Supporting Information

Evaluation of APOBEC3B recognition motifs by NMR reveals preferred substrates

Manjuan Liu,* Aurélie Mallinger, Marcello Tortorici, Yvette Newbatt, Meiron Richards, Amin Mirza, Rob L. M. van Montfort, Rosemary Burke, Julian Blagg,* Teresa Kaserer*

Cancer Research UK Cancer Therapeutics Unit, The Institute of Cancer Research, London, SM2 5NG, UK

*Corresponding authors:

E-Mail: Maggie.Liu@icr.ac.uk, Julian.Blagg@icr.ac.uk, Teresa.Kaserer@icr.ac.uk

Experimental Methods

Protein cloning, expression, and purification

The coding sequence for residues 193-382 of A3B CTD was amplified by PCR and cloned into a pET28a vector that encodes a C-terminal 6X His-tag. The A3B CTD construct 193-382 was expressed in BL21 DE3 *Escherichia coli* cells that were grown in LB medium + Kanamycin to an OD of 0.6 and induced overnight with 0.3 mM IPTG at 18 °C. The cells were then centrifuged and the dried cell pellet corresponding to three liters of cell culture was snap frozen in dry ice and stored at -80°C. Cell pellets were then resuspended in 5 volumes of cold lysis buffer (50 mM Tris pH 8.0, 300 mM NaCl, 5% (*v/v*) Glycerol, 20mM Imidazole) containing 1x complete EDTA-free protease inhibitors. The cells were lysed using an Emulsiflex C3 cell homogenizer at a pressure between 10000 and 15000 psi. To ensure proper lysis was occurring, the cell suspension was passed four times through the cell homogenizer. The lysate was then clarified by centrifugation at 60000 RCF for 40 minutes at 4°C. Following centrifugation, the supernatant was loaded on a 5 mL HisTrap column previously equilibrated on lysis buffer. Nucleic acid contamination was removed by washing the resin-bound material with lysis buffer + 1M KCl. Non-specifically bound material was removed by washing with 5 column volumes of lysis buffer plus 50 mM imidazole. Elution was performed in lysis buffer containing 250 mM imidazole.

The HisTrap eluate was subsequently concentrated to 5.0 mL and applied to a Superdex 75 HR 26/60 column equilibrated in 25 mM Hepes pH 7.5, 150 mM NaCl, 1 mM DTT. Liquid chromatography/mass spectrometry (LC/MS) was carried out on the protein as described in Newbatt et al.[1] Fractions containing

pure A3B CTD were pooled, concentrated to 10 mg/mL, snap frozen in dry ice and stored at -80 °C.

ssDNA

ssDNA were synthesized and HPLC purified by Dharmacon, Inc (up to four nucleotides) and Eurofins Genomics (longer than four nucleotides).

Real-time NMR spectroscopy

ssDNA oligonucleotide solids were dissolved in D_2O to obtain a stock solution. The D_2O stock was added to 25 mM Hepes buffer (pH 7.5 with 150 mM NaCl and 4 mM DTT). The deamination reaction was started by adding a 100 μM A3B CTD stock in the same buffer to this solution. The final composition of the reaction mixture contained 0.5 mM ssDNA oligonucleotide and 2 μM A3B CTD, a ratio of 250:1. All NMR data were recorded using a Bruker Avance 500 spectrometer equipped with a 1.7 mm TXI microprobe (Bruker biospin). Sequential ^{1}H-NMR spectra were acquired at 295 K every 10 minutes with the same NMR sample resident in the magnet. The initial rates of the deamination reactions were determined from the intensity of the H-5 resonance of the dC nucleotide (example spectra in Figures S4, S5 and S6), which were plotted as a function of reaction time. All data were analysed using MestReNova 12 reaction monitor module, the detailed data fitting method is demonstrated in Figure S7. At least two independent measurements were carried out (Table S1).

