E3}$(+1)]X | -[d] | -[d] | No inhibition |
| dZ[$C^{N3}$(–2),$H^{E3}$(+1)]X | -[d] | -[d] | 0.64 ± 0.11 |

a) dZ/FdZ-linear inhibitors were evaluated against 5'-$AT_3CAT_3$ using 50 nM of A3A-mimic; all other inhibitors were evaluated against 5'-$T_4$CAT using 200 nM of A3A-mimic using the same experimental conditions as reported in Figure 3 (Table S10 in SI).

b) 50 mM $Na^+/K^+$ phosphate buffer, pH 7.4 supplemented with 100 mM NaCl, 1 mM TCEP, 100 μM DSS and 10 % $D_2O$; A3A concentration: 140 nM; substrate concentration (dC-hairpin-2, $T(GC)_2TTC(GC)_2T$,): 500 μM. $K_m$ for dC-hairpin-2 against A3A is 33 ± 4 μM, $k_{cat}$ is 0.13 $s^{-1}$, $k_{cat}/K_m$ is 4.5 $s^{-1}$ $mM^{-1}$.

c) Analysis of A3A-mimic-catalysed deamination of dC-hairpin-1 over time using Lambert's W function (Figure S4 and Table S11 in SI). Conditions are 400 – 700 μM of dC-hairpin-2, 20 nM of A3A-mimic in a 50 mM $Na^+/K^+$ phosphate buffer (pH 6.0) containing 100 mM NaF, 1 mM TCEP, 10% $D_2O$ and 100 $\mu$M 3-(trimethylsilyl)-2,2,3,3-tetradeuteropropionic acid (TSP) at 25 °C. $K_m$ for dC-hairpin-1 against A3A-mimic is 6.6 ± 0.7 μM, $k_{cat}$ is 0.296 ± 0.015 $s^{-1}$, $k_{cat}/K_m$ is 44 $s^{-1}$ $mM^{-1}$.

d) not performed.

e) Buffer from b) contained additional additives: 0.5mM EDTA, .4 mM $ZnCl_2$, 0.1% IGEPAL and 200 µg/mL rHSA (recombinant Human Serum Albumin). In this conditions $K_m$ for dC-hairpin-2 against A3A is 12 ± 5 µM.

f) No inhibition at 5 µM concentration of dZ-containing oligonucleotide.

All calculation methods revealed that cross-linked and hairpin dZ/FdZ oligos were more powerful inhibitors of A3A-mimic and wt-A3A than linear oligos. The $K_i$ values calculated for the carbazole-containing FdZ[$C^{N3}$(–2),$H^{E1}$(+1)]X were lowest among inhibitors of both enzymes. The $K_i$ of 29 ± 5 nM for FdZ[$C^{N3}$(–2),$H^{E1}$(+1)]X against wt-A3A obtained using Lambert's W function analysis is around 80 times lower than that for the linear **FdZ**-oligo (2.4 ± 0.9 µM).[6b]

Extension of the linker in the cross-link 2 by using **dH$^{E2}$** and **dH$^{E3}$** nucleotides in dZ[$C^{N3}$(–2),$H^{E2}$(+1)]X and dZ[$C^{N3}$(–2),$H^{E3}$(+1)]X led to the loss of inhibitory potential with only the dZ-oligonucleotide based on **dH$^{E3}$** registering a $K_i$ of 640 nM against wt-A3A, a value worse than those for hairpin inhibitors and for the cross-linked 1 oligonucleotide (**Table 1)**. This demonstrates that the linker length, on the formerly alkyne arm was optimal in oligonucleotides with **dH$^{E1}$**. Indeed, molecular modelling (**Figure 4**) performed for the 4-mer dC[$C^{N3}$(–2),$H^{E1}$(+1)]X in complex with A3A-E72A showed that the linker created with **dH$^{E1}$** did not disturb key interactions of the $dC^0$ and $T^{-1}$ A3A recognition motif with the protein as observed in the crystal structures of A3A with the linear ssDNA and hairpin DNA (inhibitor and substrate binding to wt-A3A and A3A-E72A.[8b, 8d, 11c] In addition, His29 (yellow) hydrogen bonds to a phosphate oxygen linking nucleotides at positions –1 and 0 and weakly to the carbonyl of modified cytosine **dC$^{N3}$** at position –2. His29 is calculated to weakly base-stack with carbazole with a 5.6 Å distance between rings, which are not co-planar. Arg28 makes no contacts with cross-linked 2 oligonucleotide. However, an alternate conformer (not crystallographically observed in A3A and A3A-mimic crystal structures) would place Arg28 in a position to make hydrogen bonds to the phosphate linking nucleotides at –2 and –1. These features contrast with recent crystal structures of A3A with DNA hairpins, in which His29 plays a key role in interacting with DNA including base-stacking with the purine in +1 position while Arg28 forms a cation-π interaction with $T^{-2}$.[11c]

