Supplementary Data for

The Origin of Genetic Instability in CCTG Repeats

Sik Lok Lam*, Feng Wu, Hao Yang and Lai Man Chi

Department of Chemistry, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong

* To whom correspondence should be addressed. Phone: (852) 2609-8126. Fax: (852) 2603-5057.

E-mail: lams@cuhk.edu.hk

This file includes Supplementary Figures S1 to S8.

Figure S1. NMR features of the stem region of $(CCTG)_3$ hairpin. (**A**) 2D WATERGATE-NOESY spectrum shows a NOE between G8 imino and C5 amino at 0 °C, confirming the formation of CT-loop with a G8·C5 closing base pair. (**B**) 2D NOESY H6/H8–H1' fingerprint region of $(CCTG)_3$ at 15 °C shows the signals of C9 and C10 are broadened while the signals of G8, T11 and G12 remain sharp, supporting the presence of C9- and C10-bulge conformers. (**C**) The two possible pairing modes of T·T mismatch.

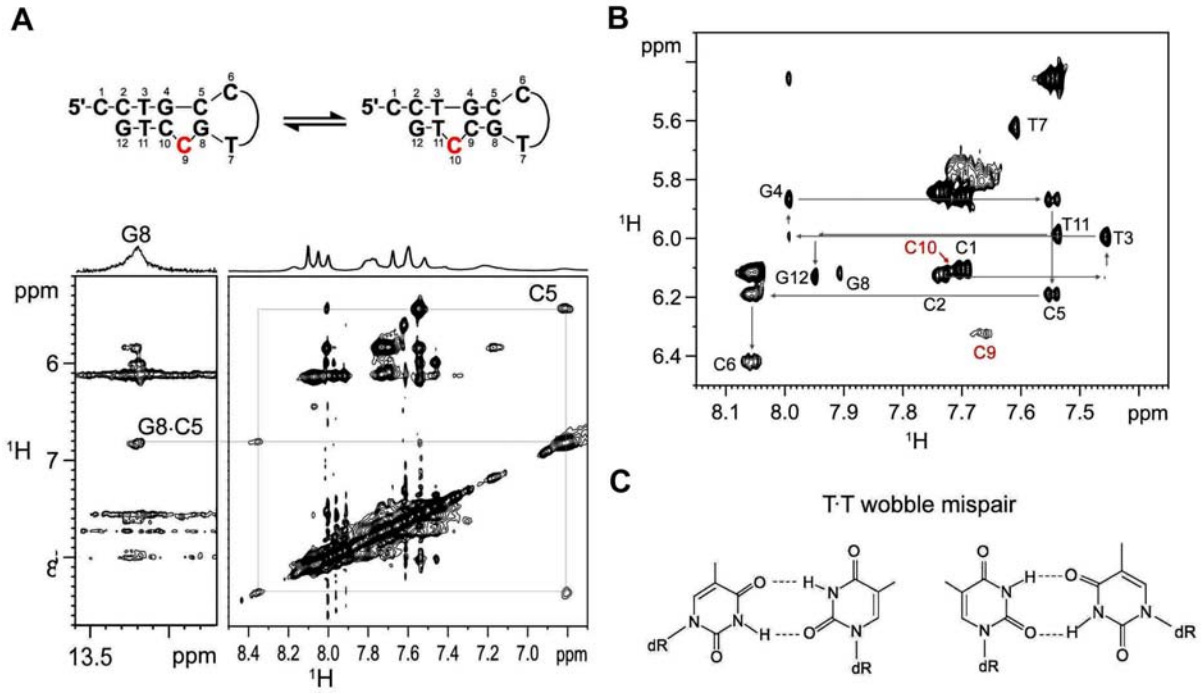

Figure S2. Effect of adding a G residue to the 3'-end of $(CCTG)_3$. (**A**) The hairpin structures were stabilized after the addition of a G. (**B**) 2D WATERGATE-NOESY imino-amino and amino-amino regions of $(CCTG)_3G$ at 0 °C. (**C**) The broadening of the C10 H5-H6 TOCSY cross peak at 15 °C was less serious after the 3'-end was stabilized, revealing the C9-bulge conformer became predominant. (**D**) At 0 °C, the T3 and T11 imino signals appeared at ~10.9 and 11.3 ppm due to the reduced dynamics of the C-bulge. (**E**) 2D NOESY H6/H8–H1' fingerprint region of $(CCTG)_3G$ at 25 °C.

