

## Supporting Information

### **Intramolecular G-quadruplex-hairpin loop structure competition of a GC-rich exon region in the TMPRSS2 gene**

Wataru Sugimoto,<sup>a</sup> Natsuki Kinoshita,<sup>a</sup> Minori Nakata,<sup>a</sup> Tatsuya Ohyama,<sup>b</sup> Hisae Tateishi-  
Karimata,<sup>b</sup> Takahito Nishikata,<sup>a</sup> Naoki Sugimoto,<sup>a,b</sup> Daisuke Miyoshi\*<sup>a</sup> and Keiko Kawauchi\*<sup>a</sup>

#### **Table of Contents**

Materials and methods

Materials

Fluorescence spectroscopy

Circular dichroism spectroscopy

Transcription assay

Table S1 and S2

Figure S1~ S14

References

## **Materials and methods**

### Materials

All high-pressure liquid chromatography (HPLC)-grade, Oligonucleotide Purification Cartridge (OPC) -grade, desalted-grade DNA strands used in this study were acquired from Eurofins Genomics (Tokyo, Japan) and Thermo Fisher Scientific (Tokyo, Japan) (Table S1). Each oligonucleotide concentration was determined by the absorbance at 260 nm at 95 °C. Extinction coefficients for single-strand DNA were calculated from mono- and dinucleotide data using the nearest-neighbor approximation model [1]. Single strand concentrations of oligonucleotides were measured at 260 nm using a UV-1800 spectrophotometer (Shimadzu Co., Ltd., Kyoto, Japan). Chemical reagents were purchased from Tokyo Chemical Industry Co., Ltd. (Tokyo, Japan) and Wako Pure Chemical Co., Ltd. (Osaka, Japan). All reagents were used without further purification.

### Fluorescence spectroscopy

Fluorescence spectra for 1  $\mu\text{M}$  ThT or 1  $\mu\text{M}$  NMM in the presence of 0, 1.0, 2.5, 3.0, 10 or 30  $\mu\text{M}$  DNA oligonucleotides in 150 mM KCl, 40 mM Tris-HCl (pH 7.2 at 37 °C) and 8 mM  $\text{MgCl}_2$  buffer were recorded on a FP-8200 spectrofluorometer (JASCO Co., Ltd., Tokyo, Japan) equipped a JASCO PTC-348 temperature controller with a 0.3 cm  $\times$  0.3 cm quartz cell at 25 °C. Before the measurement, the samples were heated to 93 °C, and then cooled at a rate of  $-0.5\text{ }^\circ\text{C min}^{-1}$ . Excitation wavelengths for ThT and NMM were 450 nm and 399 nm, respectively.

### Circular dichroism spectroscopy

CD spectra of DNA oligonucleotides were measured using a J-820 spectropolarimeter (JASCO Co., Ltd., Tokyo, Japan) at 20  $\mu\text{M}$  total DNA strand concentration in 150 mM KCl, 40 mM Tris-HCl (pH 7.2 at 37 °C) and 8 mM  $\text{MgCl}_2$  buffer. Before the measurement, the samples were heated to 93 °C, and then cooled at a rate of  $-0.5\text{ }^\circ\text{C min}^{-1}$ . The spectra at 25 °C were obtained using at least three scans between 200 to 350 nm in a cuvette with a path length of 0.1 cm.

For CD melting experiments, samples were heated from 20°C to 90°C at a rate of  $0.5\text{ }^\circ\text{C min}^{-1}$ . Before the measurements, DNA samples were heated to 93 °C, cooled to 25 °C at a rate of  $-0.5\text{ }^\circ\text{C min}^{-1}$ . Then, thermal denaturing behavior was traced at 265 nm and 290 nm. The temperature of the cell holder was regulated by a JASCO PTC-348 temperature controller, and the cuvette-holding chamber and the light source were flushed with a constant stream of dry  $\text{N}_2$  gas to avoid condensation of water on the cuvette exterior and generation of  $\text{O}_3$  gas.

### Transcription Assays

Each transcription reaction solution contained a 2.5  $\mu\text{M}$  DNA template, 150 mM KCl or LiCl, 40 mM Tris-HCl (pH 7.2 at 37 °C), 8 mM  $\text{MgCl}_2$ , with or without 10% PEG200. The samples were heated to 93 °C for 5 min, and then cooled  $-0.5\text{ }^\circ\text{C min}^{-1}$ . After annealing, 100 units of T7 polymerase was added to final concentrations of NTPs and DTT of 1 mM and 5 mM, respectively.

Each mixture was incubated at 37 °C for 0, 10, 30, 60, 90 or 120 min. Reactions were quenched by incubation with 10 units of DNaseI for 20 min, and then a 10-fold excess volume of transcription stop solution (80 wt% formamide, 10 mM  $\text{Na}_2\text{EDTA}$ , and 0.01% blue dextran) was added. Each sample was cooled rapidly after heating to 93 °C for 5 min, and then loaded onto a 12% polyacrylamide, 7 M urea gel run at 70 °C.

After electrophoresis, each gel was stained with SYBR Gold (PerkinElmer Life Sciences, Waltham, MA, U.S.A.) and fluorescence was detected with a fluorescence imager (GE Healthcare, FLA-5100, Little Chalfont, U.K.). Band intensities were quantified using ImageJ software distributed by the National Institutes of Health, USA.

Table S1

Table S1. DNA sequences used in this study

<b>Oligonucleotide</b>	<b>Sequence (5' → 3')</b>
WT	CCTCCGGGCGGGGCAGGGGGCATCGGCGGGTCCCAGGCGCCCAGGTTCCCC TCCCCAGCCCGGACCCCGAGCCGGGACCC
MT1	CCTCCGGGCGGGGCAGGGGGCATCGGCGGGTCTCAGGCGCTCAGGTTCTCC TCTCCAGCTCGGACTCTGAGCCGGGACTC
MT2	CCTCCGAGCGGAGCAGAAGGCATCGGCGAGTCTCAGGCGCTCAGGTTCTCC TCTCCAGCTCGGACTCTGAGCCGGGACTC
MT3	CCTCCAGACGCGCCAGCGCGCATCAGCGGGTCCCAGGCGCTCAGGTTGCGC TGCGCAGTCTGGACCCCGAGCCGGGACCC
MT4	CCTCGGCGCGCGCCAGCGCGCATCGCGCGGTCCCAGGGCGCCAGGTTGCGC TGCGCAGCGCCGACCCCGAGCCGGGACCC
Temp-WT (template)	CCTCCGGGCGGGGCAGGGGGCATCGGCGGGTCCCAGGCGCCCAGGTTCCCC TCCCCAGCCCGGACCCCGAGCCGGGACCCAGAGAGAGCACCCGAGCCTAGT TCGTGTCATCTCCTATAGTGAGTCGTATTATATAGTGAGTCGTATTAGTG
Temp-MT1 (template)	CCTCCGGGCGGGGCAGGGGGCATCGGCGGGTCTCAGGCGCTCAGGTTCTCC TCTCCAGCTCGGACTCTGAGCCGGGACTCCAGAGAGAGCACCCGAGCCTAGT TCGTGTCATCTCCTATAGTGAGTCGTATTATATAGTGAGTCGTATTAGTG
Temp-MT2 (template)	CCTCCGAGCGGAGCAGAAGGCATCGGCGAGTCTCAGGCGCTCAGGTTCTCC TCTCCAGCTCGGACTCTGAGCCGGGACTCCAGAGAGAGCACCCGAGCCTAGT TCGTGTCATCTCCTATAGTGAGTCGTATTATATAGTGAGTCGTATTAGTG
Temp-MT3 (template)	CCTCCAGACGCGCCAGCGCGCATCAGCGGGTCCCAGGCGCTCAGGTTGCGC TGCGCAGTCTGGACCCCGAGCCGGGACCCAGAGAGAGCACCCGAGCCTAGT TCGTGTCATCTCCTATAGTGAGTCGTATTATATAGTGAGTCGTATTAGTG
Temp-MT4 (template)	CCTCGGCGCGCGCCAGCGCGCATCGCGCGGTCCCAGGGCGCCAGGTTGCGC TGCGCAGCGCCGACCCCGAGCCGGGACCCAGAGAGAGCACCCGAGCCTAGT TCGTGTCATCTCCTATAGTGAGTCGTATTATATAGTGAGTCGTATTAGTG
85WT (template)	TCCCAGGCGCCCAGGTTCCCCCTCCCCAGCCCGGACCCCGAGCCGGGACCC AGAGAGAGCACCCGAGCCTAGTTCGTGTCATCTCCTATAGTGAGTCGTATTA TATAGTGAGTCGTATTAGTG
85MT (template)	TCTCAGGCGCTCAGGTTCTCCTCTCCAGCTCGGACTCTGAGCCGGGACTCC AGAGAGAGCACCCGAGCCTAGTTCGTGTCATCTCCTATAGTGAGTCGTATTA TATAGTGAGTCGTATTAGTG