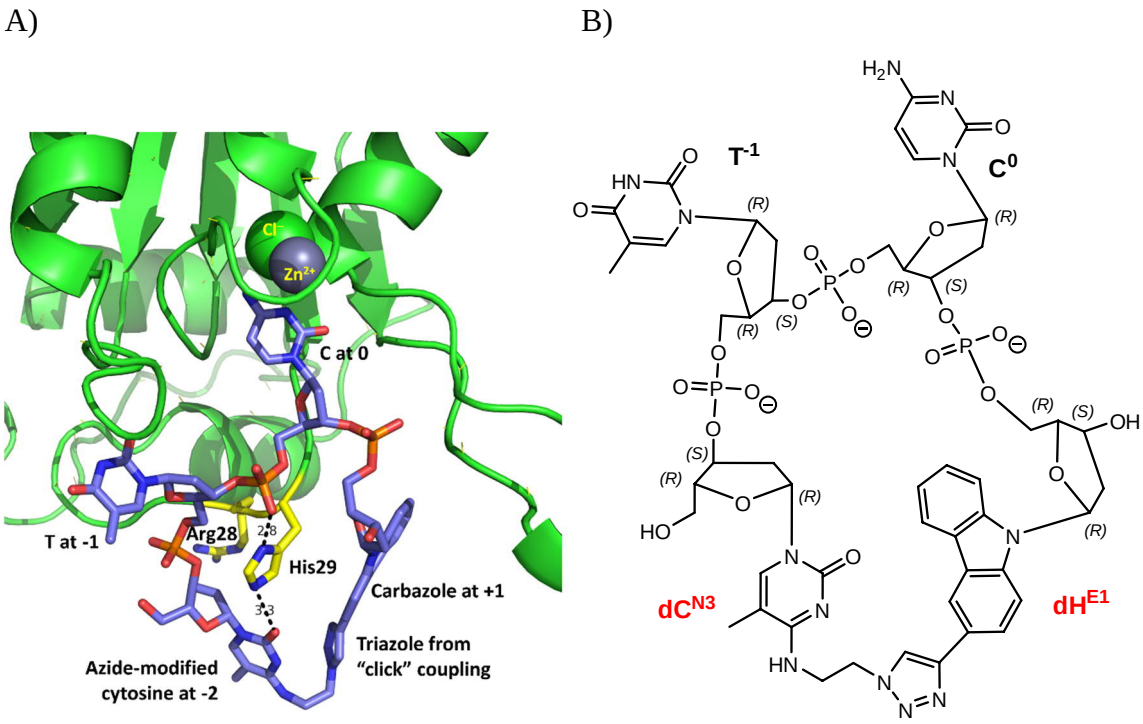



**Figure 4.** A) MM-MD-optimised A3A-E72A (green) in complex with cross-linked 4-mer dC[$C^{N3}$(–2),$H^{E1}$(+1)]X (blue). Key residues of A3A-E72A that interact with hairpin substrates and inhibitors are highlighted in yellow. B) Chemical structure of 4-mer dC[$C^{N3}$(–2),$H^{E1}$(+1)]X used for modelling in A).

*Resistance of cross-linked 2 inhibitors to nucleases*

Nuclease-stability of modified oligos was evaluated using snake venom phosphodiesterase (phosphodiesterase I, Sigma) and compared with degradation of the unmodified sequence 5'-$T_4$CAT. Under the conditions used in this experiment, 5'-$T_4$CAT and the linear 4-mer dZ oligo were completely degraded within 60 min, whereas the cross-linked 4-mer inhibitor FdZ[$C^{N3}$(–2),$H^{E1}$(+1)]X showed significantly enhanced stability towards enzymatic digestion. These data demonstrate an additional benefit of the introduction of a covalent cross-link into DNA-based inhibitors of $A3B_{CTD}$: that is, cross-linked oligos will have extended life-times in biological media in comparison with linear oligos.