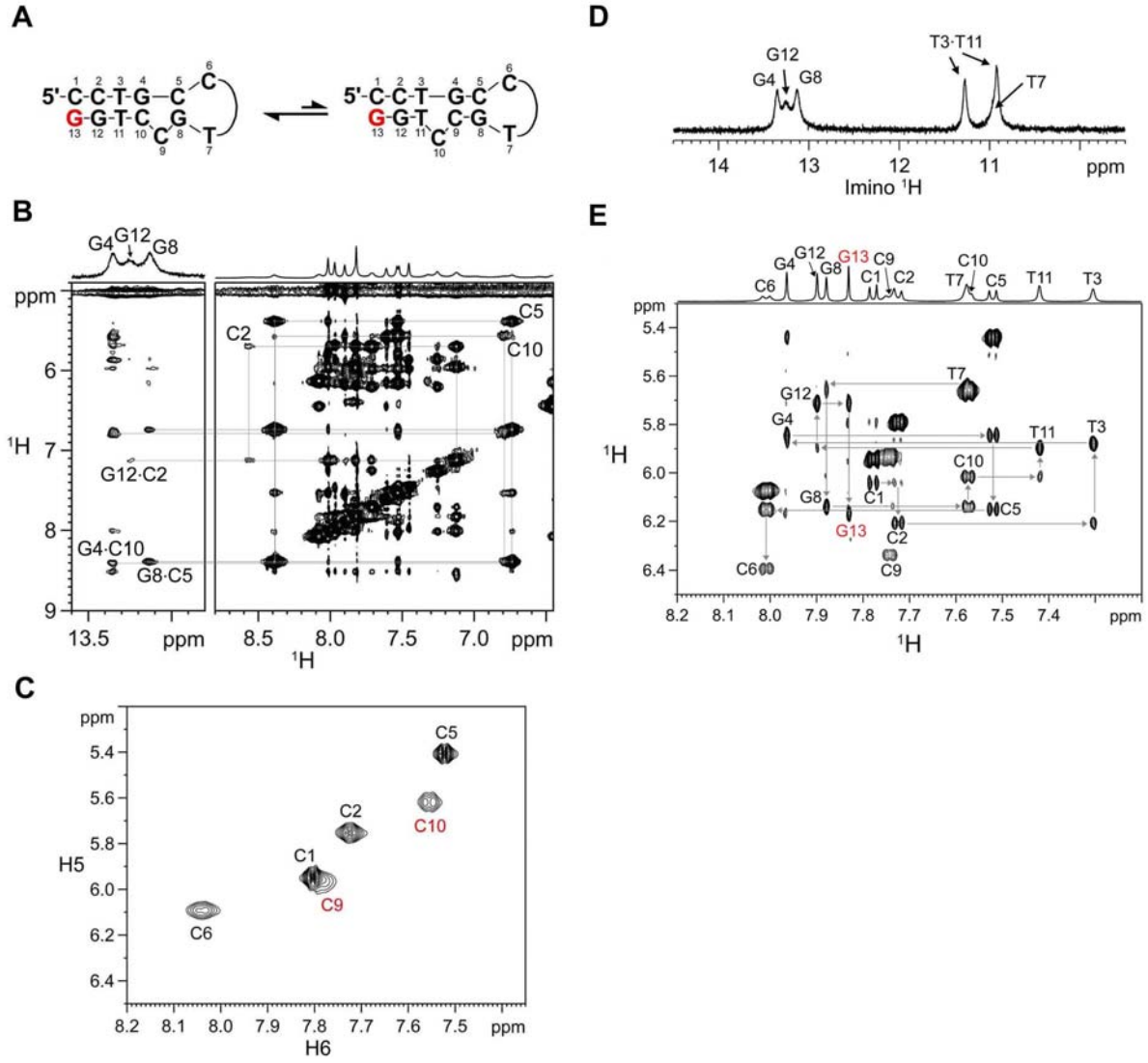

Figure S3. No TT-loop was formed in $(CCTG)_4$-C2T, suggesting the first repeat of $(CCTG)_4$ is not involved in forming CT-loop. (**A**) Variable temperature ^{1}H NMR spectra showing the imino and thymine methyl ^{1}H regions. The intensity of the imino region has been scaled up eight times. No characteristic methyl peak of TT-loop was observed. (**B**) Variable temperature ^{31}P spectra showing characteristic peaks of CT-loops. (**C**) Conformational exchange was evidenced by ROESY exchange cross peaks (in black) at 0 °C. Positive and negative contours were plotted in black and red, respectively. (**D**) The exchange process was also supported by ^{31}P–^{31}P EXSY cross peaks at 0 °C.