Fluorescence-based deamination assay

The fluorescence-based DNA cytosine deamination assay[2] was carried out (n=2) in a 384 well black proxiplate (Perkin Elmer, part number 6008289). Recombinant human A3B CTD protein 193-382aa was diluted with assay buffer (25 mM Mes, pH 6.4, 30 mM NaCl, 1 mM DTT and 0.1% triton X-100) to a final assay concentration of 15 nM in 2.4 µL. The enzyme was incubated with a final assay concentration of 375 nM ssDNA substrate 5'-6-FAM-TTATTCATAT-BHQ1-3' (Integrated DNA Technologies) and 0.2 units of UDG (NEB) both diluted together in 2.4 µL assay buffer. The plate was then centrifuged for 1 min at 1000 rpm. The reaction was incubated at 37°C for 60 mins. After incubation, 1 µL of 1.2 M NaOH was added, followed by a plate centrifugation as above and incubation at 37°C for 30 min. A neutralisation solution of 4 µL of 2 M Tris HCL (pH 7.4) was then added and the plate centrifuged as above. Relative deaminase activity was quantified by reading fluorescence with an excitation at 490 nm and emission at 520 nm on an EnVision multimode plate reader (Perkin Elmer).

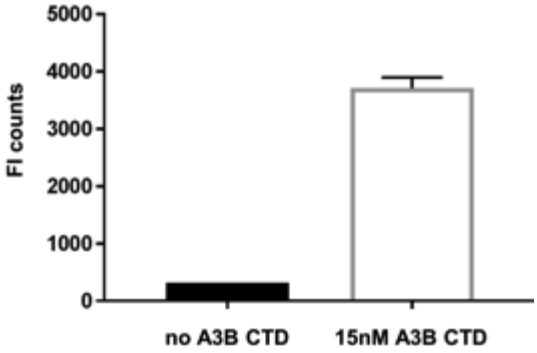

Figure S1. A3B CTD is functional in a fluorescence-based assay. In the absence of A3B CTD, the fluorescence signal of the FAM-labelled 10-mer 5'-6-FAM-TTATTCATAT-BHQ1-3' DNA substrate is quenched. Deamination by A3B CTD and subsequent break of the DNA substrate leads to loss of proximity of the quencher and the fluorescence signal increases as a consequence.

Table S1. Deamination rates of investigated oligonucleotides.

Comment	Sequence	n	Initial rate ± SD (mM*h^{-1})
10-mer benchmark	5'-TTATTCATAT-3'	3	0.174 ± 0.042
Reported preferred A3B substrate[3, 4]	5'-TCA-3'	2	0.007 ± 0.001
Mononucleotide	dCMP	2	0.000 ± 0.000
(+) Side only oligonucleotides	5'-CA-3'	2	0.000 ± 0.000
	5'-CAT-3'	3	0.023 ± 0.001
	5'-CATT-3'	2	0.020 ± 0.007
	5'-CATTT-3'	2	0.026 ± 0.001
(-) Side only oligonucleotides	5'-TC-3'	2	0.000 ± 0.000
	5'-TTC-3'	4	0.009 ± 0.002
	5'-ATTC-3'	3	0.011 ± 0.001
	5'-TATTC-3'	2	0.014 ± 0.005
Impact of +1 nucleotide	5'-TTCA-3'	3	0.107 ± 0.014
	5'-ATTCA-3'	7	0.230 ± 0.040
Impact of (-) side nucleotides	5'-TCAT-3'	3	0.011 ± 0.001
	5'-TTCAT-3'	3	0.129 ± 0.013
Directionality of ssDNA	5'-ACTT-3'	2	0.001 ± 0.000
Impact of -2 base	5'-ATCA-3'	3	0.205 ± 0.041
	5'-CTCA-3'	3	0.171 ±0.034
	5'-GTCA-3'	3	0.498 ± 0.179
Impact of -1 base	5'-AACA-3'	3	0.003 ± 0.001
	5'-ACCA-3'	3	0.024 ± 0.005
	5'-AGCA-3'	3	0.004 ± 0.001
Impact of +1 base	5'-ATCT-3'	3	0.151 ± 0.025
	5'-ATCC-3'	3	0.213 ±0.048
	5'-ATCG-3'	3	0.438 ± 0.243
Impact of -2/+1G	5'-GTCG-3'	3	0.398 ± 0.175
	5'-CGT-3'	3	0.093 ± 0.021
Drug resistant mutations	5'-TGATTTCAGA-3'	3	0.189 ± 0.0042
	5'-TAGAATCATT-3'	3	0.197 ± 0.059
	5'-GAAGTCGTAG-3'	3	0.233 ± 0.039
	5'-ACTATCGCTG-3'	3	0.298 ± 0.070
	5'-ATTACAGAAT-3'	3	0.000 ± 0.000