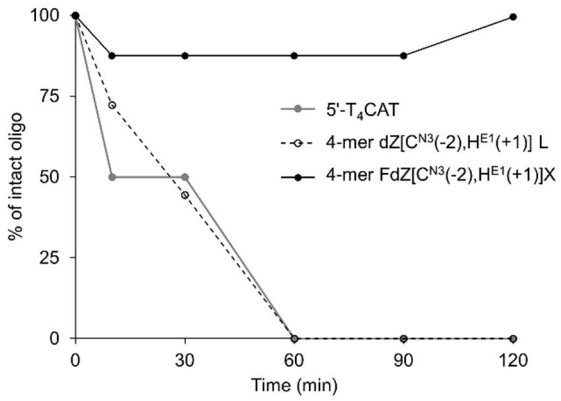



**Figure 5.** Evaluation of nuclease stability of 4-mer cross-linked oligo.

Percentage of intact oligos determined after enzymatic digestion at 37 °C using snake venom phosphodiesterase (phosphodiesterase I, 0.016 units /L, from Sigma) in 120 min. Experiments were repeated at least twice with 6% error for each point (see Supplementary Information for details).

### *Inhibition of A3B$_{CTD}$*

None of the cross-linked and hairpin dZ/FdZ-containing oligos presented in this work showed any substantial inhibition of A3B$_{CTD}$-DM, an enzyme with two mutations distant from the active site, which agrees with a recent report for 5-Me-2'-deoxyzebularine- (5Me-dZ-) containing hairpins.[11d] Such a drastic difference between inhibition of these closely related A3 enzymes points to structural variations around the active site, especially the to-date structurally-uncharacterised substrate-recognition loop 1 as noted earlier.[6c] This loop provides scope for design of A3B-specific inhibitors.

### *Conclusions*

This work demonstrates that cross-link 2 created by CuACC between a carbazole pseudo-nucleoside **dH$^{E1}$** placed in position +1 with 4-azidoethylcytosine **dC$^{N3}$** in position –2 provides not only a faster deaminating DNA substrate but also more potent inhibitors than any others described so far against A3A-mimic and wt-A3A. The inhibitory potential against wt-A3A under these conditions was improved at least 80 times for FdZ[C$^{N3}$(–2),H$^{E1}$(+1)]X in comparison with **FdZ**-linear and four times better than FdZ-hairpin-1. This demonstrates that potency of inhibitors can be further improved through introduction of modifications into DNA close to the dZ an dFdZ enhancing interactions with the proteins. The $K_i$ values for the 4-mer cross-linked oligo were worse than those for the best 9-mer cross-link 2 but still in the nM range (180 nM vs 29 nM, respectively). This means that some, but not all, interactions with the enzyme are lost. The ability of the 4-mer to inhibit A3A-mimic contrasts with recent results

where a loss of inhibitory potential was observed when the DNA was shortened and when it was circularised.[11a, 11d] Moreover, the cross-linked 4-mer oligonucleotide showed enhanced stability towards enzymatic digestion by phosphodiesterase I, which indicates that this inhibitor should have extended lifetime in biological media and thus exhibit its inhibitory effect for a longer period of time. As such, the ability of the 4-mer cross-linked dZ/FdZ-oligonucleotides to inhibit A3A represents an alternative pathway to creation of small-molecule-like A3A inhibitors targeting the active site of the enzyme. None of the small molecules tested so far against A3A[25] reaches the inhibitory potential of dZ/FdZ-DNA-based inhibitors, hairpins with short loops or cross-linked DNA.

The nanomolar inhibitory potential exhibited by our cross-linked oligos shows that properly pre-shaped, cross-linked DNA containing a hydrophobic pseudo-nucleoside in position +1 leads to more potent inhibitors of A3A than linear and hairpin inhibitors and that the better substrate becomes the better inhibitor, which supports our original hypothesis. These findings pave the way for a design of smallish-molecule inhibitors of A3A and to use them as tools to inhibit A3A in cells and in animal models.

**Acknowledgements**

NMR and mass spectrometry facilities at Massey University and the assistance of Dr Patrick J. B. Edwards and Mr David Lun are gratefully acknowledged. We thank Prof. Reuben S. Harris (HHMI and University of Texas Health, San Antonio, TX, USA) and members of his cancer research program for many helpful discussions. Authors would like to thank Dr Maksim V. Kvach for discussions and careful reading of the manuscript. We are grateful for the financial support provided by the Health Research Council of New Zealand in partnership with Breast Cancer Research (grant 20/1355), Kiwi Innovation Network with Massey Ventures Limited (grant MU002391), Maurice Wilkins Centre for Molecular Biodiscovery, as well as the support of the School of Natural Sciences, Massey University.