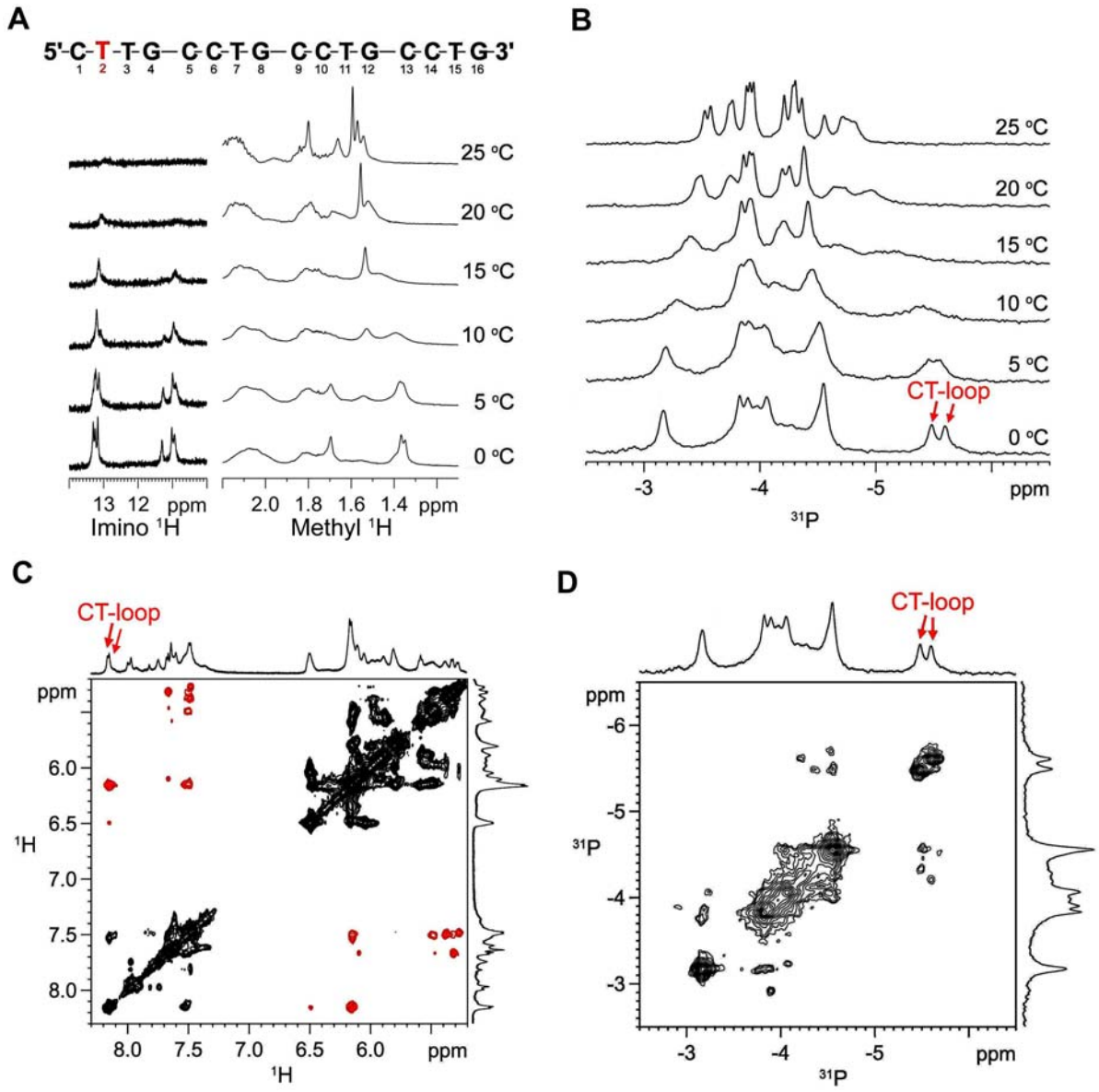

Figure S4. $(CCTG)_4$-C6T forms a stable dumbbell hairpin with a TT- and a CT-loop. (**A**) No exchange cross-peak was observed in 2D ^{31}P–^{31}P EXSY spectrum at 0 °C. The characteristic ^{31}P signals show the formation of a stable dumbbell containing a TT- and a CT-loop. (**B**) 2D NOESY H6/H8–H1′ fingerprint region of $(CCTG)_4$-C6T at 25 °C. The arrows show the sequential NOE connectivities. The G16−C2 NOE (in red) reveals the presence of intra-molecular 3′−5′ terminal stacking interaction. (**C**) Variable temperature thymine methyl ^1H spectra. (**D**) Variable temperature ^{31}P spectra. (**E**) 2D WATERGATE-NOESY spectrum at 0 °C shows the presence of G8·C5, G16·C13 base pairs, supporting the formation of a TT- and CT-loop, respectively. (**F**) Imino−imino region shows the NOEs between G8 and G4. (**G**) TOCSY spectrum shows cytosine H5−H6 cross peaks at 25 °C. (**H**) The imino signals of the loop residues T7 and T15 appeared at ~11 ppm at 0 °C, overlapping with that of T3·T11 mispair.