95WT (template)	CATCGGCGGGTCCCAGGCGCCAGGTTCCCCTCCCAGCCCGGACCCCGAG CCGGGACCCCAGAGAGAGACCCGAGCCTAGTTCGTGTCATCTCCTATAGTG AGTCGTATTATATAGTGA-GTCGTATTAGTG
95MT (template)	CATCGGCGAGTCTCAGGCGCTCAGGTTCTCCTCTCCAGCTCGGACTCTGAG CCGGGACTCCAGAGAGAGACCCGAGCCTAGTTCGTGTCATCTCCTATAGTG AGTCGTATTATATAGTGA-GTCGTATTAGTG
T7 promoter (non-template)	CACTAATACGACTCACTATATAATACGACTCACTATAGG

---

**Table S2**

Table S2. Human protease genes having G4-forming sequence

<b>Gene</b>	<b>G4-forming sequence (5' → 3')<sup>1</sup></b>	<b>Up-stream</b>	<b><math>\Delta G^{0}_{25^4}</math></b>	<b>Down-stream</b>	<b><math>\Delta G^{0}_{25^4}</math></b>
<i>CAPN2</i>	GGGCCGAGGGGATGGCCGGGAAGGACGGG	20	-17.1	17	-14.8
<i>CAPN9</i>	GGGTGATCCGGGCGTCTTGGGAACCGGG	14	-13.7	16	-14.2
<i>CAPN10</i>	GGGAGGGATCCTGGGAAGCTCTGGGATCCTGGGG	20	-15.0	14	-16.3
<i>CTSA</i>	GGGATGTACAGGGCTCTGGGAAGGGGG	14	-9.6	16	-14.9
<i>CTSD</i>	GGGTGTGTGTGGGAGGGCCGCTGGGCCAGGGG	19	-16.8	14	-15.8
<i>Furin</i> <sup>2</sup>	GGGGGCATGGGGCCGAGACGGGCAGGGCTCGGGAGGG	6	-11.4	19	-18.5
<i>Furin</i> <sup>2</sup>	GGGACTTGGGCCCTGACGGGAGAGGCGGG	10	-11.9	17	-12.5
<i>Furin</i> <sup>2</sup>	GGGCCAGGGCAGCCAGGAGGAGCAGGG	12	-10.4	23	-13.0
<i>Furin</i> <sup>2</sup>	GGGCAGGGGGATGGGAGAGGAGGGGCCACCCCTGGG	7	-13.0	19	-17.4
<i>MMP3</i>	GGGTACCAGGGGGGTCTCAGGGGAGTCAGGGGG	7	-9.3	12	-14.6
<i>MMP19</i> <sup>2</sup>	GGGCTGCCCGGGTGTGTGGGGTGAAGGGTGGAGGG	12	-9.6	12	-11.6
<i>MMP19</i> <sup>2</sup>	GGGGCCATGAGGGCCTGGGAATATCGGG	17	-18.4	14	-12.4
<i>MMP19</i> <sup>2</sup>	GGGACTGGGTTCGTGGGCACTGGGGGCACAGTGGG	19	-6.0	17	-15.8
<i>TPSB2</i> <sup>2</sup>	GGGACGGGGGGACCGGGGTGGG	14	-16.6	8	-7.9
<i>TPSB2</i> <sup>2</sup>	GGGCCTGGCCGGGCTCACAGGGCAGGGCTGGGGG	6	-16.1	12	-16.5
<i>TPSB2</i> <sup>2</sup>	GGGAGTGGGGTTGGGGGGCGGGGGCGGGGGACAGG CGGGG	20	-14.5	16	-12.4
<i>TPSB2</i> <sup>2</sup>	GGGCAGGGGAGGGCCGGAGGG	10	-8.9	16	-10.2
<i>ADAM12</i>	GGGCGCGGGCGGGGACCGGGCAGCGGG	21	-25.1	17	-19.3
<i>MMP15</i> <sup>2</sup>	GGGGGTGAAGGGCGGCCGGGCCACGTCGGGCCATCG GGGG	18	-25.0	27	-22.1
<i>MMP15</i> <sup>2</sup>	GGGCGGGGCAGACAGGGTGTGTGAGGGCTGAGAGGGG	18	-17.9	9	-7.9
<i>MMP15</i> <sup>2</sup>	GGGCCCAAGGGGAGGTGGGAGCTGGGGG	13	-15.2	12	-12.8
<i>MMP17</i>	GGGGGCTCTGGGCGGGTAGCCGGGG	14	-9.0	19	-18.0
<i>MMP25</i> <sup>2</sup>	GGGCCTGGGCAGGGTGGGG	12	-8.1	16	-7.1
<i>MMP25</i> <sup>2</sup>	GGGGGTGCTTTGGGGGGCCTGGGGGCGCGGGG	15	-15.1	16	-11.4
<i>MMP25</i> <sup>2</sup>	GGGAAGGGTTCCTGGGGATAGAAGGG	12	-12.1	11	-6.9
<i>TMPRSS2</i> <sup>2</sup>	GGGCGGGGCAGGGGGCATCGGCCGGG	27	-19.0	27	-17.0
<i>TMPRSS2</i> <sup>2</sup>	GGGGCGTACTGGGGCACGGGGGACGGG	12	-11.0	13	-13.8
<i>TMPRSS5</i> <sup>2</sup>	GGGGCAGGGAGGGGAGAGGGCAGAAGGG	21	-16.3	18	-12.3
<i>TMPRSS6</i> <sup>2</sup>	GGGCAGAGGGGGCGGGATCGGGG	11	-10.5	9	-8.9