**Supplementary information:**

Supplementary experimental details about the synthesis of nucleosides and modified oligos and enzymatic assays; and the sequence alignment of proteins used in this study, $^{1}H$, $^{13}C$, $^{31}P$ NMR, IR and HRMS (ESI) spectra of new compounds synthesised, RP-HPLC profiles and HRMS (ESI) spectra of oligos.

**Conflict of interests:** Authors filed a patent on single-stranded DNA enzyme inhibitors describing inhibitors created in this work.[11b]

## References


[1] a) R. S. Harris, K. N. Bishop, A. M. Sheehy, H. M. Craig, S. K. Petersen-Mahrt, I. N. Watt, M. S. Neuberger, M. H. Malim, *Cell* **2003**, *113*, 803-809; b) R. S. Harris, M. T. Liddament, *Nat Rev Immunol* **2004**, *4*, 868-877; c) T. Izumi, K. Shirakawa, A. Takaori-Kondo, *Mini-Rev. Med. Chem.* **2008**, *8*, 231-238; d) B. Mangeat, P. Turelli, G. Caron, M. Friedli, L. Perrin, D. Trono, *Nature* **2003**, *424*, 99-103; e) R. S. Harris, J. P. Dudley, *Virology* **2015**, *479-480*, 131-145.

[2] a) A. M. Green, M. D. Weitzman, *DNA Repair (Amst)* **2019**, *83*, 102700; b) C. Swanton, N. McGranahan, G. J. Starrett, R. S. Harris, *Cancer Disc.* **2015**, *5*, 704-712; c) L. B. Alexandrov, S. Nik-Zainal, D. C. Wedge, S. A. Aparicio, S. Behjati, A. V. Biankin, G. R. Bignell, N. Bolli, A. Borg, A. L. Borresen-Dale, S. Boyault, B. Burkhardt, A. P. Butler, C. Caldas, H. R. Davies, C. Desmedt, R. Eils, J. E. Eyfjord, J. A. Foekens, M. Greaves, F. Hosoda, B. Hutter, T. Ilicic, S. Imbeaud, M. Imielinski, N. Jager, D. T. Jones, D. Jones, S. Knappskog, M. Kool, S. R. Lakhani, C. Lopez-Otin, S. Martin, N. C. Munshi, H. Nakamura, P. A. Northcott, M. Pajic, E. Papaemmanuil, A. Paradiso, J. V. Pearson, X. S. Puente, K. Raine, M. Ramakrishna, A. L. Richardson, J. Richter, P. Rosenstiel, M. Schlesner, T. N. Schumacher, P. N. Span, J. W. Teague, Y. Totoki, A. N. Tutt, R. Valdes-Mas, M. M. van Buuren, L. van 't Veer, A. Vincent-Salomon, N. Waddell, L. R. Yates, J. Zucman-Rossi, P. A. Futreal, U. McDermott, P. Lichter, M. Meyerson, S. M. Grimmond, R. Siebert, E. Campo, T. Shibata, S. M. Pfister, P. J. Campbell, M. R. Stratton, *Nature* **2013**, *500*, 415-421.