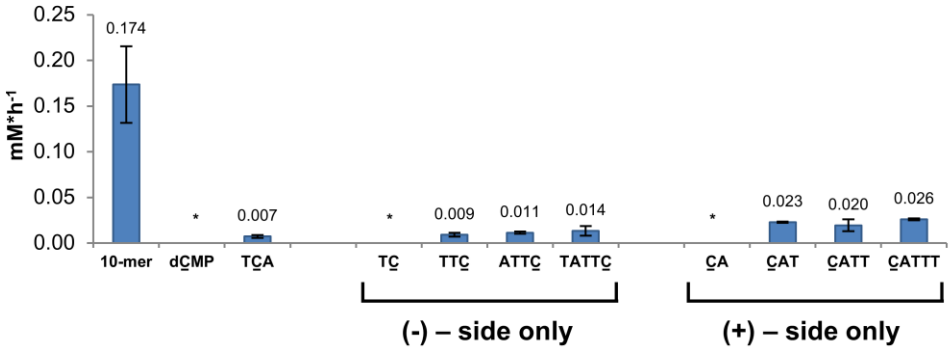

Figure S2. Initial deamination rates of the 10-mer 5'-TTATTCATAT-3', dCMP, 5'-TCA-3', and (-) - and (+) - side only oligonucleotides. *no deamination observed after 20 hours.

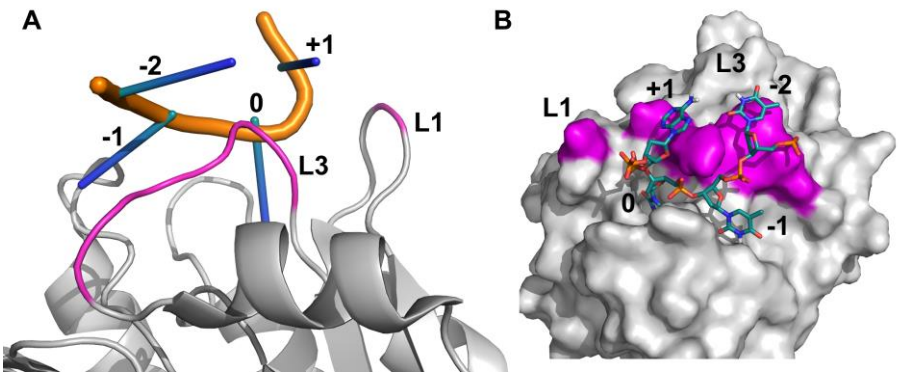

Figure S3. In the A3B CTD (gray cartoon (A) or surface (B))–ssDNA (orange and blue cartoon (A) or sticks (B)) complex (PDB entry 5TD5[4]), A3B CTD loop 1 (L1) residues 242 to 250 are replaced by a single serine. In addition, A3B loop 3 (L3) is replaced by the A3A L3. The engineered protein loops (highlighted in magenta) are proximal to the ssDNA bases at positions -2 and +1.