<i>TMPRSS6<sup>2</sup></i>	GGGGAAGGGGAGAGGCTGGGTAGGG	12	-12.0	7	-4.0
<i>TMPRSS6<sup>2</sup></i>	GGGAATCAGGGCATGGGACGGG	5	-4.2	9	-6.1
<i>TMPRSS6<sup>2</sup></i>	GGGCCCCAGGGGAGGGGGTATATGGGGAGGG	16	-18.1	17	-18.8
<i>TMPRSS6<sup>2</sup></i>	GGGCTGGGCAGGGGAGCCGGGTGGGG	20	-13.0	16	-17.8
<i>TMPRSS6<sup>2</sup></i>	GGGCTGGGGTCTGGGAGGGGAGTGGCAGGGAGGG	9	-9.3	15	-13.2
<i>TMPRSS6<sup>2</sup></i>	GGGTGCCAAGGGGAGAGGGCAGGGAGGGGGTGGGGG CAGAGGGGAGGGG	9	-9.0	18	-12.8
<i>TMPRSS6<sup>2</sup></i>	GGGCAGGTGGGCAGGCAGGGTGGGG	9	-9.6	18	-9.3
<i>TMPRSS6<sup>2</sup></i>	GGGCACCTGGGAGGGAGGAGCGGG	16	-19.8	15	-9.39
<i>TMPRSS6<sup>2</sup></i>	GGGAAGAGAGGGAGGGGGAGGG	7	-8.6	9	-4.1
<i>TMPRSS6<sup>2</sup></i>	GGGATGGGCAGGGAGAGGGATGCGGGGCGGG	18	-13.0	8	-9.7
<i>TMPRSS6<sup>2</sup></i>	GGGTGGATGGGTGGGTGGGTAGATGGGTGGG	5	-2.4	2	0.3
<i>TMPRSS6<sup>2</sup></i>	GGGACAGAGGGGATGGGGCAGGG	13	-10.2	19	-12.0
<i>TMPRSS6<sup>2</sup></i>	GGGTCCTAGTGGGGAGCGGGTGGG	9	-11.7	14	-9.7
<i>TMPRSS6<sup>2</sup></i>	GGGGTGGGGTGGGGTGGGGTGGGGTGGGG	14	-12.6	19	-14.5
<i>TMPRSS6<sup>2</sup></i>	GGGCTGGGTGTGGGCCTGGGTCCTCAGGGG	22	-12.2	4	-5.7
<i>TMPRSS9</i>	GGGGGAGACATGGGAAGGGGCTGGG	14	-9.5	13	-13.1
<i>ADAM9<sup>3</sup></i>					
<i>ADAM10<sup>3</sup></i>					
<i>ADAM17<sup>3</sup></i>					
<i>CAPN1<sup>3</sup></i>					
<i>CAPN3<sup>3</sup></i>					
<i>CAPN6<sup>3</sup></i>					
<i>CAPN7<sup>3</sup></i>					
<i>CAPN11<sup>3</sup></i>					
<i>CAPN12<sup>3</sup></i>					
<i>CASP1<sup>3</sup></i>					
<i>CASP2<sup>3</sup></i>					
<i>CASP3<sup>3</sup></i>					
<i>CASP4<sup>3</sup></i>					
<i>CASP6<sup>3</sup></i>					
<i>CASP7<sup>3</sup></i>					
<i>CASP8<sup>3</sup></i>					
<i>CASP9<sup>3</sup></i>					
<i>CASP10<sup>3</sup></i>					

<i>CASP12</i> <sup>3</sup>					
<i>CELA3A</i> <sup>3</sup>					
<i>CELA3B</i> <sup>3</sup>					
<i>CTRC</i> <sup>3</sup>					
<i>CTSB</i> <sup>3</sup>					
<i>CTSK</i> <sup>3</sup>					
<i>CTSV</i> <sup>3</sup>					
<i>MMP1</i> <sup>3</sup>					
<i>MMP7</i> <sup>3</sup>					
<i>MMP8</i> <sup>3</sup>					
<i>MMP10</i> <sup>3</sup>					
<i>MMP12</i> <sup>3</sup>					
<i>MMP13</i> <sup>3</sup>					
<i>MMP16</i> <sup>3</sup>					
<i>MMP20</i> <sup>3</sup>					
<i>MMP24</i> <sup>3</sup>					
<i>PLG</i> <sup>3</sup>					
<i>PRSS1</i> <sup>3</sup>					
<i>TMPRSS1</i> <sup>3</sup>					
<i>TMPRSS3</i> <sup>3</sup>					
<i>TMPRSS4</i> <sup>3</sup>					
<i>TMPRSS7</i> <sup>3</sup>					
<i>TMPRSS11D</i> <sup>3</sup>					
<i>TMPRSS12</i> <sup>3</sup>					
<i>TMPRSS13</i> <sup>3</sup>					

<sup>1</sup>G4 forming G-rich sequence ( $G_{<2}X_{1-7}G_{<2}X_{1-7}G_{<2}X_{1-7}G_{<2}$ ) identified by our sequence analysis for 60 human protease genes.

<sup>2</sup>These genes have multiple G4-forming sequences.

<sup>3</sup>These genes do not have the G4-forming sequence.

<sup>4</sup>Values of  $\Delta G^0_{25}$  (kcal/mol) evaluated by m-fold (2) for the hairpin loop structure of oligonucleotides involving the G4-forming sequence and the upstream or the downstream sequence with 50 nucleotide long.

Figure S1

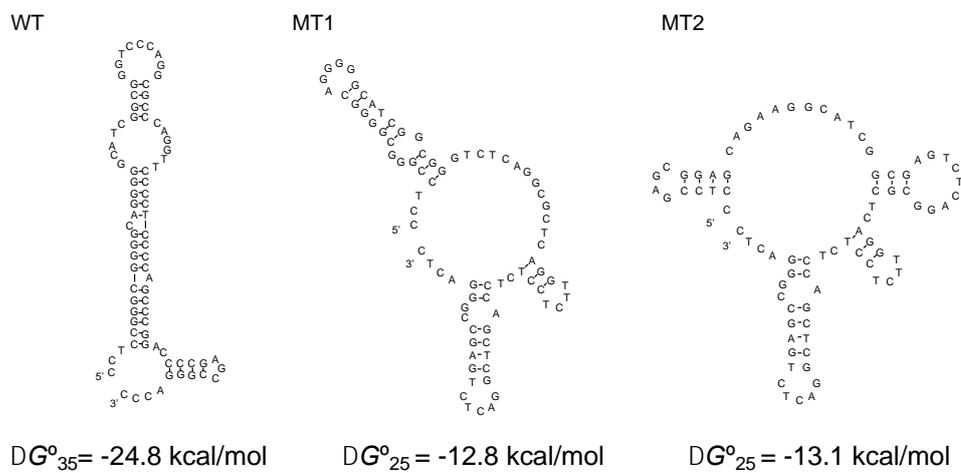


Fig S1. Possible structures and their thermodynamic stability at 25 °C for WT, MT1 and MT2. These structures were predicted using m-fold (2).

