Electronic Supplementary Information

Construction of a G-quadruplex-specific DNA endonuclease

Dung Thanh Dang[abc], Le Tuan Anh Nguyen[a], Tinh Tuom Thi Truong[bd], Hoang Duc Nguyen[bd] and Anh Tuân Phan[ae]*

a. School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore

b. University of Science, VNUHCM, Ho Chi Minh City, Vietnam

c. Ho Chi Minh City Open University, Ho Chi Minh City, Vietnam

d. Vietnam National University, Ho Chi Minh City, Vietnam

e. NTU Institute of Structural Biology, Nanyang Technological University, Singapore

Construction of plasmid

The endonuclease protein with RHAU peptide motif was generated by incorporating RHAU140 (53-192) peptide to Fok1 endonuclease cleavage domain. DNA encoding for RHAU140-Fok1 was amplified by PCR using *RHAU140-Fok1* template (synthesized from IDT company) and a pair of primer ON1/ON2. This PCR product was then cloned into treated pET-Duet1 (Merck) at *BamHI* and *XhoI*, resulting in plasmid pRHAU140-Fok1. (Fig. S1).

The negative control protein was generated by incorporating RHAUΔRSM (RHAU without G4 binding domain) to Fok1 endonuclease cleavage domain. DNA encoding for RHAUΔRSM-Fok1 was amplified by PCR using *RHAU140-Fok1* template (synthesized from IDT company) and a pair of primer ON9/ON2. This PCR product was then cloned into treated pET-Duet1 (Merck) at *BamHI* and *XhoI*, resulting in plasmid pRHAUΔRSM -Fok1.