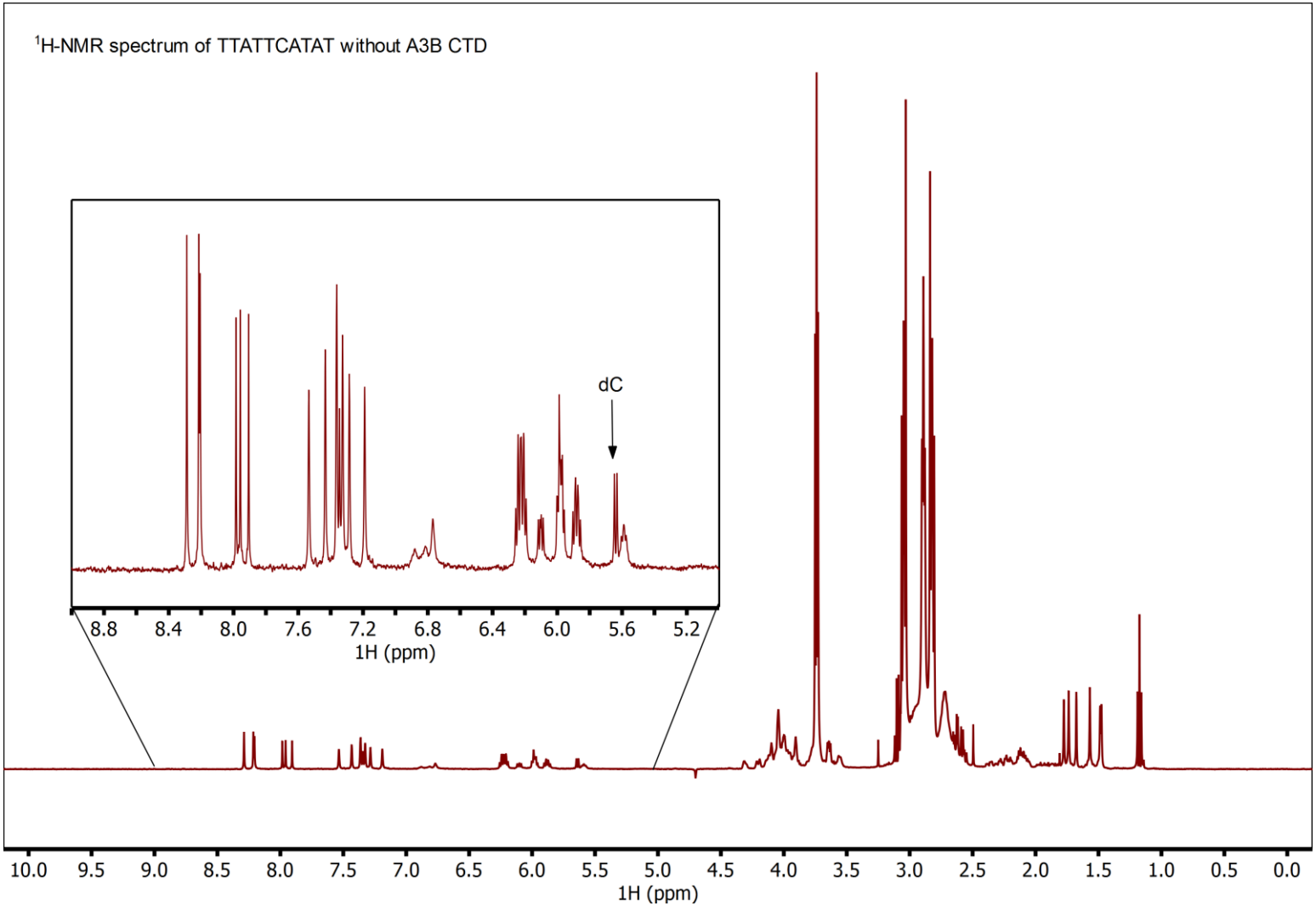

Figure S4. ^{1}H-NMR spectrum of 5'-TTATTCATAT-3' without A3B CTD.