Figure S2

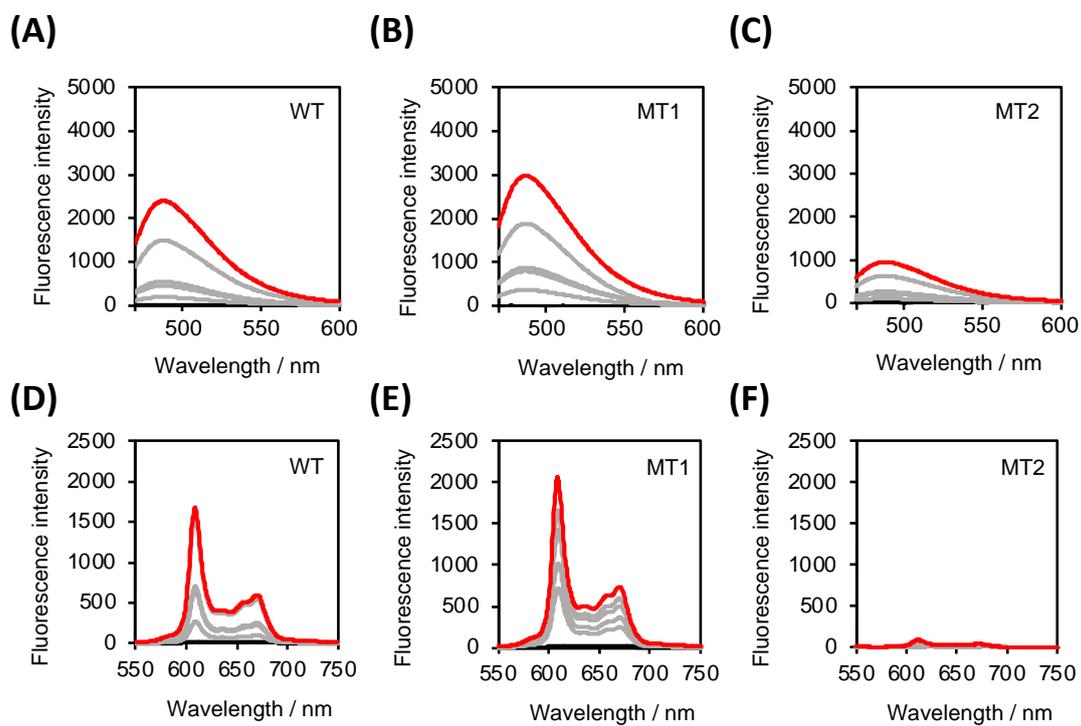


Fig. S2 Fluorescence spectra for 1  $\mu$ M ThT (A-C), or 1  $\mu$ M NMM (D-F) in the presence of 0-30  $\mu$ M WT, MT1, MT2 DNA oligonucleotides in buffer containing 150 mM KCl, 40 mM Tris-HCl (pH 7.2 at 37°C), and 8 mM MgCl<sub>2</sub>. ThT: Ex = 450 nm, Em = 470-600 nm. NMM: Ex = 399 nm, Em = 550-750 nm.

Figure S3

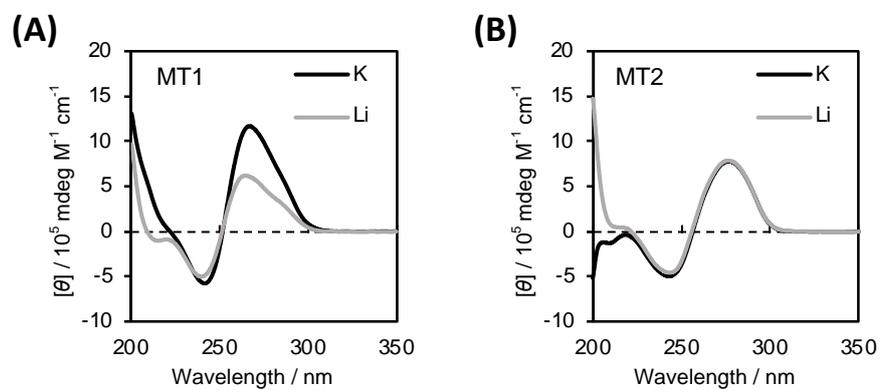


Fig. S3 CD spectra of 20  $\mu\text{M}$  MT1 (A) and MT2 (B) in buffers containing 150 mM KCl (black line) or 150 mM LiCl (gray line), 40 mM Tris-HCl (pH 7.2), and 8 mM  $\text{MgCl}_2$  at 25  $^\circ\text{C}$ .

Figure S4

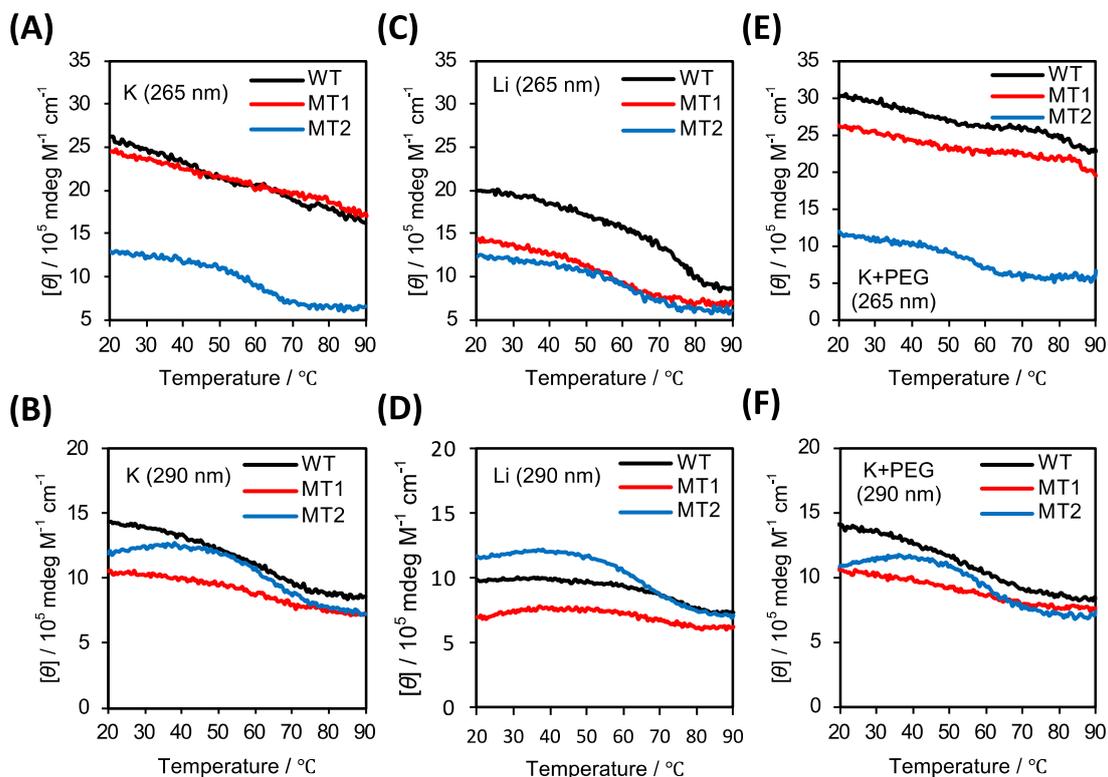


Fig. S4 Thermal denaturing profile traced at 265 nm (A) (C) (E) and 290 nm (B) (D) (F) of 20  $\mu\text{M}$  WT (black) or MT1 (red) or MT2 (blue) in buffers containing 150 mM KCl (A) (B) or 150 mM LiCl (C) (D) or 150 mM KCl with 10% PEG200 (E) (F), 40 mM Tris-HCl (pH 7.2), and 8 mM  $\text{MgCl}_2$ .