[3] a) C. Martínez-Ruiz, J. R. M. Black, C. Puttick, M. S. Hill, J. Demeulemeester, E. Larose Cadieux, K. Thol, T. P. Jones, S. Veeriah, C. Naceur-Lombardelli, A. Toncheva, P. Prymas, A. Rowan, S. Ward, L. Cubitt, F. Athanasopoulou, O. Pich, T. Karasaki, D. A. Moore, R. Salgado, E. Colliver, C. Castignani, M. Dietzen, A. Huebner, M. Al Bakir, M. Tanić, T. B. K. Watkins, E. L. Lim, A. M. Al-Rashed, D. Lang, J. Clements, D. E. Cook, R. Rosenthal, G. A. Wilson, A. M. Frankell, S. de Carné Trécesson, P. East, N. Kanu, K. Litchfield, N. J. Birkbak, A. Hackshaw, S. Beck, P. Van Loo, M. Jamal-Hanjani, N. McGranahan, C. Swanton, M. A. Bakir, E. L. Lim, A. M. Frankell, K. Litchfield, N. J. Birkbak, P. Van Loo, J. F. Lester, A. Bajaj, A. Nakas, A. Sodha-Ramdeen, K. Ang, M. Tufail, M. F. Chowdhry, M. Scotland, R. Boyles, S. Rathinam, C. Wilson, D. Marrone, S. Dulloo, D. A. Fennell, G. Matharu, J. A. Shaw, J. Riley, L. Primrose, E. Boleti, H. Cheyne, M. Khalil, S. Richardson, T. Cruickshank, G. Price, K. M. Kerr, S. Benafif, K. Gilbert, B. Naidu, A. J. Patel, A. Osman, C. Lacson, G. Langman, H. Shackleford, M. Djearaman, S. Kadiri, G. Middleton, A. Leek, J. D. Hodgkinson, N. Totten, A. Montero, E. Smith, E. Fontaine, F. Granato, H. Doran, J. Novasio, K. Rammohan, L. Joseph, P. Bishop, et al., *Nature* **2023**; b) M. Petljak, A. Dananberg, K. Chu, E. N. Bergstrom, J. Striepen, P. von Morgen, Y. Chen, H. Shah, J. E. Sale, L. B. Alexandrov, M. R. Stratton, J. Maciejowski, *Nature* **2022**, *607*, 799-807; c) M. C. Jarvis, M. A. Carpenter, N. A. Temiz, M. R. Brown, K. A. Richards, P. P. Argyris, W. L. Brown, D. Yee, R. S. Harris, *bioRxiv* **2022**, 2022.2004.2026.489523; d) S. Sharma, S. K. Patnaik, R. Thomas Taggart, E. D. Kannisto, S. M. Enriquez, P. Gollnick, B. E. Baysal, *Nat. Commun.* **2015**, *6*, 6881; e) A. Maiti, S. Hou, C. A. Schiffer, H. Matsuo, *Curr. Opin. Struct. Biol.* **2021**, *67*, 195-204.

[4] E. K. Law, R. Levin-Klein, M. C. Jarvis, H. Kim, P. P. Argyris, M. A. Carpenter, G. J. Starrett, N. A. Temiz, L. K. Larson, C. Durfee, M. B. Burns, R. I. Vogel, S. Stavrou, A. N. Aguilera, S. Wagner, D. A. Largaespada, T. K. Starr, S. R. Ross, R. S. Harris, *J. Exp. Med.* **2020**, *217*, e20200261.

[5] a) M. B. Burns, L. Lackey, M. A. Carpenter, A. Rathore, A. M. Land, B. Leonard, E. W. Refsland, D. Kotandeniya, N. Tretyakova, J. B. Nikas, D. Yee, N. A. Temiz, D. E. Donohue, R. M. McDougle, W. L. Brown, E. K. Law, R. S. Harris, *Nature* **2013**, *494*, 366-370; b) R. S. Harris, *Genome Med* **2013**, *5*, 87.

[6] a) 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; b) M. V. Kvach, F. M. Barzak, S. Harjes, H. A. M. Schares, H. M. Kurup, K. F. Jones, L. Sutton, J. Donahue, R. T. D'Aquila, G. B. Jameson, D. A. Harki, K. L. Krause, E. Harjes, V. V. Filichev, *ChemBioChem* **2020**, *21*, 1028-1035; c) F. M. Barzak, S. Harjes, M. V. Kvach, H. M. Kurup, G. B. Jameson, V. V. Filichev, E. Harjes, *Org. Biomol. Chem.* **2019**, *17*, 9435 - 9441.

[7] a) J. J. Barchi, A. Haces, V. E. Marquez, J. J. McCormack, *Nucleosides Nucleotides* **1992**, *11*, 1781-1793; b) V. E. Marquez, J. J. Barchi, Jr., J. A. Kelley, K. V. Rao, R. Agbaria, T. Ben-Kasus, J. C. Cheng, C. B. Yoo, P. A. Jones, *Nucleosides Nucleotides Nucl. Acids* **2005**, *24*, 305-318.