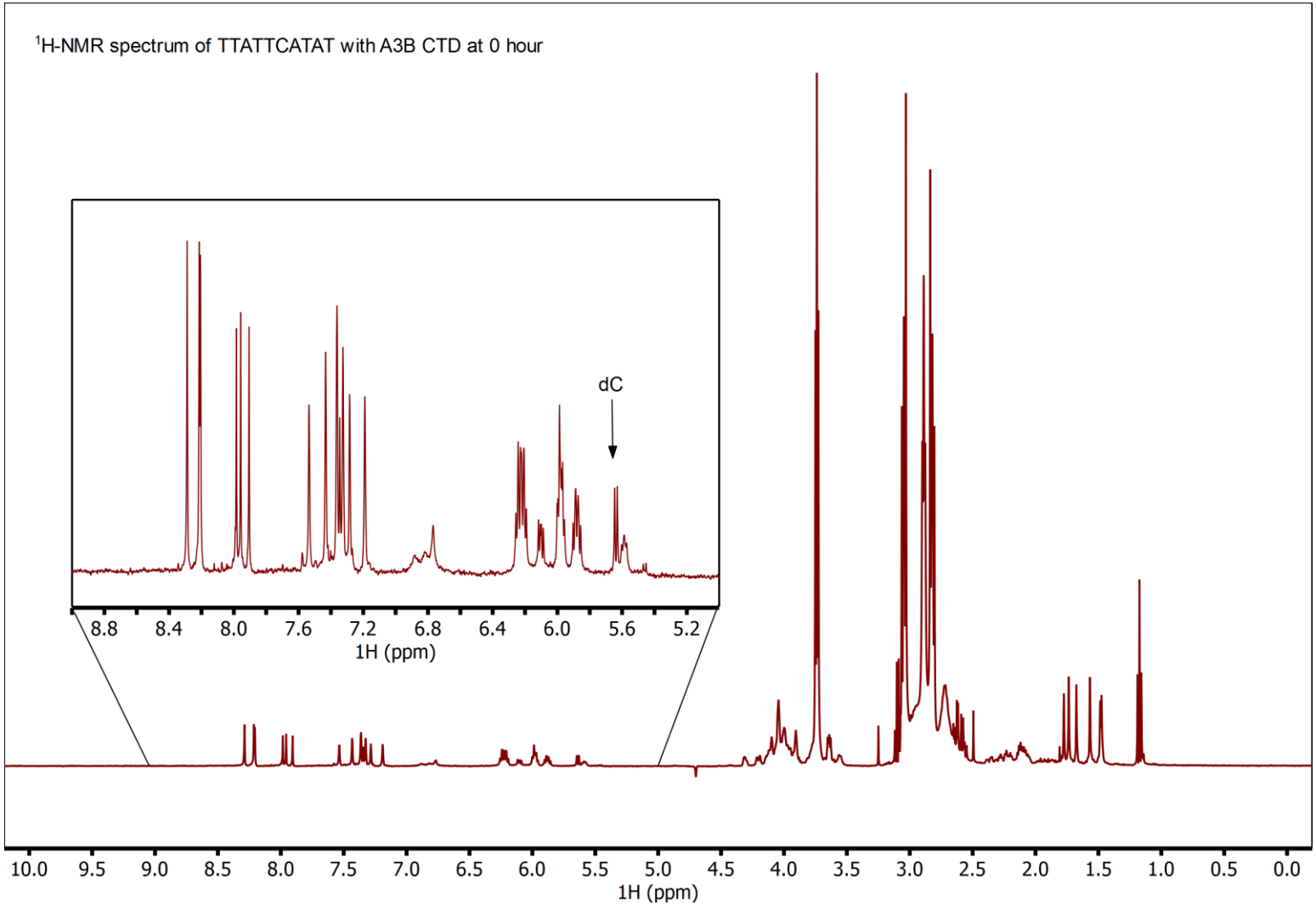

Figure S5. ^{1}H-NMR spectrum of 5'-TTATTCATAT-3' with A3B CTD at 0 hour.