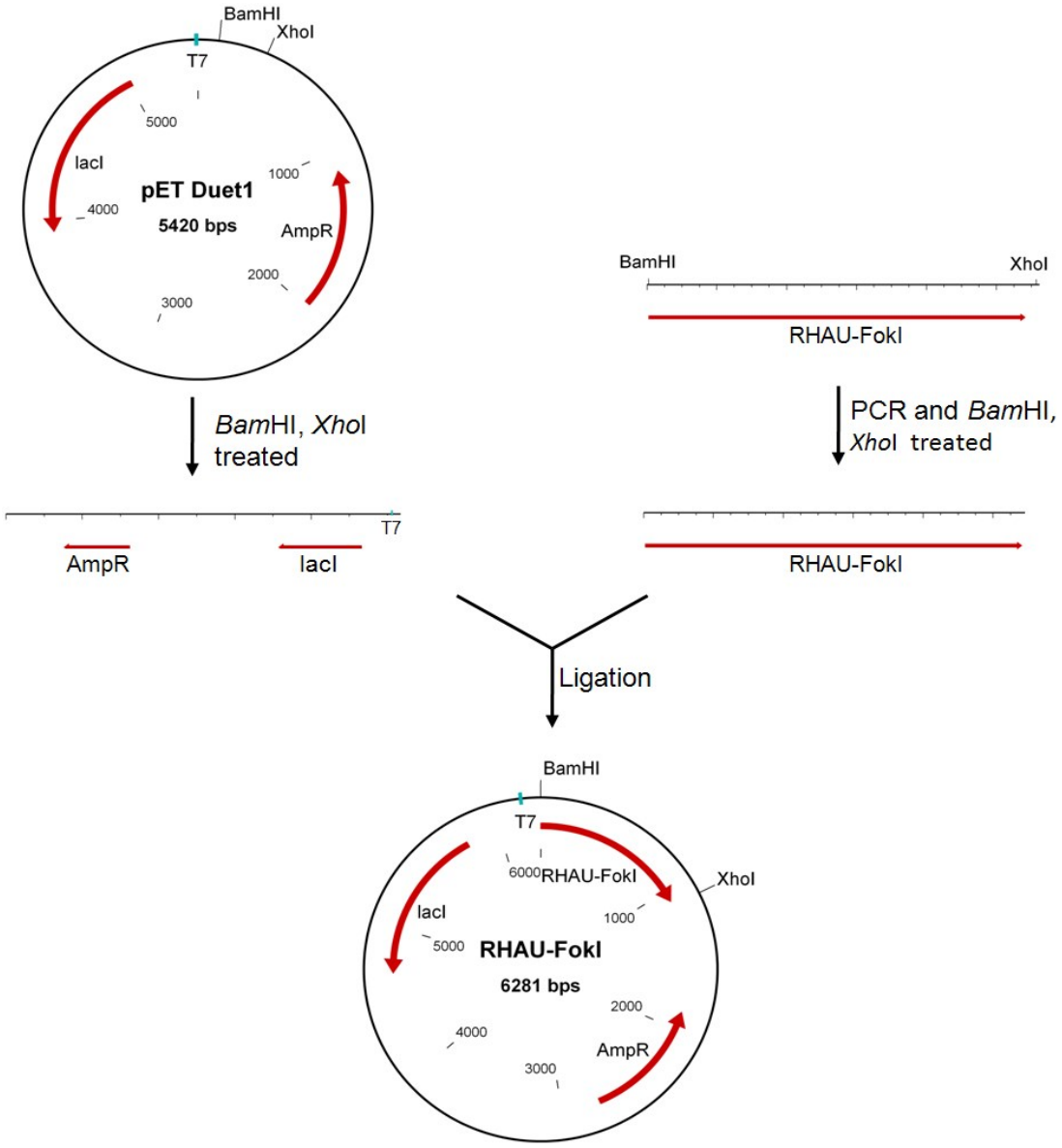

Fig. S1 Construction of plasmid for expression of RHAU140-Fok1.

ON1: 5’- gcgtggatccgtccatgcatcccgggcacctgaaag -3’
ON2: 5’- ccatctcgagttacttgtacagctcgtccatgccgagagtg -3’
ON9: 5’- gtgtggatccgaaacaggggcagaagaacaag-3’

Amino acid sequences of RHAU140-Fok1 (Fok1 sequence is boldfaced, the linker is colored in red):

MGSSHHHHHHSQDPSM*HPGHLKGREIGMWYAKKQGQKNKEAERQERAVVHMDERREEQIVQLLNSVQAKNDKESE AQISWFAPEDHGYGTEVSTKNTPCSENKLDIQEKKLINQEKKMFRIRNRSYIDRDSEYLLQENEPDGTLDQKLLEDLQKKK*S GGGSGGGSGGGSGGG**QLVKSELEEKKSELRHKLKYVPHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYRGKHLGGSR KPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRYVEENQTRNKHINPNEWWKVYPSSVTEFKFLFVSGHF KGNYKAQLTRLNHITNCNGAVLSVEELLIGGEMIKAG TLTLEEVRRKFNNGEINFRS**

Amino acid sequences of RHAUΔRSM-Fok1 (Fok1 sequence is boldfaced, the linker is colored in red):
SDS-PAGE
MGSSHHHHHHSQDP*KQGQKNKEAERQERAVVHMDERREEQIVQLLNSVQAKNDKESEAQISWFAPEDHGYGTEVST KNTPCSENKLDIQEKKLINQEKKMFRIRNRSYIDRDSEYLLQENEPDGTLDQKLLEDLQKKK*SGGGSGGGSGGGSGGG**QL VKSELEEKKSELRHKLKYVPHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYRGKHLGGSRKPDGAIYTVGSPIDYGVIV DTKAYSGGYNLPIGQADEMQRYVEENQTRNKHINPNEWWKVYPSSVTEFKFLFVSGHFKGNYKAQLTRLNHITNCN GAVLSVEELLIGGEMIKAG TLTLEEVRRKFNNGEINFRS**

SDS-PAGE experiments were performed using 12% gel polyacrylamide with 5% of stacking gel (Fig. S2).