In the presence of  $\text{K}^+$  at 265 nm, WT did not show denaturation behavior even at 90  $^{\circ}\text{C}$ , indicating that G4 is highly stable. On the other hand, a clear denaturation behavior was observed at 290 nm, showing WT forms two structural components, which could be G4 and hairpin loop with high and modest thermal stabilities, respectively. A similar trend was observed for MT1, while CD intensity at 290 nm was smaller than WT, indicating that lower ratio of G4 in MT1. These results are consistent with the results of the ThT fluorescence intensity.

In the presence of  $\text{Li}^+$  at 265 nm, WT and MT1 showed denaturation behaviors, consistent with that lower stability of G4 in the presence of  $\text{Li}^+$  than  $\text{K}^+$ , indicating that the denaturation curve of WT and MT1 traced at 265 nm reflects thermal denaturation of G4, and thus that the curve traced at 290 nm reflects thermal denaturation of hairpin loop. Denaturation curves of MT2 at both 265 nm and 295 nm showed clear melting behaviors both in the presence of  $\text{K}^+$  and in the presence of  $\text{Li}^+$ , confirming that stability of a structure of MT2, which is hairpin loop, is independent of cation species.

In the presence of  $\text{K}^+$  with PEG200 at 265 nm and 290 nm, WT and MT1 showed similar denaturation

behaviors with ones without PEG200, although CD intensities at 295 nm were higher, showing that G4 was stabilized by molecular crowding with PEG200. MT2 showed similar denaturation curves at both wavelengths with and without PEG200, although the melting behavior with PEG200 was observed at slightly lower temperature than that without PEG200. These results are consistent with the stabilization and destabilization effects of molecular crowding on G4 and hairpin loop, respectively. Therefore, it is reasonable to consider that WT forms G4 and hairpin loop, depending on the surrounding condition, whereas the dominant structures of MT1 and MT2 are G4 and hairpin loop, respectively, almost independent of the surrounding condition. These CD melting profiles of WT, MT1, MT2 were consistent with the results of ThT fluorescence.

Figure S5

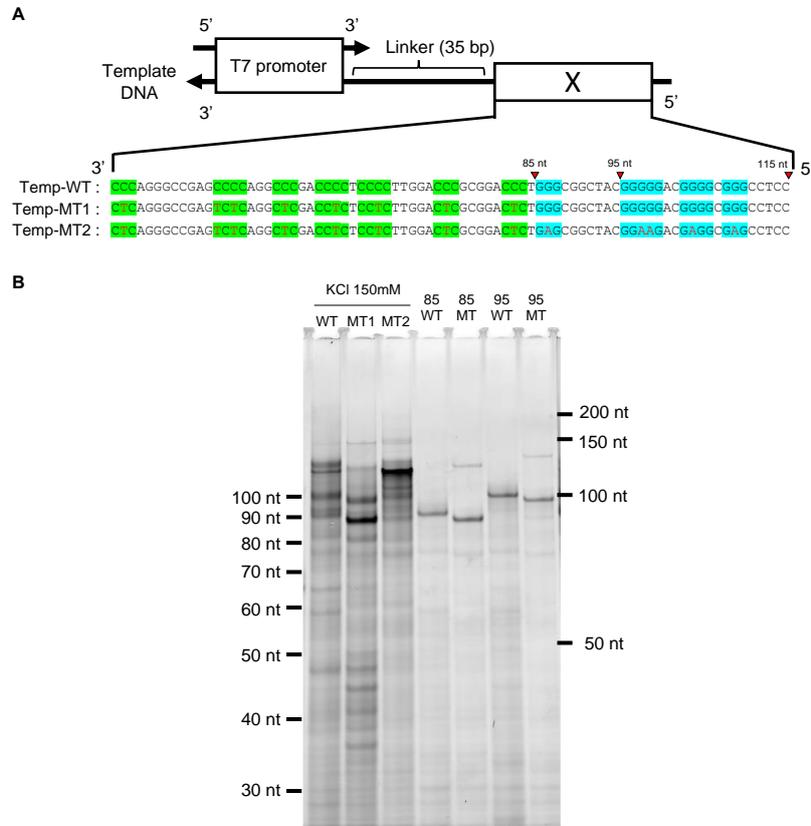


Fig. S5 (A) Design of the template strand for the T7 transcription experiments. Temp-WT, Temp-MT1 and Temp-MT2 contain WT, MT1 and MT2, respectively. The T7 RNA polymerase-binding site is separated from the GC-rich region by a linker sequence with a length of 35 nucleotides. The cytosine- and guanine-stretches are shown in green and blue, respectively. The 85th, 95th, and 115th nucleotides are highlighted by red arrows. The total length of all templates is 152 nucleotides. (B) Transcripts from Temp-WT (lane 1), Temp-MT1 (lane 2), Temp-MT2 (lane 3), and truncated template strands (lanes 4 ~ 7).

Lanes 4 and 5: Transcripts from templates that have the same sequence as Temp-WT (lane 4) or Temp-MT1 (lane 5) but are truncated after the 85th nucleotide. The main bands in lane 4 and lane 5 show the same migration as the third main bands observed with Temp-WT (lane 1) and Temp-MT1 (lane 2), respectively. These results demonstrate that the third bands are products arrested just before the first guanine stretch in Temp-WT and Temp-MT1.

Lanes 6 and 7: Transcripts from templates which have the same sequence as Temp-WT and Temp-MT1, respectively, but are truncated after the 95th nucleotide. The main bands in lane 6 and lane 7 show the same migration as the second main bands observed with Temp-WT (lane 1) and Temp-MT1 (lane 2), respectively. These results demonstrate that the second bands are products arrested just before the first guanine stretch in Temp-WT and Temp-MT1.

Figure S6

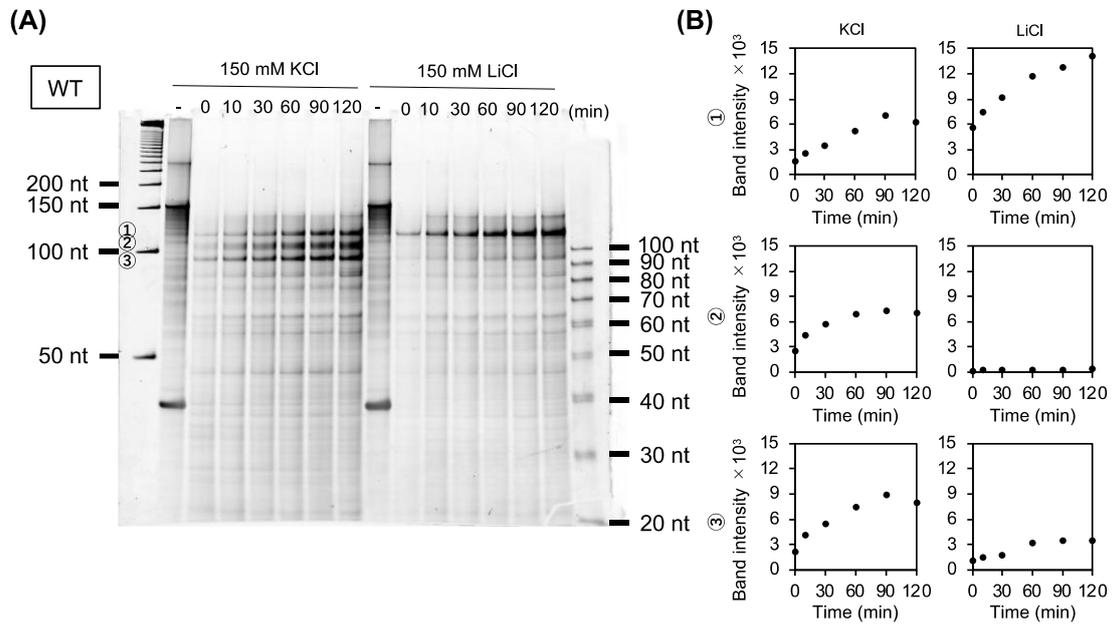


Fig. S6 (A) Transcripts from Temp-WT after a reaction time of 10, 30, 60, 90 or 120 min at 37 °C. The transcription reactions were carried out in the presence of  $K^+$  (left) or  $Li^+$  (right). (B) Intensities of the three main bands after various reaction times in the presence of  $K^+$  (left) or  $Li^+$  (right).