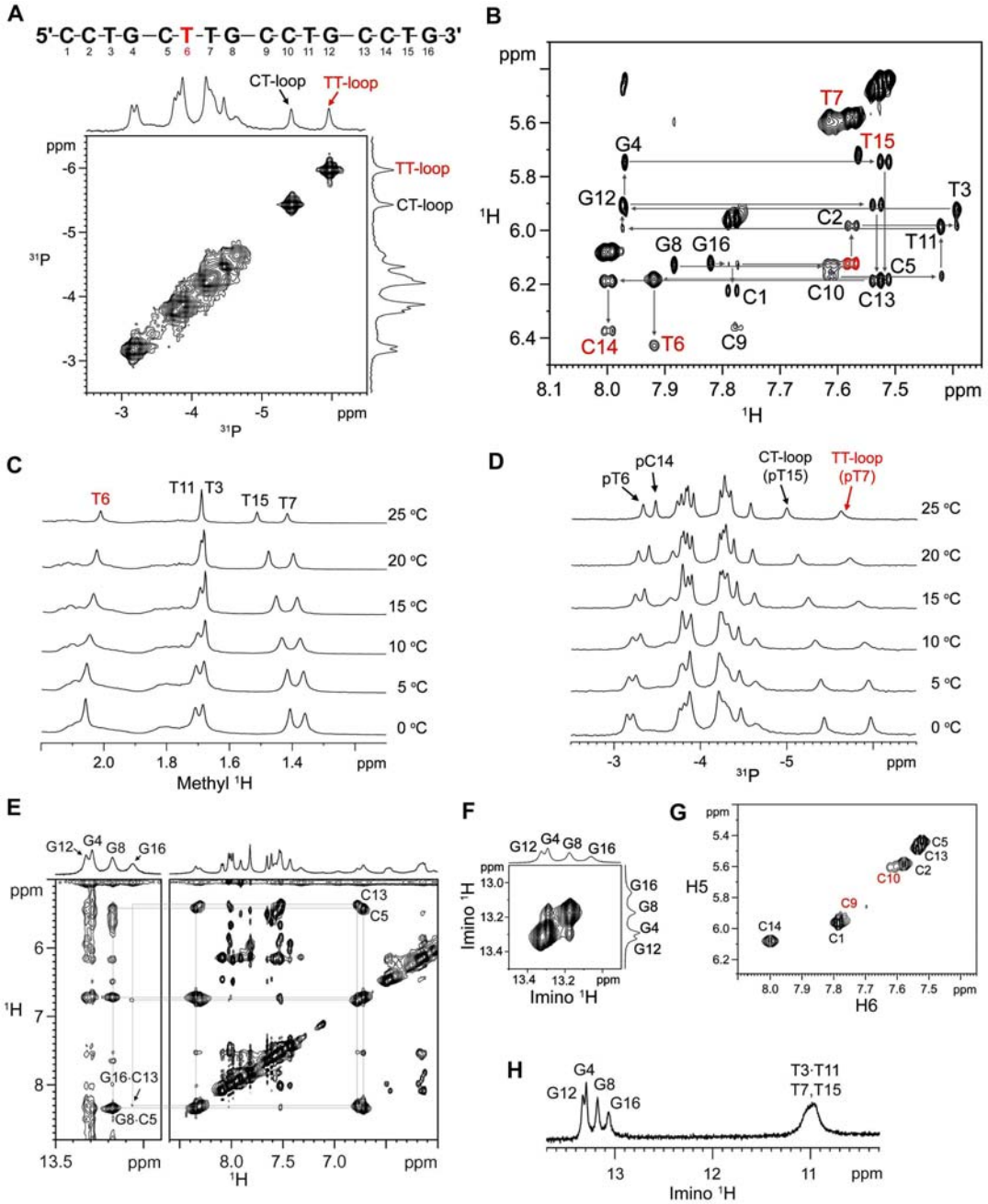

Figure S5. $(CCTG)_4$-C10T adopts both hairpin and dumbbell conformations. (**A**) The formation of a minor conformer with a TT-loop was evidenced by an enormously upfield small ^{31}P signal at ~-6.0 ppm. (**B**) Variable temperature ^{1}H NMR spectra showing the imino and thymine methyl ^{1}H regions. The intensity of the imino region has been scaled up eight times. A downfield thymine methyl ^{1}H signal at ~2.08 ppm appeared at 0 °C also supports the TT-loop formation. (**C**) Two intense downfield cytosine H5−H6 TOCSY signals were observed at 0 °C, suggesting the major conformer contains two CT-loops. (**D**) 2D ^{31}P–^{31}P EXSY spectrum at 0 °C shows exchange cross peaks. (**E**) Conformational exchange was evidenced by 2D ROESY exchange cross peaks (in black) at 0 °C. Positive and negative contours were plotted in black and red, respectively. (**F**) The thymine methyl protons also show ROESY exchange cross peaks at 0 °C.