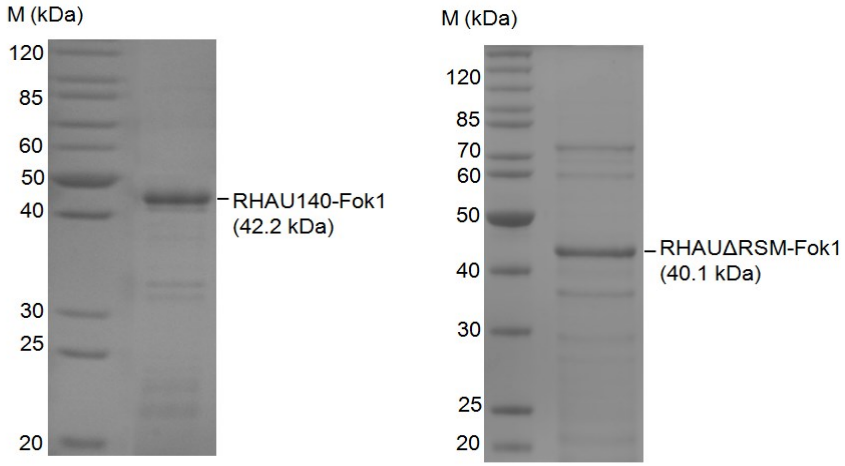

Fig. S2 SDS-PAGE of RHAU140-Fok1 (left) and RHAUΔRSM-Fok1 (right)

Construction of substrates

Myc-dx2

Firstly, a short DNA consisting of *Myc* sequence was generated by hybridization of two primers ON3/ON4. This short DNA was then inserted into the pSmart HC Kan vector by CloneSmart® Blunt Cloning Kits, resulting in pMyc. The *Myc-dx2* substrate was generated by the PCR with a pair of primer ON5 and FAM-labelled ON6 using the pMyc template.

dx2

Firstly, a short DNA consisting of mutated *Myc* sequence was generated by hybridization of two primers ON7/ON8. This short DNA was then inserted into the pSmart HC Kan vector by CloneSmart® Blunt Cloning Kits, resulting in pNon-Myc. The *dx2* substrate was generated by the PCR with a pair of primer ON5 and FAM-labelled ON6 using the pNon-Myc template.

ON3:
5'_caacgtctatatcatggccgacaagcagaactt**ggggagggtggggagggtggggaagg**aatgacgacgtcaactacaagggacgacagc_3'

ON4:
5'_gctgtcgtcccttgtagttgacgtcgtcattccttccccaccctccccaccctccccaagttctgcttgtcggccatgatatagacgttg_3'

ON5: 5'_ gagccagtgagttgattgca_3'

ON6: 5'_**FAM**-ctgctatggaggtcaggtat_3'

ON7:
5'_caacgtctatatcatggccgacaagcagaactt**gaagagcatgtagagcttgtcgaagg**aatgacgacgtcaactacaagggacgacagc_3'

ON8: 5'_gctgtcgtcccttgtagttgacgtcgtcattccttcgacaagctctacatgctcttcaagttctgcttgtcggccatgatatagacgttg_3'

Cleavage of enzyme

G4-dx1 duplex bearing a G-quadruplex motif was cut in dose-dependence of RHAU140-Fok1. Efficiency of DNA cleavage was enhanced in increasing protein concentration (Fig. S3).

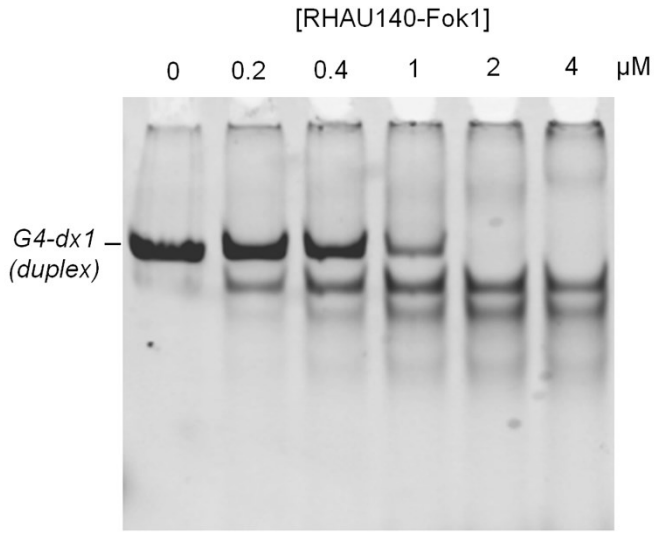

Fig. S3 DNA cleavage as a function of enzyme concentration. Cleavage of *G4-dx1* labeled with FAM at the 3'-end of the top strand (1 μM) bearing a G4 motif by different RHAU140-Fok1 concentrations (0, 0.2, 0.4, 1, 2, 4 μM) were carried out at 37°C in 1x CutSmart® buffer in 30 minutes and analyzed by native gel electrophoresis (18%).