Figure S7

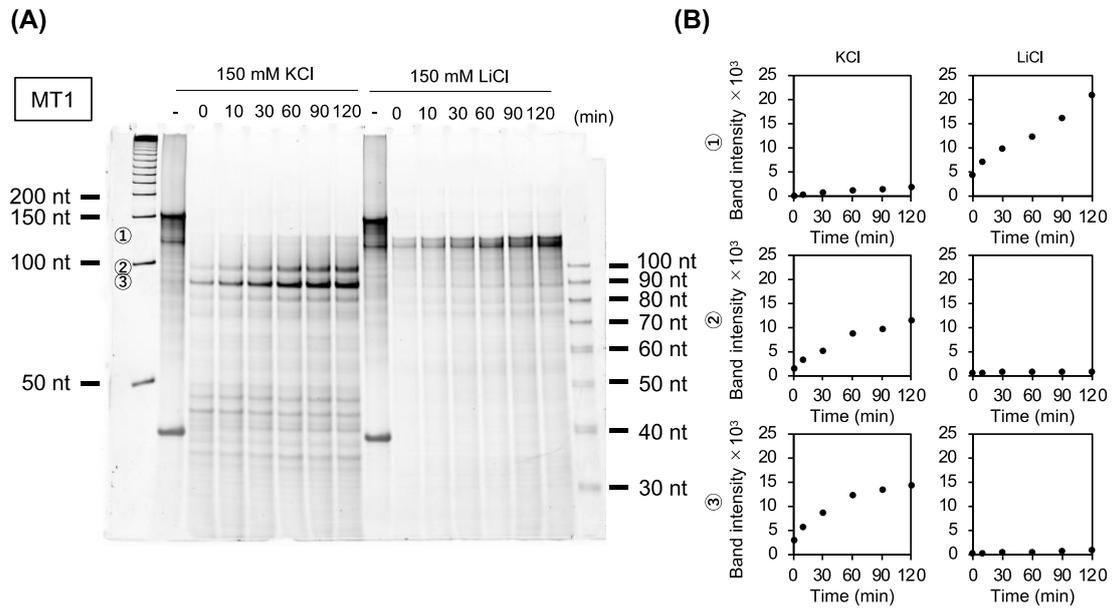


Fig. S7 (A) Transcripts from Temp-MT1 after a reaction time of 10, 30, 60, 90 or 120 min at 37 °C. The transcription reactions were carried out in the presence of  $K^+$  (left) or  $Li^+$  (right). (B) Intensities of the three main bands after various reaction times in the presence of  $K^+$  (left) or  $Li^+$  (right).

Figure S8

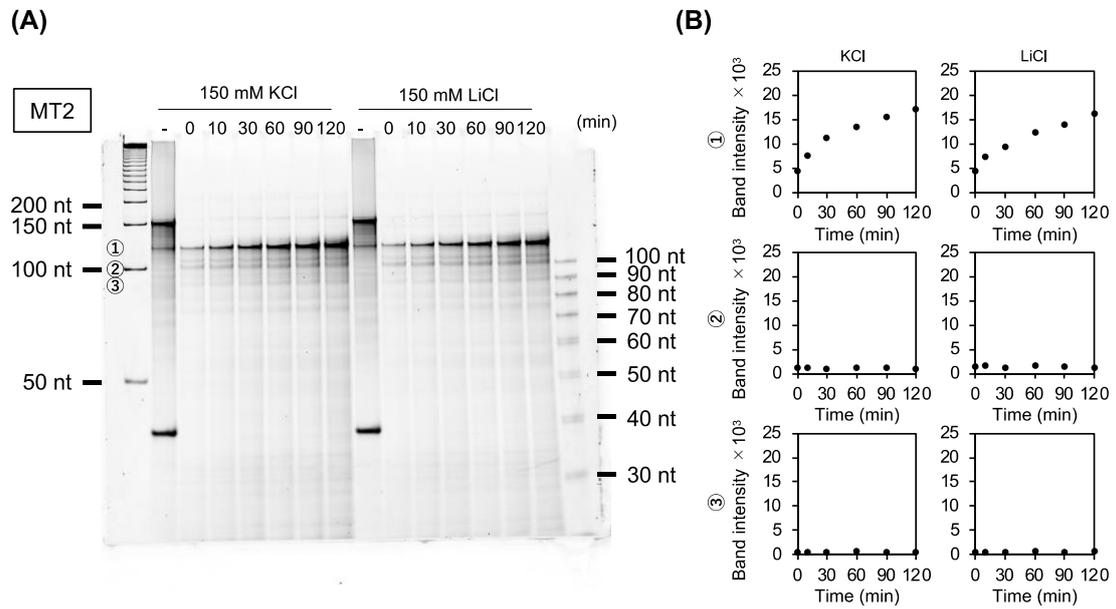


Fig. S8 (A) Transcripts from Temp-MT2 after a reaction time of 10, 30, 60, 90 or 120 min at 37 °C. The transcription reactions were carried out in the presence of  $K^+$  (left) or  $Li^+$  (right). (B) Intensities of the three main bands after various reaction times in the presence of  $K^+$  (left) or  $Li^+$  (right).