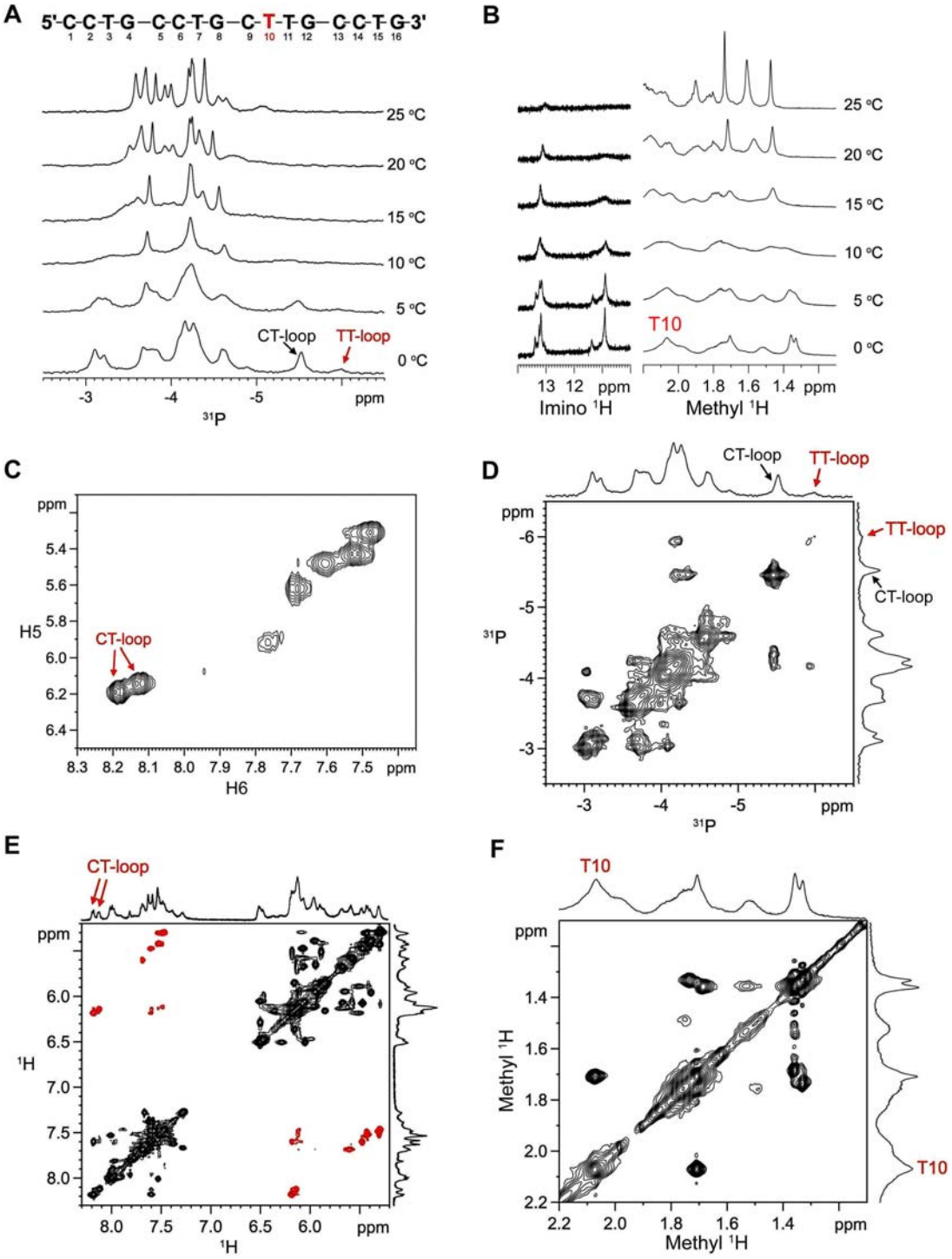

Figure S6. $(CCTG)_4$-C14T forms a stable dumbbell hairpin with a TT- and a CT-loop. (**A**) Variable temperature ^{1}H NMR spectra showing the thymine methyl ^{1}H region. T14 has a characteristic downfield signal of TT-loop at 2.02 ppm. (**B**) Variable temperature ^{31}P NMR spectra showing the characteristic ^{31}P signals of CT- and TT-loops. (**C**) No exchange cross peak was observed in 2D ^{31}P–^{31}P EXSY spectrum at 0 °C. (**D**) 2D NOESY H6/H8–H1′ fingerprint region at 25 °C shows the sequential NOEs. The G16−C2 NOE (in red) reveals the presence of intra-molecular 3′−5′ terminal stacking interaction. (**E**) TOCSY spectrum shows cytosine H5−H6 cross peaks at 25 °C. (**F**) The imino signals of the loop residues T7 and T15 appeared at ~11 ppm, overlapping with those of T3·T11 mispair at 0 °C.