[8] a) S. Harjes, G. B. Jameson, V. V. Filichev, P. J. B. Edwards, E. Harjes, *Nucleic Acids Res.* **2017**, *45*, 5602-5613; b) T. Kouno, T. V. Silvas, B. J. Hilbert, S. M. D. Shandilya, M. F. Bohn, B. A. Kelch, W. E. Royer, M. Somasundaran, N. Kurt Yilmaz, H. Matsuo, C. A. Schiffer, *Nat. Commun.* **2017**, *8*, 15024; c) A. Maiti, W. Myint, T. Kanai, K. Delviks-Frankenberry, C. Sierra Rodriguez, V. K. Pathak, C. A. Schiffer, H. Matsuo, *Nat. Commun.* **2018**, *9*, 2460-2460; d) 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.

[9] R. Buisson, A. Langenbucher, D. Bowen, E. E. Kwan, C. H. Benes, L. Zou, M. S. Lawrence, *Science* **2019**, *364*, eaaw2872.

[10] a) A. Langenbucher, D. Bowen, R. Sakhtemani, E. Bournique, J. F. Wise, L. Zou, A. S. Bhagwat, R. Buisson, M. S. Lawrence, *Nat Commun* **2021**, *12*, 1602; b) S. M. Wörmann, A. Zhang, F. I. Thege, R. W. Cowan, D. N. Rupani, R. Wang, S. L. Manning, C. Gates, W. Wu, R. Levin-Klein, K. I. Rajapakshe, M. Yu, A. S. Multani, Y. Kang, C. M. Taniguchi, K. Schlacher, M. D. Bellin, M. H. G. Katz, M. P. Kim, J. B. Fleming, S. Gallinger, R. Maddipati, R. S. Harris, F. Notta, S. R. Ross, A. Maitra, A. D. Rhim, *Nat Cancer* **2021**, *2*, 1338-1356.

[11] a) H. M. Kurup, M. V. Kvach, S. Harjes, F. M. Barzak, G. B. Jameson, E. Harjes, V. V. Filichev, *Biochemistry* **2022**, *61*, 2568-2578; b) H. M. Kurup, M. V. Kvach, S. Harjes, F. M. Barzak, Y. Su, G. B. Jameson, E. Harjes, V. V. Filichev, **2022**, PCT/IB2022/050656 - P000179WO; c) S. Harjes, H. M. Kurup, A. E. Rieffer, M. Bayarjagal, J. Filitcheva, Y. Su, T. K. Hale, V. V. Filichev, E. Harjes, R. S. Harris, G. B. Jameson, *bioRxiv* **2023**, 2023.2002.2017.528918; d) J. C. Serrano, D. von Trentini, K. N. Berrios, A. Barka, I. J. Dmochowski, R. M. Kohli, *ACS Chem Biol* **2022**, *17*, 3379-3388.

[12] a) B. Leonard, S. N. Hart, M. B. Burns, M. A. Carpenter, N. A. Temiz, A. Rathore, R. I. Vogel, J. B. Nikas, E. K. Law, W. L. Brown, Y. Li, Y. Zhang, M. J. Maurer, A. L. Oberg, J. M. Cunningham, V. Shridhar, D. A. Bell, C. April, D. Bentley, M. Bibikova, R. K. Cheetham, J.-B. Fan, R. Grocock, S. Humphray, Z. Kingsbury, J. Peden, J. Chien, E. M. Swisher, L. C. Hartmann, K. R. Kalli, E. L. Goode, H. Sicotte, S. H. Kaufmann,

R. S. Harris, *Cancer Res.* **2013**, *73*, 7222-7231; b) K. Chan, S. A. Roberts, L. J. Klimczak, J. F. Sterling, N. Saini, E. P. Malc, J. Kim, D. J. Kwiatkowski, D. C. Fargo, P. A. Mieczkowski, G. Getz, D. A. Gordenin, *Nat. Genet.* **2015**, *47*, 1067-1072.

[13] I.-J. L. Byeon, C.-H. Byeon, T. Wu, M. Mitra, D. Singer, J. G. Levin, A. M. Gronenborn, *Biochemistry* **2016**, *55*, 2944-2959.

[14] a) J. F. Mouscadet, C. Ketterle, H. Goulaouic, S. Carteau, F. Subra, M. Lebret, C. Auclair, *Biochemistry* **1994**, *33*, 4187-4196; b) S. Nakamura, H. Kawabata, K. Fujimoto, *ChemComm* **2017**, *53*, 7616-7619; c) K. Fujimoto, K. Hiratsuka-Konishi, T. Sakamoto, T. Ohtake, K.-i. Shinohara, Y. Yoshimura, *Mol. Biosyst.* **2012**, *8*, 491-494; d) A. S. Gouda, Ł. Przypis, K. Walczak, P. T. Jørgensen, J. Wengel, *Org. Biomol. Chem.* **2020**, *18*, 6935-6948; e) M. D. Matteucci, U. von Krosigk, *Tetrahedron Lett.* **1996**, *37*, 5057-5060.