Figure S9

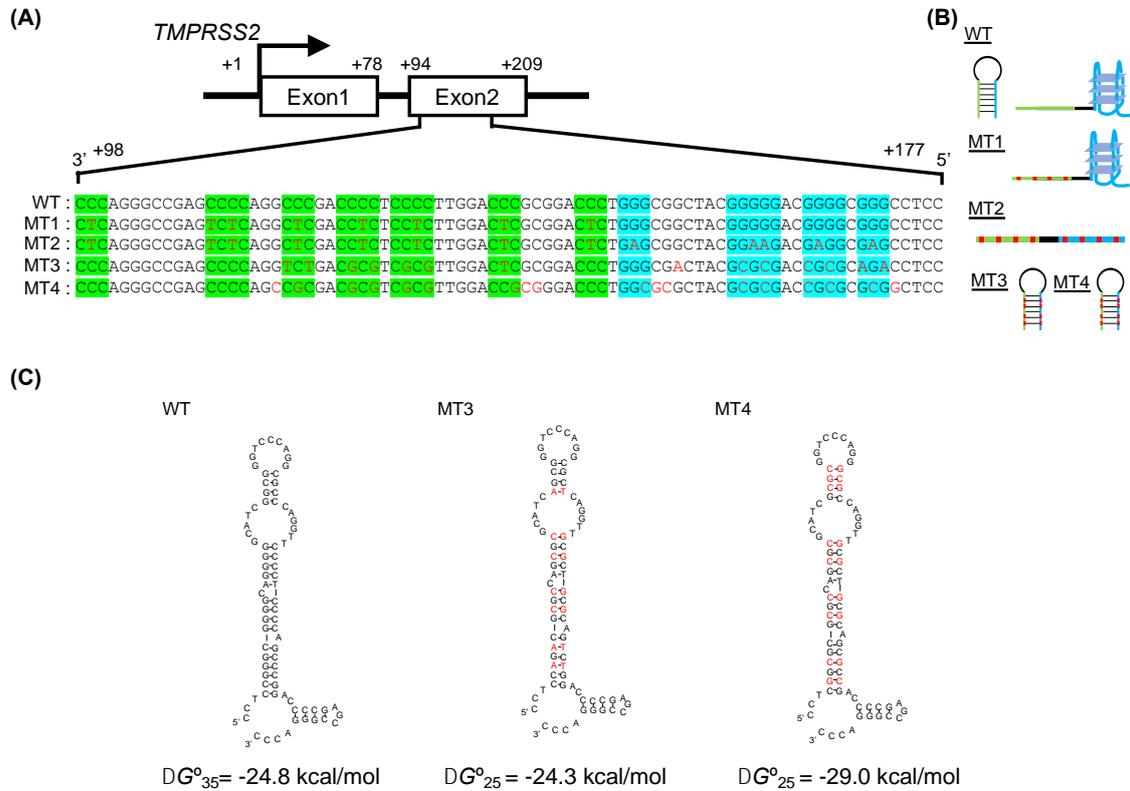


Fig. S9 (A) Schematic diagram showing the location of the GC-rich region in the *TMPRSS2* gene. The guanine and cytosine stretches are highlighted in blue and green, respectively. WT is a wild-type sequence. MT1 has the the C-to-T mutations. MT2 has the C-to-T and the G-to-A mutations. MT3 has the same secondary structure and thermodynamic stability ( $\Delta G_{37}^{\circ}$ ) with WT. MT4 is designed as MT3 but with higher thermodynamic stability than WT. Note that MT3 and MT4 do not have any guanine stretch, and thus cannot form G4. The mutation sites are indicated by the red letter. Sequence of WT, MT1 and MT2 are shown for comparison. (B) Possible structure of WT, MT1, MT2, MT3 and MT4. (C) Possible structures and their thermodynamic stability at 25 °C for WT, MT3 and MT4. These structures were predicted using m-fold<sup>2</sup>. The mutation sites are indicated by the red letter. WT is shown for comparison.

Figure S10

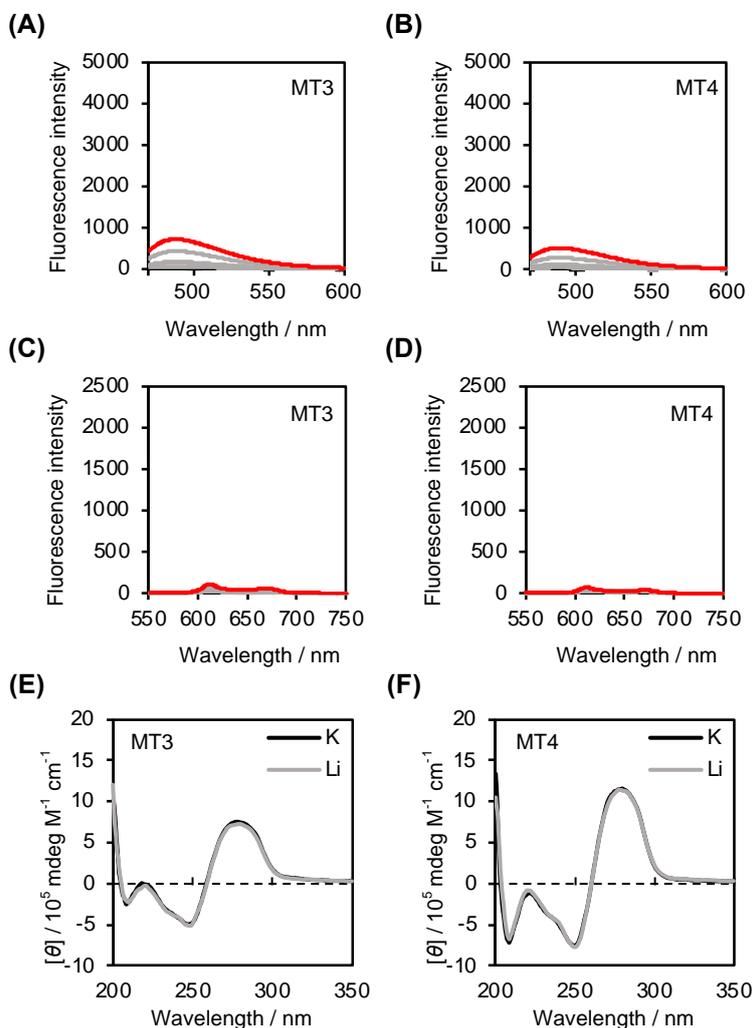


Fig. S10 Fluorescence spectra for 1  $\mu\text{M}$  ThT (A, B), or 1  $\mu\text{M}$  NMM (C, D) in the presence of 0-30  $\mu\text{M}$  MT3, MT4 DNA oligonucleotides in the buffer containing 150 mM KCl, 40 mM Tris-HCl (pH 7.2), and 8 mM  $\text{MgCl}_2$ . ThT:  $E_x = 450 \text{ nm}$ ,  $E_m = 470\text{-}600 \text{ nm}$ , NMM:  $E_x = 399 \text{ nm}$ ,  $E_m = 550\text{-}750 \text{ nm}$ . CD spectra of 20  $\mu\text{M}$  MT3 or MT4 in the buffers containing 150 mM KCl (black line) or 150 mM LiCl (gray line), 40 mM Tris-HCl (pH 7.2), and 8 mM  $\text{MgCl}_2$  (E, F). All measurements were carried out at 37°C.

Figure S11

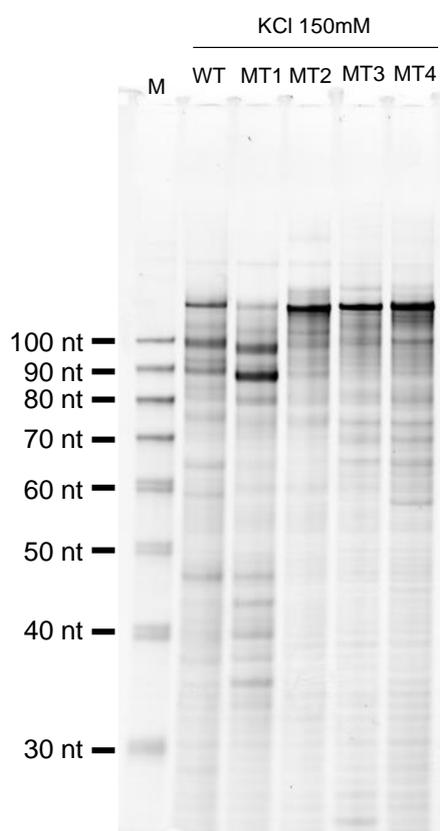


Fig. S11 Transcript production from Temp-WT, Temp-MT1, Temp-MT2, Temp-MT3, Temp-MT4 and transcriptional length marker (M). Reaction mixtures contains 100 units T7 polymerase, 2.5  $\mu$ M DNA template in the buffer of 150 mM KCl, 40 mM Tris-HCl (pH 7.2), 8 mM MgCl<sub>2</sub>. Denaturing gel electrophoresis were carried out at 70°C of transcription reaction products after 120 min at 37°C.