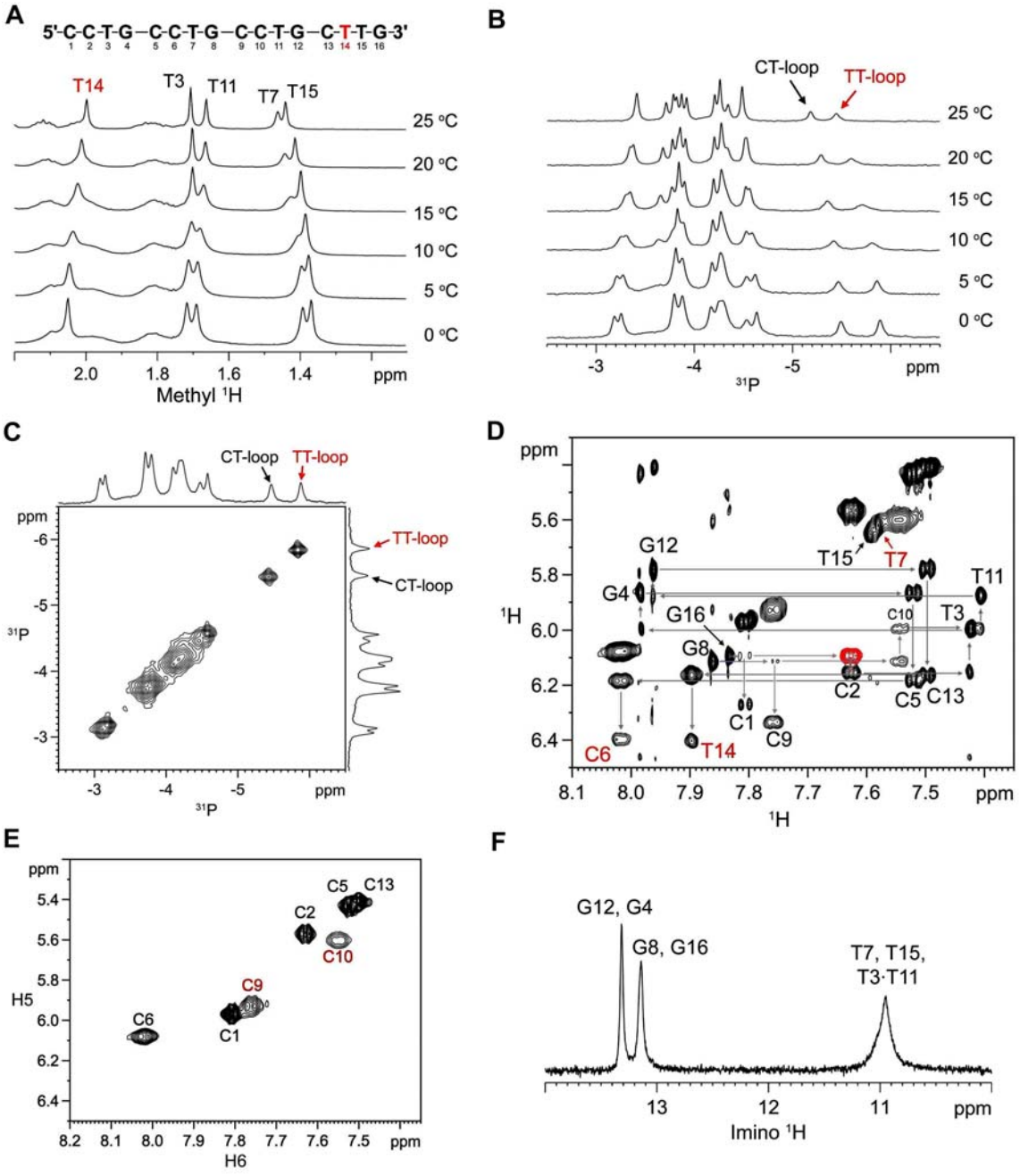

Figure S7. $(CCTG)_5$-C18T possesses NMR characteristics of a TT- and a CT-loop. (**A**) Variable temperature ^{1}H NMR spectra showing the imno and thymine methyl ^{1}H regions. The intensity of the imino region has been scaled up eight times. T18 has a characteristic downfield signal of TT-loop at ~2.05 ppm. (**B**) TOCSY spectrum shows cytosine H5−H6 cross peaks at 0 °C. (**C**) Variable temperature ^{31}P NMR spectra showing the characteristic ^{31}P signals of CT- and TT-loops. (**D**) Exchange cross peaks were observed in 2D ^{31}P–^{31}P EXSY spectrum at 0 °C.