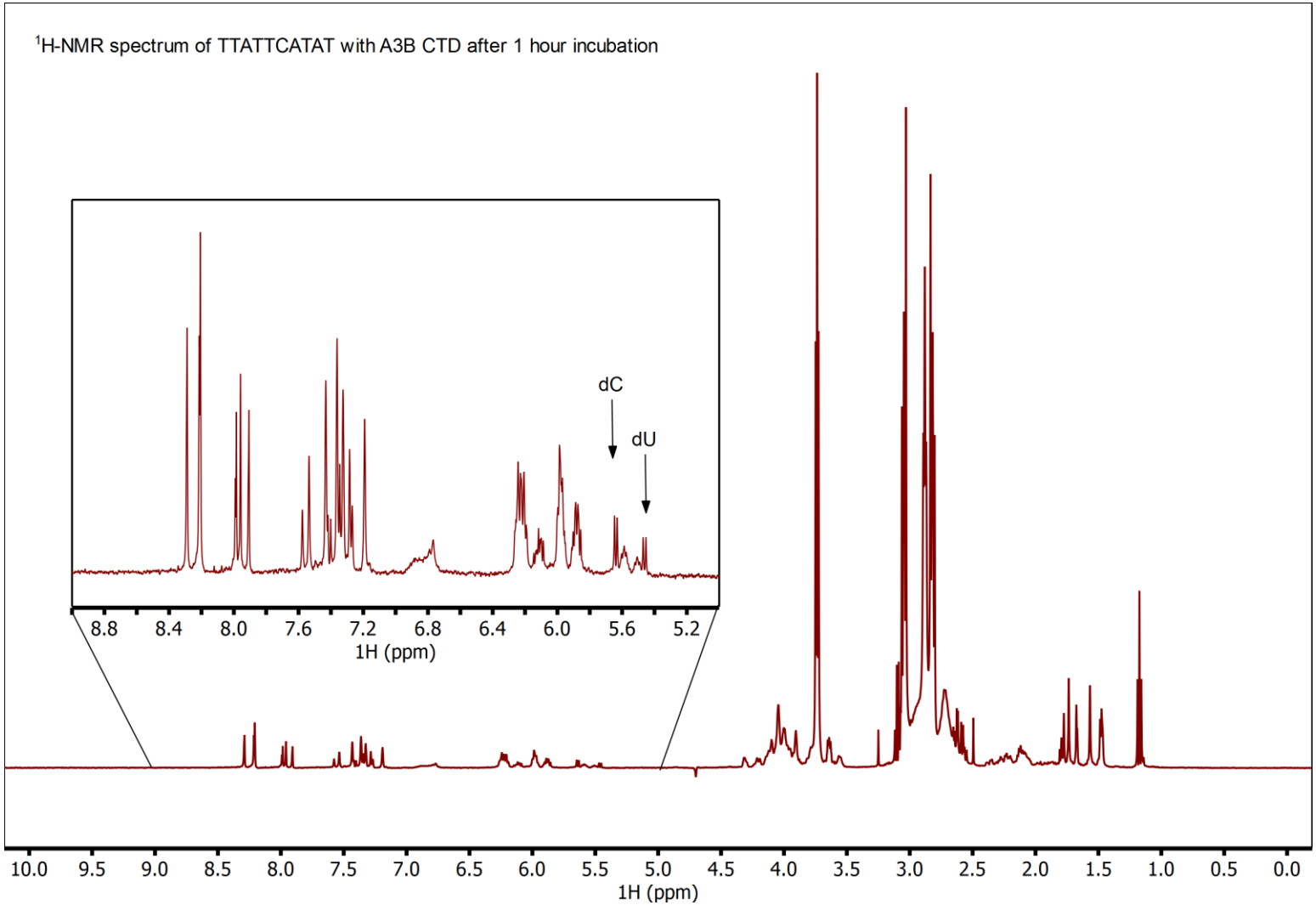

Figure S6. ^{1}H-NMR spectrum of 5'-TTATTCATAT-3' with A3B CTD after 1 hour incubation.