The specific cleavage double-strand *G4-dx1* bearing a G4 motif was not observed when incubation of this substrate with RHAU140ΔRSM-Fok1. RHAUΔRSM-Fok1 (lack of G4 binding domain) did not recognize and bind G4 structure of substrate, therefore, specific cleavage did not occur in reaction (Fig. S4).

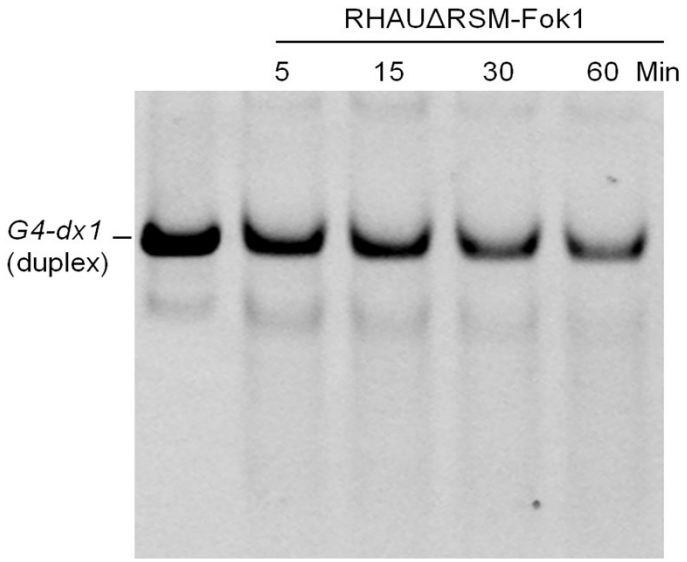

Fig. S4 Time-dependence of DNA cleavage by RHAUΔRSM-Fok1. Cleavage of *G4-dx1* labeled with FAM at the 3'-end of the top strand (1 μM) bearing a G4 motif by RHAUΔRSM-Fok1 (RHAU140 lacking the G4 binding domain) (2 μM) was carried out at 37°C in 1x CutSmart® buffer and analyzed by native gel electrophoresis (18%). Samples were sequentially collected at 5, 15, 30, 60 minutes.

Cleavage of DNA by RHAU140-Fok1 was carried out with single and double DNA (Fig. S5). RHAU140-Fok1 showed high efficiency of specific cleavage with double strand DNA (lane 4 and 6). In contrast, minor cleavage was observed with single DNA under the same condition (lane 2 and 8). Cleavage products of double-strand DNA *G4-dx1* labeling with FAM at the 3'-end of the top strand and at the 5'-end of the bottom strand (C-strand) were observed (lane 4 and 6) that demonstrates RHAU140-Fok1 cleaved both top and bottom strands of DNA duplex.