Figure S12

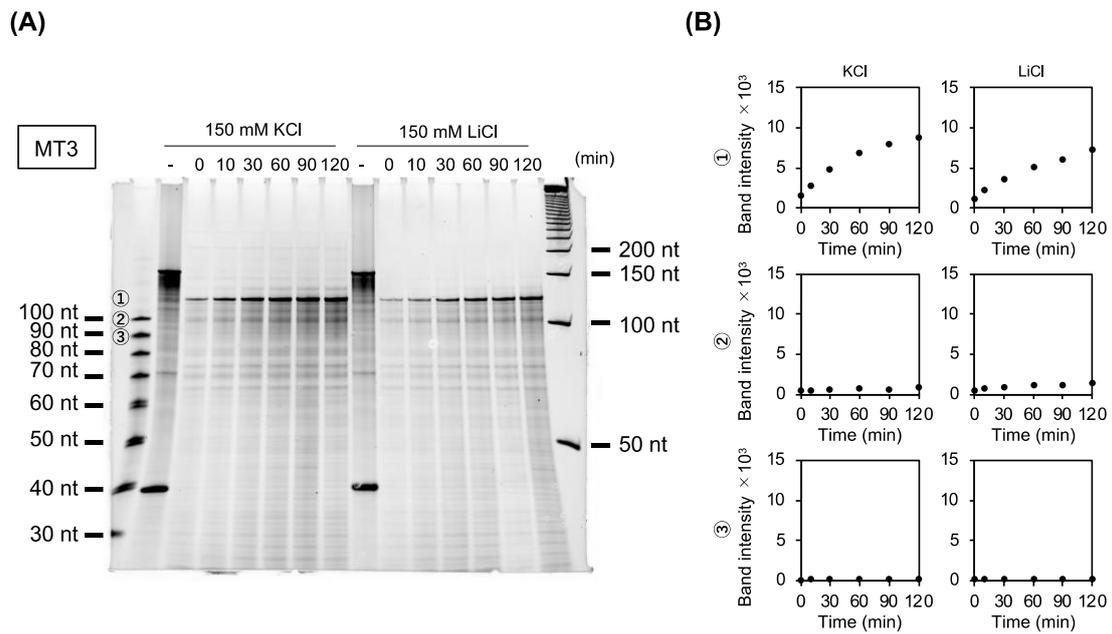
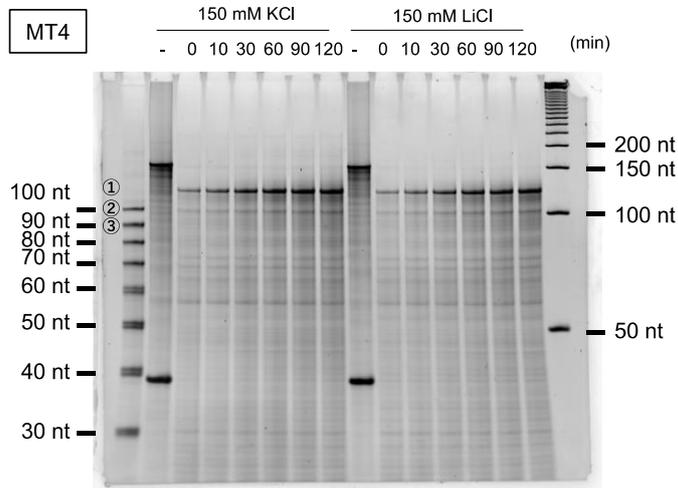


Fig. S12 (A) Transcripts from Temp-MT3 after a reaction time of 10, 30, 60, 90 or 120 min at 37 °C. The transcription reactions were carried out in the presence of  $K^+$  (left) or  $Li^+$  (right). (B) Intensities of the three main bands after various reaction times in the presence of  $K^+$  (left) or  $Li^+$  (right).

Figure S13

(A)



(B)

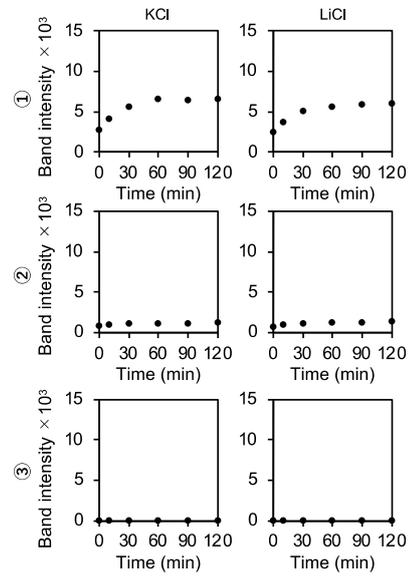


Fig. S13 (A) Transcripts from Temp-MT4 after a reaction time of 10, 30, 60, 90 or 120 min at 37 °C. The transcription reactions were carried out in the presence of  $K^+$  (left) or  $Li^+$  (right). (B) Intensities of the three main bands after various reaction times in the presence of  $K^+$  (left) or  $Li^+$  (right).

Figure S14

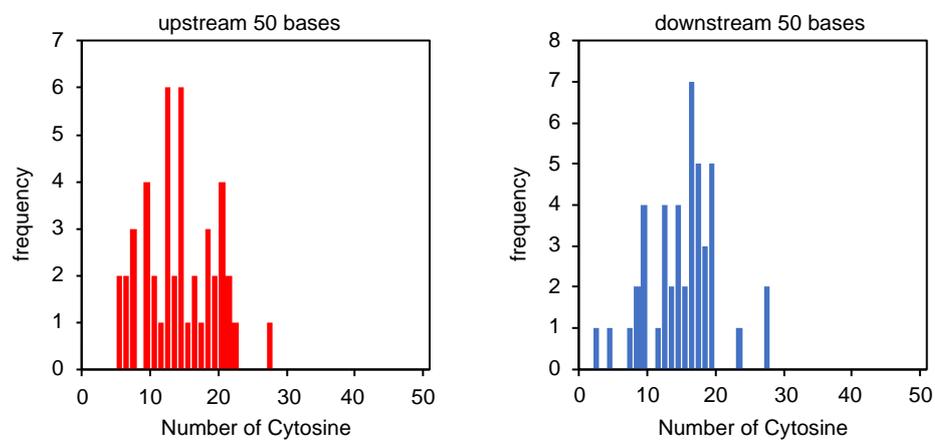


Fig. S14 The numbers of cytosine bases in the upstream (A) and in the downstream (B) 50 bases region from the putative quadruplex G4-forming sequence observed in the 62 human proteases.

## References

1. N. Sigimoto, M. Nakano, S. Nakano, *Biochemistry* 2000, 39, 11270-11281.
2. M. Zuker, *Nucleic Acids Res.*, 2003, 31, 3406-3415.