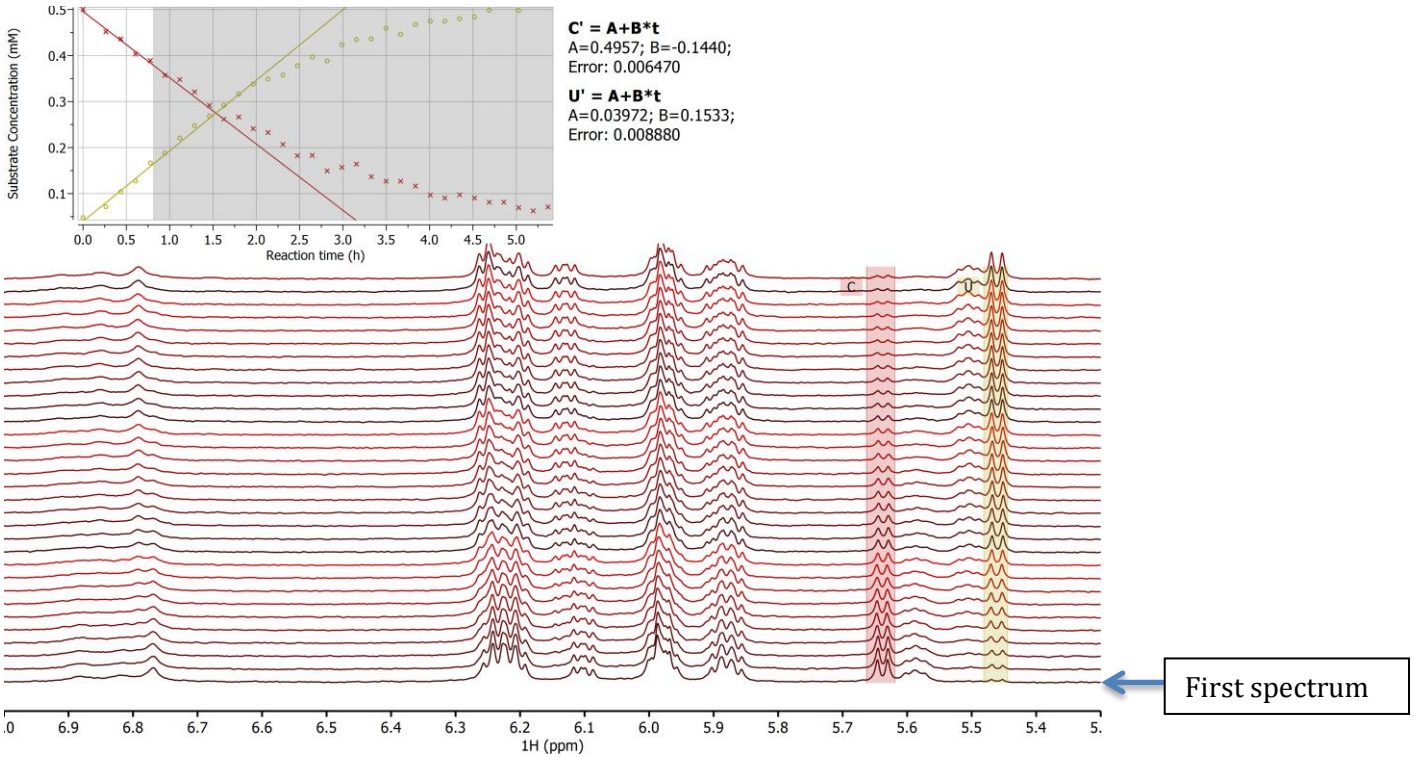

Figure S7. 5'-TTATTCATAT-3' deamination data analysis exemplar demonstrating data fitting by ***Mnova 12 reaction monitor module***.

1. A series of ^{1}H-NMR spectra obtained at different time points from the same sample were loaded to generate the stacked spectra.
2. Two regions were selected manually to cover the NMR signal of the H-5 resonance of the dC nucleotide and the corresponding signal of the dU nucleotide.
3. The integral of the dC signal from the first spectrum (time 0) was calibrated to represent 0.5 mM of concentration and all other integrals were referenced to this value using ***Response Factor mode***.
4. The time points were generated automatically from data acquisition time stamp.
5. The Concentrations were plotted as a function of reaction time; the data was fitted with ***Linear Fit Function*** using only data points within 30% substrate conversion.
6. The dC reduction rate was reported as the deamination rate.

References

1. Newbatt, Y., Hardcastle, A., McAndrew, P. C., Strover, J. A., Mirza, A., Morgan, G. J., Burke, R., Davies, F. E., Collins, I., and van Montfort, R. L. M. (2012) Identification of autophosphorylation inhibitors of the inositol-requiring enzyme 1 alpha (IRE1α) by high-throughput screening using a DELFIA assay. *J. Biomol. Screening 18*, 298-308.

2. Li, M., Shandilya, S. M. D., Carpenter, M. A., Rathore, A., Brown, W. L., Perkins, A. L., Harki, D. A., Solberg, J., Hook, D. J., Pandey, K. K., Parniak, M. A., Johnson, J. R., Krogan, N. J., Somasundaran, M., Ali, A., Schiffer, C. A., and Harris, R. S. (2012) First-in-class small molecule inhibitors of the single-strand DNA cytosine deaminase APOBEC3G. *ACS Chem. Biol. 7*, 506-517.

3. Burns, M. B., Lackey, L., Carpenter, M. A., Rathore, A., Land, A. M., Leonard, B., Refsland, E. W., Kotandeniya, D., Tretyakova, N., Nikas, J. B., Yee, D., Temiz, N. A., Donohue, D. E., McDougle, R. M., Brown, W. L., Law, E. K., and Harris, R. S. (2013) APOBEC3B is an enzymatic source of mutation in breast cancer. *Nature 494*, 366-370.

4. Shi, K., Carpenter, M. A., Banerjee, S., Shaban, N. M., Kurahashi, K., Salamango, D. J., McCann, J. L., Starrett, G. J., Duffy, J. V., Demir, Ö., Amaro, R. E., Harki, D. A., Harris, R. S., and Aihara, H. (2016) Structural basis for targeted DNA cytosine deamination and mutagenesis by APOBEC3A and APOBEC3B. *Nat. Struct. Mol. Biol. 24*, 131-139.