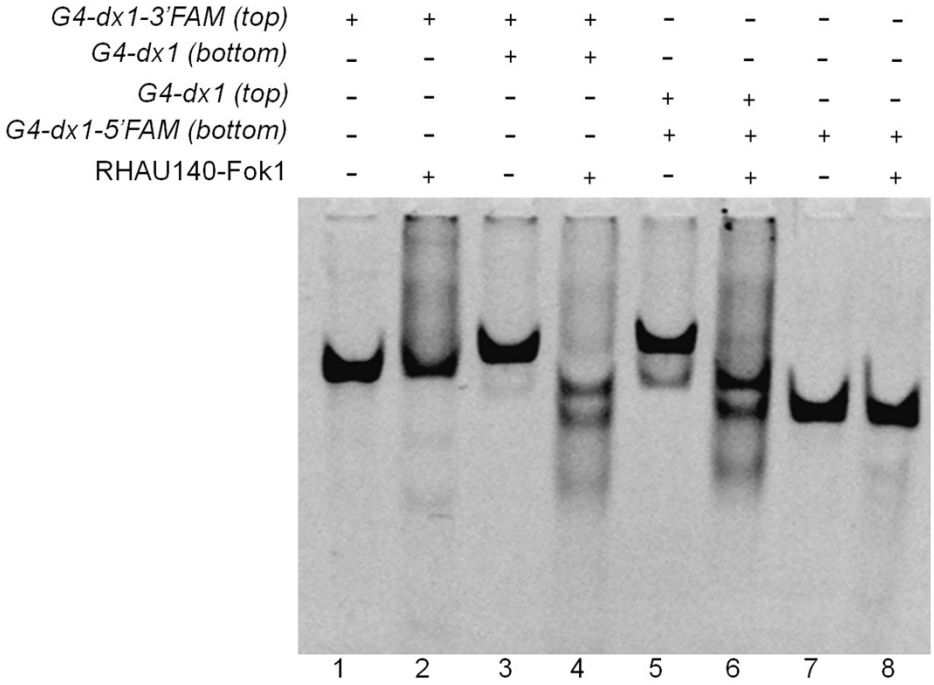

Fig. S5 Cleavage assay by RHAU140-Fok1 on various DNA substrates. Lane 1 and 2, only the top strand of *G4-dx1* labeled with FAM at the 3'-end (1 μM) were incubated without and with RHAU140-Fok1 (2 μM); Lane 3 and 4, G4-contining duplex *G4-dx1* in the presence the bottom strand where the 3'-end of the top strand was labelled by FAM (1 μM) were incubated without and with RHAU140-Fok1 (2 μM); Lane 5 and 6, same as lane 3 and 4, except that the DNA substrate was labeled with FAM at the 5'-end of the bottom strand; Lane 7 and 8, bottom strand labeled with FAM at the 5'-end (1 μM) were incubated without (lane 7) and with RHAU140-Fok1 (2 μM) (lane 8). The samples were carried out at 37 ^{o}C in 1x CutSmart® buffer in 30 minutes and analyzed by native gel electrophoresis (18%).

The *Myc-dx2* was labeled at the 5'-end of the C-rich strand with a FAM fluorescent dye. The control *dx2*, where guanines in the G-tracts were mutated, was also labeled with FAM. After incubation, RHAU140-Fok1 protein did not show any specific cleavage activity to both substrates in 1x CutSmart buffer condition.

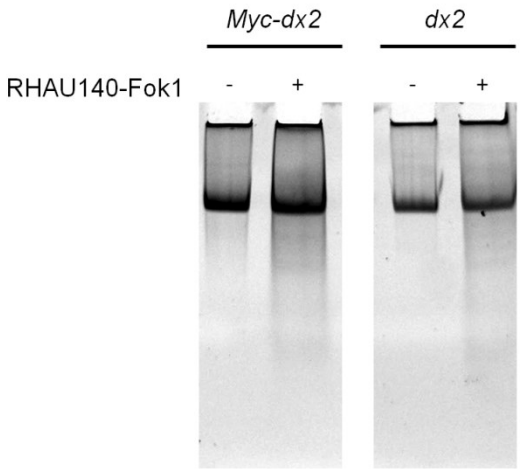

Fig. S6 Cleavage of RHAU140-Fok1 protein (500 nM) for *Myc-dx2* (640 nM) bearing parallel G-quadruplex structure and *dx2* (640 nM) were carried out at 37°C in 1x CutSmart buffer. All the samples were labelled with FAM and analyzed by the native gel electrophoresis (15%).

DNA ladder production

The DNA ladders were made based on the Maxam – Gilbert method.[1] Briefly, 2 nanomoles of FAM-labelled DNA (top strand or bottom strand) were chemically modified in 2 separate reactions: (1) at Adenine and Guanine (A and G) by 50% formic acid, and (2) at Cytosine and Thymine (C and T) by 60% hydrazine, for 10 minutes at room temperature. Subsequently, the modified DNA were precipitated by 70% ethanol, then, the DNA pellets were washed twice by 70% ethanol and air dried before being dissolved in 1 M piperidine. The DNA – Piperidine solutions were incubated at 90°C for 30 mins, by which piperidine specifically cleaves the DNA at the modified bases. 20 picomoles of the resulted DNA ladders were run on denaturing gels along with the reported enzymatic reactions for sequencing analysis.

References

1. A.M. Maxam and W. Gilbert, *Proc Natl Acad Sci USA*, 1977, **74**, 560-564.