[15] a) L. S. Tsutsumi, D. Gündisch, D. Sun, *Curr Top Med Chem* **2016**, *16*, 1290-1313; b) A. Głuszyńska, *European Journal of Medicinal Chemistry* **2015**, *94*, 405-426; c) S. Issa, A. Prandina, N. Bedel, P. Rongved, S. Yous, M. Le Borgne, Z. Bouaziz, *J Enzyme Inhib Med Chem* **2019**, *34*, 1321-1346.

[16] M. A. Fomich, M. V. Kvach, M. J. Navakouski, C. Weise, A. V. Baranovsky, V. A. Korshun, V. V. Shmanai, *Org. Lett.* **2014**, *16*, 4590-4593.

[17] A. Kiviniemi, P. Virta, H. Lonnberg, *Bioconjugate Chem.* **2008**, *19*, 1726-1734.

[18] G. Pourceau, A. Meyer, J. J. Vasseur, F. Morvan, *J. Org. Chem.* **2009**, *74*, 6837-6842.

[19] a) S. A. Ingale, H. Mei, P. Leonard, F. Seela, *J. Org. Chem.* **2013**, *78*, 11271-11282; b) M. I. Fatthalla, E. B. Pedersen, *Eur. J. Org. Chem.* **2016**, n/a-n/a; c) S. A. Ingale, P. Leonard, H. Yang, F. Seela, *Org. Biomol. Chem.* **2014**, *12*, 8519-8532.

[20] A. G. Myers, S. D. Goldberg, *Angew. Chem. Int. Ed. Engl.* **2000**, *39*, 2732-2735.

[21] a) V. V. Filichev, E. B. Pedersen, *J. Am. Chem. Soc.* **2005**, *127*, 14849-14858; b) M. Rist, N. Amann, H. A. Wagenknecht, *Eur. J. Org. Chem.* **2003**, 2498-2504; c) V. V. Filichev, I. V. Astakhova, A. D. Malakhov, V. A. Korshun, E. B. Pedersen, *Chem.—Eur. J.* **2008**, *14*, 9968-9980.

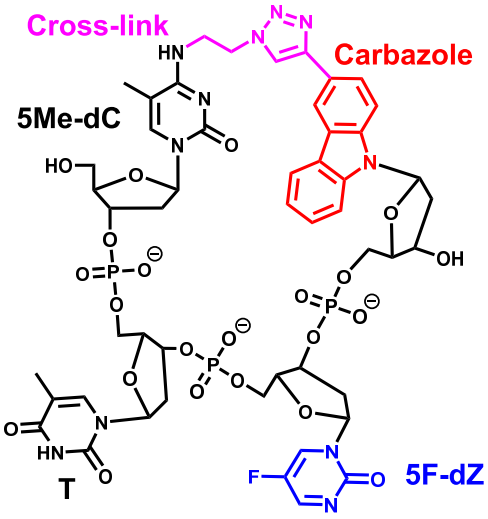
[22] A. Maiti, A. K. Hedger, W. Myint, V. Balachandran, J. K. Watts, C. A. Schiffer, H. Matsuo, *Nat Commun* **2022**, *13*, 7117.

[23] 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**, *13*, 2427-2432.

[24] C. T. Goudar, S. K. Harris, M. J. McInerney, J. M. Suflita, *J Microbiol Methods* **2004**, *59*, 317-326.

[25] a) J. J. King, F. Borzooee, J. Im, M. Asgharpour, A. Ghorbani, C. P. Diamond, H. Fifield, L. Berghuis, M. Larijani, *ACS Pharmacol Transl Sci* **2021**, *4*, 1390-1407; b) Y.-H. Zhang, X.-C. Guo, J.-B. Zhong, D.-X. Zhong, X.-H. Huang, Z.-Y. Fang, C. Zhang, Y.-J. Lu, *ChemistrySelect* **2022**, *7*, e202201456.

**Graphical Abstract:**




Inhibitory potential of cross-linked oligonucleotide against DNA mutator enzyme APOBEC3A was improved by incorporation of carbazole pseudonucleoside reaching $K_i$ of 29 nM.