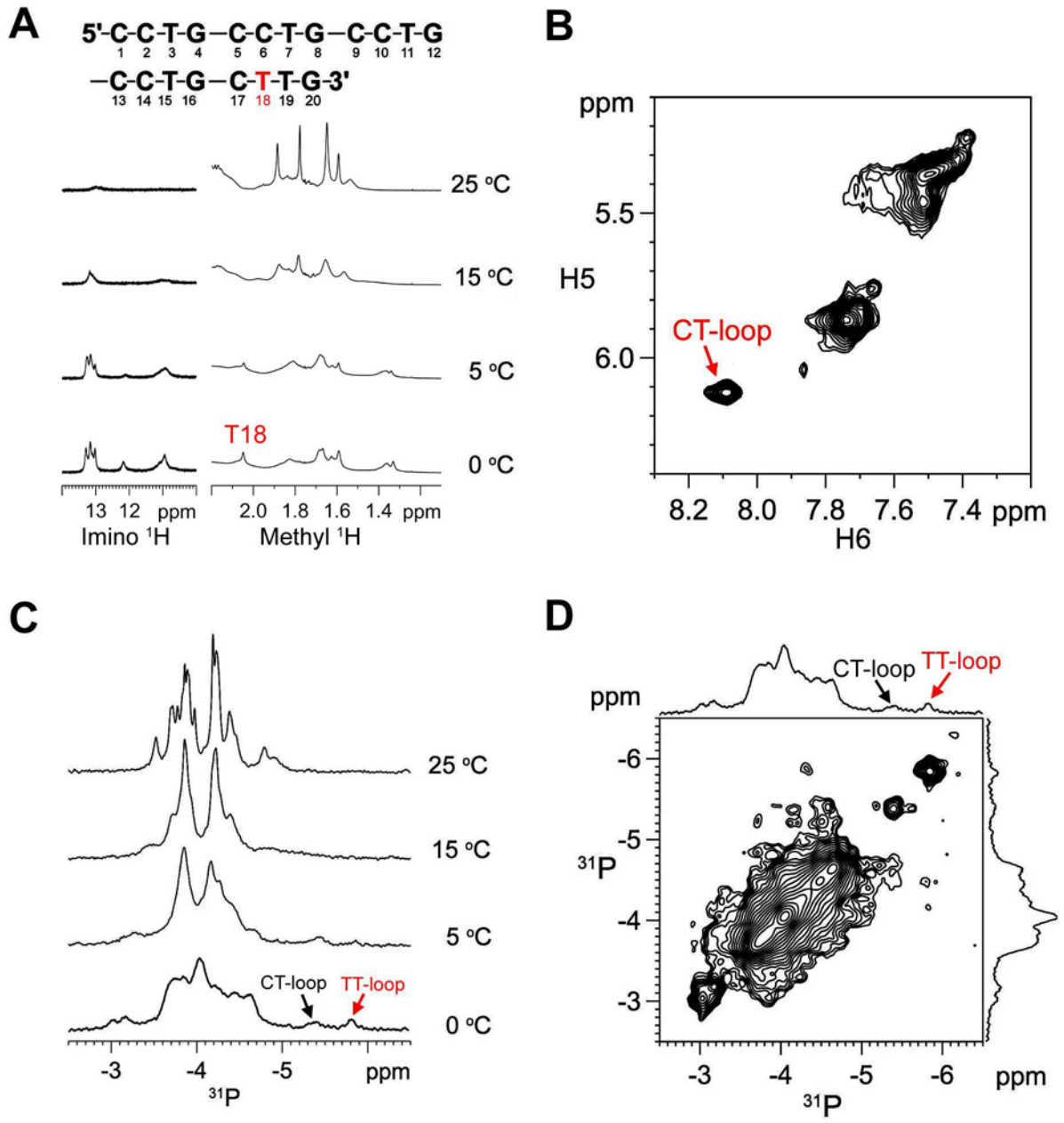

Figure S8. $(CCTG)_6$-C22T possesses NMR characteristics of a TT- and a CT-loop. (**A**) Variable temperature ^{1}H NMR spectra showing the imino and thymine methyl ^{1}H regions. The intensity of the imino region has been scaled up eight times. T22 has a characteristic downfield signal of TT-loop at ~2.06 ppm. (**B**) TOCSY spectrum shows cytosine H5−H6 cross peaks at 0 °C. (**C**) Variable temperature ^{31}P NMR spectra showing the characteristic ^{31}P signals of CT- and TT-loops. (**D**) Exchange cross peaks were observed in 2D ^{31}P–^{31}P EXSY spectrum at 10 °C.